

MAGYAR NYELVŐR

JOURNAL OF THE INTERNATIONAL SOCIETY
OF HUNGARIAN LANGUAGE AND CULTURE

ARTIFICIAL INTELLIGENCE AND LANGUAGE: INTERDISCIPLINARY PERSPECTIVES



György Csepeli: The silent province

Agnes Czine: Words, phrases, thoughts without heart and soul...

Geoff Vaughan – Rita Szűts-Novák: What we talk about when we talk about AI. Artificial intelligence and language

László Z. Karvalics: Artificial intelligence and language – A different approach to the problem

Balázs Arató – Géza Balázs: The communicative-linguistic modes of artificial intelligence with a focus on justice

Lajos Toldi – Tünde Lengyel-Molnár: Development options for artificial intelligence supported intelligent tutoring systems via the integration of the ChatGPT

Nikoletta Tolner – Monika Pogátsnik: The large language models in the service of education

Lilla Király: Artificial intelligence as a virtual language teacher

Géza Balázs – Ottó Beke – Zoltán Szűts: ChatGPT language recognition limitations. Prompts in Hungarian in a network of Latin and Cyrillic letters. (Netlinguistics case study. The “Chinese room” experiment)

Zoltán Trenka – Viktória Varsányiné Szeredi – Monika Pogátsnik: The impact and potential of artificial intelligence in language education

Renátó Vági – István Üveges: Laws Clearly: Large language models and plain language transformation

János Tamás Papp: Navigating the Digital Contract Maze: How AI can guide users to informed decisions

Balázs Arató – Tamás Sajben: The connection points between smart contracts and artificial intelligence – Opportunities and challenges

Benjamin Szilágyi – Rozália Lakner: NLP-enhanced classification of remote employment opportunities

Béla Pokol: The layers of being and the questions of robot ethics

MAGYAR NYELVŐR – JOURNAL OF THE INTERNATIONAL SOCIETY OF HUNGARIAN LANGUAGE AND CULTURE

Founded in 1872

Published quarterly (in January, April, July and December) in printed and online versions

Double-blind peer-reviewed open-access journal

Editors: Szarvas, Gábor (1872–1895), Simonyi, Zsigmond (1896–1919),
Balassa, József (1920–1940), Beke, Ödön (1946–1953),
Lőrincze, Lajos (1954–1993), Keszler, Borbála (1994–2021)

Special Issue: Artificial Intelligence and Language: Interdisciplinary Perspectives (2024-5)

Editor-in-Chief (2022):

Prof. dr. Balázs, Géza, e-mail: balazs.geza@szfe.hu

Special Issue Editors:

Arató, Balázs PhD, e-mail: arato.balazs@kre.hu

Piglerné Lakner, Rozália PhD, e-mail: lakner.rozalia@amk.uni-obuda.hu

Szűts, Zoltán PhD, e-mail: szuts.zoltan@uni-eszterhazy.hu

Linguistic Proofreader: Vesszős, Balázs PhD

Editing and image: Fancsek, Krisztina

Cover photo: Artificial intelligence and language (The image was generated by OpenAI software.)

Press: Könyvpont Nyomda, Budapest

The address of the Editorial Staff: 38. Rákóczi Street, H-1072 Budapest, Hungary

The email of the Editorial Staff: magyarnyelvor@mnyknt.hu

The direct email of the Editor-in-Chief: balazs.geza@szfe.hu

Website, archive: www.mnyknt.hu



Publisher: The International Society of Hungarian Language and Culture

Chief Publisher: Dr. Arató, Balázs

Address: 38 Rákóczi Street, 1072 Budapest, Hungary

Phone: 36-30-318-9666, e-mail: irodavezeto@mnyknt.hu

ISSN (printed volume) 0025-0236

ISSN (online volume) 1585-4515

Subscription: from 2022 exclusively at the publisher's office, the International
Society of Hungarian Language and Culture.

The price of each volume: HUF 1500, annual subscription (4 volumes): HUF 6000.

We charge postage abroad.

Európa elrablása

Történelem, kultúra, média, jog, retorika

3. Zamárdi szabadegyetem
2024. szeptember 20–22.

Hethland szálloda
Zamárdi, Kiss Ernő utca 3.

2024. szeptember 20. péntek

Szakály Sándor történész: Trianontól Párizsig

2024. szeptember 21. szombat

Gózon Ákos újságíró, szerkesztő: Európai tükrök

G. Etényi Nóra történész: Európa-szimbólumok a kora újkorban

Csepeli György szociálpszichológus: Útvesztés- útkeresés

Kövecses Zoltán nyelvész: A szeretet fogalma
– kognitív nyelvészeti megközelítés

Korzenszky Richárd, a Tihanyi Apátság volt előjárója:
Európa - értékrend nélkül? Anyanyelv, otthon, haza

2024. szeptember 22. vasárnap

Balázs Géza nyelvész: A Kelet és a Nyugat

Csicsely Tamás mérnök: Európa elrablása a Gottscheer kápolna
és a máriacelli kegytemplom között

Paár Ádám politológus, történész: A népi írók és Európa

Az előadások nyilvánosak,
ingyenesen látogathatók.

Frissülő információk:
www.mnyknt.hu



PI PETŐFI
KULTURÁLIS
ÖSZÖNKÖDÉS



Bolyai
Onképző Műhely



Magyar Nyelv és Kultúra
Nemzetközi Társasága

MAGYAR NYELVÖR – JOURNAL OF THE INTERNATIONAL SOCIETY
OF HUNGARIAN LANGUAGE AND CULTURE
SPECIAL ISSUE: ARTIFICIAL INTELLIGENCE AND LANGUAGE:
INTERDISCIPLINARY PERSPECTIVES

CONTENTS

<i>György Csepei</i> : The silent province	565
<i>Ágnes Czine</i> : Words, phrases, thoughts without heart and soul... ..	570
<i>Geoff Vaughan – Rita Szűts-Novák</i> : What we talk about when we talk about AI. Artificial intelligence and language	579
<i>László Z. Karvalics</i> : Artificial intelligence and language – A different approach to the problem	589
<i>Balázs Arató – Géza Balázs</i> : The communicative-linguistic modes of artificial intelligence with a focus on justice	601
<i>Lajos Toldi – Tünde Lengyel-Molnár</i> : Development options for artificial intelligence supported intelligent tutoring systems via the integration of the ChatGPT	618
<i>Nikoletta Tolner – Monika Pogátsnik</i> : The large language models in the service of education	629
<i>Lilla Király</i> : Artificial intelligence as a virtual language teacher	649
<i>Géza Balázs – Ottó Beke – Zoltán Szűts</i> : ChatGPT language recognition limitations. Prompts in Hungarian in a network of Latin and Cyrillic letters. (Netlinguistics case study. The “Chinese room” experiment)	666
<i>Zoltán Trenka – Viktória Varsányiné Szeredi – Monika Pogátsnik</i> : The impact and potential of artificial intelligence in language education	682
<i>Renátó Vági – István Üveges</i> : Laws Clearly: Large language models and plain language transformation	696
<i>János Tamás Papp</i> : Navigating the Digital Contract Maze: How AI can guide users to informed decisions	717
<i>Balázs Arató – Tamás Sajben</i> : The connection points between smart contracts and artificial intelligence – Opportunities and challenges	728
<i>Benjamin Szilágyi – Rozália Lakner</i> : NLP-enhanced classification of remote employment opportunities	747
<i>Béla Pokol</i> : The layers of being and the questions of robot ethics	761

MAGYAR NYELVÖR

Journal of the International Society of Hungarian Language and Culture

VOLUME 148

*

2024

*

NUMBER 5 (SPECIAL ISSUE)

GYÖRGY CSEPELI

THE SILENT PROVINCE

Abstract

Digitalisation, datafication and networking have profoundly changed the spaces of human drama on Earth, and the linguistic space is no exception. ChatGPT speaks in the language of the average speaker, referred to by Heidegger as “das Man”, incapable of innovation or discovery, in contrast to the multitude of unspoken sentences waiting to be uttered, with the excitement of discovery, novelty, and adventure. Yet, behind the unspoken but pronounceable utterances, there is a silent realm whose events are the unspeakable.

Keywords: space, generative language models, Heidegger, Wittgenstein, silence, linguistic innovation, silent domain

Introduction

John Maynard Smith and Eörs Szathmáry correctly state, that: “It is impossible to imagine our society without language” (Smith–Szathmáry 1999: 149). Moreover, we should add that imagining human beings without language is impossible. Language elevates man from the constraints of the eternal “here and now”, which rules in the animal kingdom. Deprived by the fixed environment (Umwelt) set by genes, human beings, by the power of language, can create their open environment (Welt) where past, present, future, and here and there exist. Reality by linguistic means becomes constructed and shared. Heidegger realised this, stating that “language is the house of being. In its home Man dwells” (Heidegger 1978: 217).

Smith and Szathmáry treat language as the last significant inheritance transition system evolving into one with potentially unlimited heredity. In

their somewhat limited approach, the authors fail to understand that language is more than a means of transmitting information. Nevertheless, they are right to emphasise two significant changes in how data has been transmitted since the emergence of language in human life. The first significant change was the invention of writing, which made the appearance of large-scale civilisations possible. The authors wrote in 1999, that the second change, “through which we are living today, is the use of electronic means for storing and transmitting information.” They were right, assuming that the effects of this change “will be as profound as were those after the origin of the genetic code, or of language, but we are not rash enough to predict what they will be (Smith–Szathmáry 1999: 179).

The changes brought about by digitisation, datafication, and networking reaching all corners of the world have fulfilled Smith and Szathmáry’s prophecy. Digital transformation has radically changed almost every aspect of people’s lives. The latest wave of change reached the language itself.

On November 30, 2022, OpenAI launched a chatbot using statistical linguistic patterns to extrapolate and generate new texts prompted by user requests. By January 2023, the magic device had gained over 100 million users, each contributing to developing software constantly enriched by each new prompt.

Generative language models, which are constantly being expanded and improved by users, are based on sentences previously written and spoken by the users of a given language, which Chomsky identifies as performance for individual language users, as distinct from competence, which is the condition for the creation of new sentences (Chomsky 1957). Generative language models are based on the collective performative performance of the users of a given language but lack the constant readiness to innovate, surprise, and create, which is characteristic of competence.

The creator and beneficiary of the collective linguistic performance embodied in generative language models is the faceless collective being that Heidegger calls *das Man* “who is not this one, not that one, not oneself, not some people and not the sum of them all. The ‘who’ is the neuter, the ‘they’”. In contrast, the “real others, as distinguishable and explicit, vanish more and more. In this inconspicuousness and ascertainability, the real dictatorship of the ‘they’ is unfolded. We take pleasure and enjoy ourselves as *they* take pleasure; we read, see, and judge literature and art as they see and judge. Likewise, we shrink back from the ‘great mass’ as they shrink back, we find ‘shocking’ what they find shocking” (Heidegger 1962: 164).

Being familiar and accessible to everyone, the sentences of ChatGPT produced by the prompts never get to the heart of the matter; the sentences are

completely insensitive to genuineness. The language of ChatGPT never goes beyond the averageness “with which it prescribes what can and may be ventured, it keeps watch over everything exceptional and thrusts itself into the fore. Every kind of priority gets noiselessly suppressed” (Heidegger 1962: 165).

The language of ChatGP, undoubtedly useful in everyday routine cognitive operations, is counterposed to a living language, which selects from an inexhaustible wealth of competence, and constantly expands Man’s scope for adventure in search of sentences never uttered before (Hankiss 2014).

Wittgenstein famously said, “Whereof one cannot speak, thereof one must be silent” (Wittgenstein 1921: 5). In the context of the silence required by Wittgenstein, Géza Balázs raises the point that in the silent domain, speakers must also struggle with the challenges of sentences not yet uttered, but which can be uttered (Balázs 2020).

The silent domain can be divided into two halves. In one-half of the domain are the utterable sentences, which, sooner or later, when uttered, get entered into the database of generative language models. In the other, more exciting half of the domain, there are the unutterable contents, which are impossible to force into the language machinery. At the same time, their existence is implicitly indicated by linguistic manifestations whose code needs to be deciphered.

Nietzsche would claim that the house of existence has two inhabitants, one being Apollo and the other Dionysus. The house of existence consists not only of the living quarters inhabited by the god Apollo, which can be imagined and represented by language. In the house of existence are the silent quarters of Dionysus outside of language, where individuals and communities have their secrets, sins, and fears.

The appropriate procedure for addressing the contents of the mind, which are banished to the depths of the individual consciousness and cannot be expressed in sentences, is psychoanalysis and its predecessor, confession. The unspeakable contents locked up in the unconscious of communities can break out of the silent realm through the mediation of myths and rituals. A unique example is the occasion of the meeting between Paul Celan and Martin Heidegger on 24 July 1967. Celan read his poetry in the main auditorium of the University of Freiburg, where Martin Heidegger, who was rector of the university in 1933, was seated in the front row.

The poet and the philosopher mutually admired each other, but apparently could not cope with the fact that fate had led them on divergent paths in the past. Heidegger was an active supporter of Hitler’s National Socialist dictatorship, which persecuted Celan because of his Jewish origins. Heidegger had been writing his monumental philosophical works in peaceful conditions

in Freiburg. Meanwhile, the poet, who lived in Bukovina, was forced into a ghetto after the German troops had entered in 1941 and deported to a labour camp after the ghetto was eliminated. In the meantime, the parents of the poet were deported to an internment camp, where Celan's father died of typhus and his mother was shot. Celan remained imprisoned until February 1944.

At Heidegger's invitation the next day, Celan went up the mountain, where they talked for an hour and a half at Heidegger's farmhouse, between the hut and the well. There were no witnesses to the conversation, but according to the participants' retrospective accounts, it seems that their sentences were only meant to be spoken, and they talked about the landscape, the birds, the well, and the trees in the garden. However, they were silent about the past (Bíró 2000).

Following the visit, the poet wrote these words in the philosopher's guest book: "In the hut's book, glancing towards the well's star, in the hope a word to come". Later, the poet wrote a poem about the visit, a poem of poignant beauty. In the poem, the poet reiterated the hope that the thinker would utter the word he had been unable to utter on 25 July 1967 in Freiburg.

Celan had been waiting in vain. Not even three years after their meeting, he committed suicide by throwing himself into the Seine in Paris on 20 April 1970. He did not know that on 30 January 1968, 35 years to the day after Hitler came to power, Heidegger had written him an unsent letter, a letter in which he did not, however, write the words the poet had been waiting for. The philosopher remained a prisoner of the silent province.

Celan's poem *Todtnauberg* in German and in English

Arnika, Augentrost, der
Trunk aus dem Brunnen mit dem
Sternwürfel drauf,

in der
Hütte,

die in das Buch
– wessen Namen nahm's auf
vor dem meinen? –
die in dies Buch
geschriebene Zeile von
einer Hoffnung, heute,
auf eines Denkenden
kommendes
Wort
im Herzen,

Arnica, eyebright, the
draft from the well with the
star-die on top,

in the
Hütte,

written in the book
– whose name did it record
before mine – ?
in this book
the line about
a hope, today,
for a thinker's
word
to come,
in the heart,

Waldwasen, uneingeebnet, Orchis und Orchis, einzeln,	forest sward, unleveled, orchis and orchis, singly,
Krudes, später, im Fahren deutlich,	crudeness, later, while driving, clearly,
der uns fährt, der Mensch, der's mit anhört,	he who drives us, the man, he who also hears it,
die halb- beschrittenen Knüppel- pfade im Hochmoor,	the half- trod log- trails on the highmoor,
Feuchtes, viel.	humidity, much.

[URL1]

References

- Balázs, Géza 2020. *Nyelvszokások. A nyelvszokások mint szimbolikus nyelvi cselekvések*. Írások az antropológiai nyelvészet köréből. Budapest. Inter.
- Bíró, Zsuzsa 2000. Hazatérés? *Liget* 8. <https://ligetmuhely.com/liget/hazateres/>
- Chomsky, Noam 1957. *Syntactic Structures*. Hague. Mouton.
<https://doi.org/10.1515/9783112316009>
- Hankiss, Elemér 2014. *Az emberi kaland*. Budapest. Helikon.
- Heidegger, Martin 1962. *Being and Time*. Hoboken. Blackwell Publishing.
- Heidegger, Martin 1978. Letter on Humanism. In: D. F. Krell (ed.): *Basic Writings*. London. Routledge. 213–65.
- Smith, John Maynard – Szathmáry, Eörs 1999. *The Origins of Life. From the Birth of Life to the Origins of Language*. London. Oxford University Press.
- Wittgenstein, Ludwig 1921. *Tractatus Logico-Philosophicus*. Oxford. Routledge.

[URL1] <https://writing.upenn.edu/epc/authors/joris/todtnauberg.html>

György Csepeli

Professor Emeritus

Eötvös Loránd University

Faculty of Social Sciences

E-mail: csepeli.gyorgy@gmail.com

<https://orcid.org/0000-0002-9331-058X>

ÁGNES CZINE

WORDS, PHRASES, THOUGHTS WITHOUT HEART AND SOUL...

Abstract

Artificial intelligence now permeates almost all of our lives in some form or another. The title also discusses the impact of the emergence of artificial intelligence on legal language. On the other hand, whether AI is capable of interpreting legal concepts that are not mere terms but have a number of underlying meanings.

Keywords: Language and AI, legal language and artificial intelligence, AI, our mother tongue, our common national treasure, the relevant legal interpretation.

1. Language and AI

When Isaac Asimov formulated the three laws of robotics in his 1941 novel *Runaround*, he could foresee that a world was imminent in which we would have to live with artificial intelligence, its technical vehicles and all the moral and legal issues that would arise from it. We are now at a point where the European Parliament on 13 March 2024 and the Council of the EU on 21 May 2024 (URL1) have approved a regulation, the artificial intelligence Act [2]. This Act aims to establish a minimum level of regulation, with different categories of risk for each AI system (hereinafter referred to as “artificial intelligence”) and applies to all such systems placed on the market or used in the EU (URL2).

There is therefore a guideline for EU Member States to act upon, but as lawyers we cannot sit back complacently because the work is only just beginning. For now, there are more questions than answers in the legislation. There are many new legal questions about self-directing, self-improving, decision-making programmes modelled on human beings and their physical embodiment, robots. Each branch of law will have to think through how to create guarantees and how to adjudicate the legal liability associated with the decisions taken by self-improving machines, learning algorithms. For example, how to assess statements made by AI in terms of freedom of expression, civil contractual relations, criminal liability. What about the copyright of a work written, painted or created by a machine? Can such an entity inherit, have legal capacity or limited legal capacity? Can it be put on the right moral path by punishment, can it recognise the consequences of its actions or is the

only sanction for it to be unplugged? Thousands of detailed questions arise when we consider the rights of the human being and the future entitlements of the AI in relation to those rights, to which legal experts are seeking answers in the different legal systems of different countries.

This of course requires not only the development of legislation but also of legal language. The term of artificial intelligence has been translated and the Hungarian abbreviation MI has become a part of the Hungarian legal languages, but what do we do with a concept such as deepfake? Do we need to create a Hungarian equivalent of this word, replacing it with the Hungarian equivalent, which has a different connotation and different associations? Do we simply adopt the mostly English-derived terms without changing them or perhaps “explaining” them? Especially in the light of what follows.

We asked the best expert, the ChatGPT itself, about the impact of the development of artificial intelligence on legal language. Its short response was that artificial intelligence (AI) is having a significant impact on the development and use of legal language in a number of ways. It increases the accuracy, efficiency and effectiveness of legal language by automating routine tasks, improving access to information and providing tools to better understand and apply legal concepts. This transformation will not only benefit legal professionals, but will also make legal services more accessible and understandable to the general public. This can also serve to meet the requirement of norm clarity (Czine 2022; Arató 2022; Arató–Balázs 2022) [3].

Standardisation, automation, accuracy, handling information dumping, accessibility, speed of information flow are the key words that characterise the benefits that AI can bring to the development of legal language. This could be a significant advantage, since it enables us to know and understand in a matter of seconds the legislation that applies to a legal relationship in a foreign environment, even in a distant country. With a click of a mouse and without any ambiguity, legal experts in different jurisdictions can clarify legal issues between themselves with great precision. However, there may be, who or which argue with each other will no longer be human beings, but legal AI systems?

Well, the answer is not so simple, because we know that the interpretation of the law by judges, legal practitioners and legislative amendments behind legal concepts can and do constantly shape the content of legal concepts. When the legislator amends the law, the legal norm may also acquire a new content as a result. This is also a continuous process in judicial jurisprudence, because judicial application of the law interprets and applies different legal norms on a daily basis and could slightly modified their contents. For an

algorithm [URL3] it is necessary to constantly monitor the relevant legal interpretation and language rules related to the solution of the given task. It is necessary to follow the changes in language, concepts that are developed by the legislator and the law enforcer.

1.1. To approach the impact of AI on legal language, we need to look at some of the risks of AI, so we need to start from a more distant perspective in search of an answer.

Why we afraid of this tool, although we can use it for translations, analysis, drafting documents and countless other applications? To give a short answer, it would be that we fear the autonomy of the machine. The loss of human control. What do I mean?

If we simplify it, we can say that an artificial-intelligence system is a software that can be created (evolved or developed) in different ways and that produces different outputs related to the goals set by humans, such as making decisions or recommendations, or even creating other content and thus influencing its environment (Tóth 2022: 164).

The “soul” of the software is a set of algorithms that contain the steps to be executed. Given that these sequences of steps are specified and controlled by the programming human intelligence, as long as the algorithms themselves are under control, we can be considered to be in control of the machine. However, there are already AI systems that can make autonomous decisions and react autonomously to the environment. The question is, however, to what extent this autonomous behaviour can be extended, to what extent it can deviate from the will of its creator. The extent of autonomy that can be achieved by the machine and how it can be kept under human control is therefore the key problem.

It is no coincidence that, after various formulations, this is the definition of artificial intelligence finally adopted by the European Parliament and the Commission [6]:

“AI system means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments” [Art. 3(1)] (URL4).

AI systems with a high degree of autonomy can certainly be a cause for concern. The most tangible example comes from the military. Since human

lives are at stake, cases have attracted the attention of the world, such as an automated drone attack. Autonomous weapon systems powered by artificial intelligence have huge risks. Modern armed conflict is becoming a computer game: destruction is monitored only remotely, via a display, and the lives of the soldiers carrying it out are not threatened by the collision, making the killing of others a matter of no consequence. I do not want to refer simply to the moral hazard here.

As one expert writes about the risk of such life-and-death situations:

“[The] machine’s decisions may not even be understood, this is called the black box problem. This means that because AI does not follow instructions, but learns and synthesizes, after a while the developer himself will not understand why the AI made this or that decision” (URL5).

1.2. Speaking of shortcomings, I will refer to the commonly known drawback to which the authors usually draw attention, namely that AI does not have the emotional intelligence of humans. AI is not capable, for example, of extracting the essence and identifying the legal problem from the often difficult to unravel, often incoherent, confused verbal communication of a litigant. It is not suited to dealing with an angry or desperate client with empathy, all within a tight timeframe. It cannot deal with a traumatised and then re-traumatised client who has been in the courtroom and then collapses as a result. The machine has no sense of fairness, no sense of vocation, no sense of companionship, no ability to draw the very fine line between law and morality (Péter 2022: 570).

I would like to add it: for the time being.

2. Nowadays, however, the two main shortcomings of AI: the lack of emotional intelligence and the risk of losing control over its autonomy, are behind all its actions and decisions. How does this relate to the issue of legal language?

Our mother tongue is our common national treasure, which has been polished, expanded, carved or even coloured by generations. Words are not just cold information, they have a mood, they are linked to historical or personal memories, they evoke emotions. Anyone who doubts this allegation should read a poem. Legal language is not without these overtones, moments, emotions, and if you don’t believe it, visit a courtroom.

The current Penal Code also uses the concept of thief in defining robbery – a qualified case – as follows: “Robbery is also when the thief caught in the act uses violence or direct threats against life or limb to keep the property.”[10]

Although the word “thief” is a perfectly ordinary word in this norm, to understand the true meaning of the word “thief” within the norm, we need a much broader knowledge.

In linguistics, thief language is a separate category (URL6) and has its own literature (URL7). It takes a huge amount of knowledge and a huge amount of intelligence for the AI to really, as I have indicated in this example, deal with legal concepts properly. Not to mention the fact that even within the European Union, the national translations and thus the content and interpretation of certain legal concepts are often not the same. Not to mention, for example, how can the principles of criminal procedure be interpreted by the MI? Article 1 of the current Criminal Procedure Act lays down the presumption of innocence, which must prevail and be enforced until the court has ruled that the accused is criminally liable.

The formulation of the concept does not pose a problem for the MI, but how can this concept, together with its prosecutorial content, be reflected in a substantive decision closing criminal proceedings delegated to the MI?

Another example could be the legal metaphors or legal technical terms. To say that there is a “chilling effect” on freedom of expression by some punishment, or to apply the “fruit of the poisonous tree” principle to evidence, is much more emphatic and effective than to talk about the prevention of freedom of expression or that evidence was obtained illegally. There is a significant difference between someone making a request and someone presenting it. The latter carries greater respect in Hungarian. The series of examples can be continued as you like it.

An unemotional intelligence cannot create a term, develop a concept, that does not require mere cold information. One might say to this, let us write an algorithm that includes emotion. Think it through, how would we describe happiness or love?

For example, different dictionaries interpret the word like differently. According to the wiki dictionary (URL8), the word like means “to feel a sincere affection for a person, living or not living thing. He or she is attached to, identifies with, appreciates; wants to be with.”

According to the Interpretive Dictionary of the Hungarian Language “it loves someone (person) is to feel affection for someone or something; to have a tender feeling or affection for him or her” (URL9).

In the dictionary of synonyms, the following terms and concepts can be found for the word love: likes, adores, is attracted to, adores, is inclined to, is fond of, is worshipping, idolizes, adores, is sympathetic to, is attracted to, etc. (URL10).

This would have to be coded for a machine that has no human body, that cannot feel the warmth that fills the human “heart” when it loves or is happy, not to mention the fact that the hormones and chemical processes involved are different in each person.

The other deficiency is not easy to remedy. At most, we would find that our machines think and act according to rules and legal concepts that we do not understand. It delivers justice according to its own arbitrary law, which is impermissible in law. However, it is capable of performing tasks that seem monotonous, and emotional intelligence is irrelevant, based on a large amount of data from given criteria.

It is well known that the development and role of AI is now unstoppable. Where is it present in everyday life? Most of us encounter its activity when using our mobile phones to run various search engines. Another important area is the world of translation systems. Official translators are now using these translation systems. Smart cars that brake against our will if they detect a collision hazard or park for us. They are also slowly becoming indispensable in healthcare, and more.

In the months, perhaps years to come, the rise and constant renewal of generative artificial intelligence is already a reality. Recent news includes Meta, which is not trying to catch up with Microsoft and OpenAI, Google and the Chinese giants. The focus of development is a new language model that will form the basis of an AI system (URL11).

The new system will be accessible to LLaMA researchers, governments, NGOs and academics. The significance of the system lies in the fact that it analyses huge amounts of text, but does not stop there, it then summarises information and generates content to answer questions, and does so in a very human and direct-looking style. Moreover, the diversity of languages is not an obstacle in this work, as it works in almost 20 languages. We must therefore feel a sense of responsibility towards national languages, our own languages.

Well, I think it is important to underline that the already mentioned European Union Regulation [17] emphasizes the following ideas regarding human decision making and AI:

“For the purposes of this Regulation, an artificial intelligence system that does not substantially influence the outcome of a decision shall be understood as an artificial intelligence system that has no influence on the substance of the decision and thus on its outcome, whether human or automated. AI systems that do not substantially affect the outcome of decision-making may be situations in which one or more of the following conditions are met.” [18]

The first assumption is that the AI system is intended to perform only a narrow procedural task. For example, these systems convert unstructured data sets into structured sets, organise incoming receipts and documents. A second condition is that the task carried out by the AI system must be aimed at improving the result of a human activity that has been completed previously. The EU regulation contains important messages for decision-makers.

2. Is the language of AI our language?

For years, scientific articles have been published about how OpenAI and UC Berkeley have created AI that uses its own language. In fact, it not only uses it, but can even develop it further when needed. Well, it doesn't need any instructions to do that (URL12). Language has been a mediator since time immemorial, expressing thoughts, emotions and so much more. A human being is incapable of learning all the languages on earth, but the human brain has been able to enable AI to do so. An important question is whether the language of the AI can evolve without a human brain and human thoughts, or whether thoughts can exist without language. Research seems to suggest that mechanical language can arise and evolve without a conscious brain.

However, it must be said that writing scientific papers, summarising research results and formulating *de lege ferenda* proposals, or solving mathematical problems is not an easy task even for AI. A good example of this is the fact that the Galactica thunderbolt published by Meta was quickly cancelled because it was unsuitable for solving problems, it kept giving wrong answers. In the answering process, linguistic content plays a huge role in the interpretation of texts, the fact that languages and our language are not a simple algorithm that gives an exact answer to every question, but a wonderful world with which we express our thoughts, emotions, creativity or emotional intelligence. In the process, we “manage” our own emotions, we also strive to get into the emotional state that will help us best solve the problem at hand, and we also recognise the emotions of others in the process of finding the keys to a solution. As a result, we also manage relationships. A smile, an encouraging look, or as Dezső Kosztolányi wrote in *Sweet Anna*, “... when the eyes vote”, is no substitute with the most intelligent AI.

References

- [1] URL1. <https://www.coe.int/en/web/portal/-/council-of-europe-adopts-first-international-treaty-on-artificial-intelligence> (Downloaded: 28 June 2024)
- [2] Position of the European Parliament adopted at first reading on 13 March 2024 with a view to the adoption of Regulation (EU) No 2024/EU of the European Parliament and of the Council laying down harmonised rules on artificial intelligence and amending Regulation (EC) No 300/EC /2008, 167/2013/EU, 168/2013/EU, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (1)
- [3] Czine, Ágnes 2022. Constitutional aspects of clarity of legal provisions. *Magyar Nyelvőr*, 46 (special issue). 65–80. <https://doi.org/10.38143/Nyr.2022.5.65>
- Arató, Balázs 2022. Norm clarity in the light of Hungarian case law. *Magyar Nyelvőr*, 46 (special issue). 81–90. <https://doi.org/10.38143/Nyr.2022.5.81>
- Arató, Balázs – Balázs, Géza 2022. The linguistic norm and norm of legal language. In: *Magyar Nyelvőr* 46 (special issue). 91–103. <https://doi.org/10.38143/Nyr.2022.5.91>
- [4] URL2. Kitti Mezei 2023. Current issues in the regulation of artificial intelligence in the European Union. *In Medias Res* 1. 53–71. <https://orcid.org/10.59851/imr.12.1.4> (Downloaded: 28 June 2024)
- [5] URL3. An algorithm is a precise description of the solution to a problem as a sequence of simpler steps. An algorithm specifies which operations are to be performed in which predefined order. Source: <https://lexiq.hu> (Downloaded: 28 June 2024)
- [6] Norbert Tóth: Artificial Intelligence in the Service of Linguistic Rights – Opportunities and Risks (MJ, 2022/3, p. 164)
- [7] REGULATION (EU) 2024/... OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of ... laying down harmonised rules for artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139, (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (the Artificial Intelligence Regulation)
- [8] URL4. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138-FNL-COR01_HU.pdf (Downloaded: 28 June 2024)
- [9] URL5. Anton Bendarzhevsky, Autonomous military systems: the time of impunity has come, Index, 20 March 2023. <https://index.hu/kulfold/2023/03/20/autonom-katonai-rendszerek-dronok-mesterseges-intelligencia/> (Downloaded: 28 June 2024)
- [10] Kitti Péter: Judicial Competences in the 21st Century, with Special Reference to the Rise of Artificial Intelligence MJ, 2022/10, p. 570.
- [11] Act C of 2012, § 365(2).

- [12] URL6. <https://www.arcanum.com/hu/online-kiadvanyok/Lexikonok-a-magyar-nyelv-szotara-czuczorfogarasi-55BEC/t-6E0A8/tolvajnyelv-6F034/> (Downloaded: 28 June 2024)
- [13] URL7. https://mnytud.arts.unideb.hu/szleng/szl_kut/04_111ev/05lihacs.pdf (Downloaded: 28 June 2024)
- [14] URL8. <https://wikiszotar.hu/ertelmezo-szotar/Szeret> (Downloaded: 28 June 2024)
- [15] URL9. <https://www.arcanum.com/hu/online-kiadvanyok/Lexikonok-a-magyar-nyelv-ertelmezo-szotara-1BE8B/> (Downloaded: 28 June 2024)
- [16] URL10. <https://szinonimaszotar.hu/keres/szeret> (Downloaded: 28 June 2024)
- [17] URL11. Zuckerberg, who is still the majority shareholder and CEO of the company, will be the CEO of “Meta Platforms Inc.” in October 2021. “, or Meta for short, to refer to the company’s future efforts to build the metaverse, in which he sees the future of Facebook (however, there was already so much under the name “Facebook” that the name change was justified anyway). https://itcafe.hu/hir/meta_generativ_ai_mesterseges_intelligencia.html (Downloaded: 28 June 2024)
- [18] REGULATION (EU) No 2024/... OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of ... laying down harmonised rules for artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) No 2018/858, (EU) No 2018/1139, (EU) No 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Regulation) (Artificial Intelligence Regulation) (para 53)
- [19] REGULATION (EU) No 2024/... of the EUROPEAN PARLIAMENT AND OF THE COUNCIL of ... laying down harmonised rules for artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139, (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (AI Regulation) (paragraph 53)
- [20] URL12. https://itcafe.hu/hir/sajat_nyelvet_fejleszt_a_mesterseges_intelligencia.html (Downloaded: 28 June 2024)

Ágnes Czine

Constitutional Judge, Professor, Head of Department
Károli Gáspár University of the
Reformed Church in Hungary
E-mail: czine.agnes@kre.hu
<https://orcid.org/0000-0001-5155-0377>

WHAT WE TALK ABOUT WHEN WE TALK ABOUT AI. ARTIFICIAL
INTELLIGENCE AND LANGUAGE

Abstract

With the introduction of AI tools and software into an increasingly large number of applications that are used across society, including education, it is of significant importance to consider the nature, role and use of language, both in terms of the language at the heart of large language models and the language at the heart of humans. Language has been the evolutionary miracle that has set humans apart from the rest of the animal kingdom. It has led to the production of language in the form of written records and spoken output, which has shaped our cultural and social life. The use of AI large language models and tools has now given us the potential to regenerate a human interest in what Merlin Donald referred to as “our collective memory palace”. It may bring us closer to a deeper, linguistic appreciation of the riches of humanity’s creative and intellectual treasures. This article takes a brief look at some critical areas of language history and theory. It identifies the importance of the library as a crucial public institution that can help shape the relationship between AI and humans.

Keywords: artificial intelligence, large language models, language, humanity, libraries, language theory

Background

The initial, global buzz around generative AI tools that were released in November 2022 was the excitement of having instant linguistic interactions with a screen. Here was a brave new world where questions were typed into a simple display panel, and an impressive array of text instantly flowed back on the screen. To many, the experience was one of amazement often followed by a desire to immediately discover the extent that interactions could go to. Here we were, interacting with a machine that talked, or messaged, back to us and communicated in such a way that we had never experienced before. A human-to-computer language interface. It was not long before many discovered the microphone icon embedded on the interface page, and the interaction with the AI program became a two-way voice dialogue. Not only could the AI voice give impressive responses, but it could easily be asked to reframe its responsive style and even the language used. Ask a question in English, receive an answer, even in another language, and adapt the style

and voice register at will. Here we were in a new linguistic paradigm with technology. At the very heart of the new generative AI tools, or LLMs, is the word language. The importance of language and the importance of our understanding of these new forms of technology with regards to language are extremely significant as these tools become more present in our everyday lives. In this article, we look at some important issues of language and make an individual case for seeing how our libraries and their potential use of AI as an area of significant importance in this relationship may develop (Kőkuti et al. 2023; Rajcsányi-Molnár et al. 2024a, 2024b).

Linguistic knowledge

What exactly can we do as we are faced with the inextricable rise of AI and the inevitable challenges to everything that we have previously considered at least academically stable with regards to language. One crucial area of consideration is language use. In linguistics, the field of pragmatism is concerned with the 'use' of language; what Saussure termed 'parole'. This external use of language, or surface structure, in linguistics has a far more settled contemporary understanding than the more complex and challenging linguistic study of deep structures of language and is neatly summarised by Berwick and Chomsky thus:

“[...]certain key biological components associated with language, in particular the 'input-output' system of vocal learning and production that constitutes part of the system we will call 'externalizaion', have been clarified biologically and genetically, so much so that we can effectively use a 'divide-and-conquer' strategy and place this sensorimotor aspect of externalization aside while we focus on language's more central properties” (Berwick and Chomsky 2011: 2).

Although of crucial relevance to the understanding of how language actually is formed in the human mind, issues of acquisition and deep structure can generally be put to one side when looking at the pragmatics of language; that is, our everyday use of language and what that may have to say about our relationship with generative technology and language.

It is the overwhelming and fundamental success of homo sapiens that we have developed a language ability that has allowed us to convey a thought or idea from one person to another or many. The evolution of human language shows that this capacity for using language to convey a thought or an idea has been a long and hard-won process. Estimates for the human ability to

speech range between 100,000 and 50,000 years ago. It is the unique human capacity, and it has been a process of change and refinement ever since. It is also a uniquely social process. As Tomasello (2011: 343) explains:

“Language, or better linguistic communication, is (thus) not any kind of object, formal or otherwise; rather it is a form of social action constituted by social conventions for achieving social ends, premised on at least some shared understanding and shared purposes among users.”

Humans are bombarded with language from the minute they are born. A process of symbolic interpretation of spoken language starts in the early months of a new-born child’s life, and the growth in linguistic comprehension and production is one of the most incredible human cognitive achievements that truly separates us from the rest of the animal kingdom. Utterances begin to occur as early as eight months, and a fully developed spoken syntax generally forms in a child by the age of three. Young humans are truly remarkable language creators. Dehaene (2019: 67) explains this incredible human linguistic feat as:

“In a few months, children quickly manage to surpass any existing artificial intelligence algorithm. By the time they blow out the candle, they have already laid down the foundation for the main rules of their native language at several levels, from elementary sounds (phonemes) to melody (prosody), vocabulary (lexicon), and grammar rules (syntax).”

Becoming literate

Our ability to represent speech with visual or written symbols is even more recent in terms of evolutionary time; our first written symbols appeared approximately 5000 years ago. Again, the development and refinement of the written word has been a long, hard-won process. Our biologically primary spoken capacity for language and our non-biological secondary ability to use a written symbolic code for writing are uniquely human attributes. No other animal can do anything remotely what we do in terms of speech and symbolic representation.

The technological invention of the printing press in the 15th century dramatically transformed our human relationship with text. From that point onward, written language, which had been the preserve of the elite, often religious, scribes, gradually allowed the general population to become literate. Again, this has been a hard-won transformation that is still being

played out in education systems across the globe, as can be seen in the huge efforts that are poured into elementary education in order to achieve a level of functional literacy.

The education system is infused with language both from without and within. We are inextricably tied to language, as is evidenced by the insistence on learning the foundations of written language and literacy in most elementary stages of education globally. Donald (2001a: 232) explains this as a process which eventually becomes automatic, and

“once they are learned, the demons will go downstairs, the ideograms will become invisible, processed down below, in the automatized basement of the mind, and your attention will focus only on the meaning of the text”.

It is this immense cognitive battle that has been at the heart of education, arguably since classical times for some and certainly since the inception of mass education in the nineteenth century. Donald makes the important point that once basic decoding of an alphabet has been mastered the attention turns to meaning. Quine (1968: 143) explains this crucial point thus:

“There are two parts to knowing a word. One part is being familiar with the sound of it and being able to reproduce it... The other part, the semantic part, is knowing how to use the word.”

Developing and mastering language use is at the core and heart of schooling. From the initial teaching of the ability to decode the alphabet through to a functioning literacy by the end of the elementary schooling period students are subjected to a relentless focus on language. The relentless addiction to written signs and symbols continues to be the major element of education in the upper years. Students are asked to read increasingly complex text, asked to answer questions that assess their comprehension and appreciation of those increasingly complex texts and asked to become adept at producing a wider range of written texts in the form of essays, arguments, persuasive texts, among a few of the tasks that students are faced with as they progress through the higher stages of the formal schooling stage. What is being increasingly asked of students is that they attend to a precision of language. As Austin (1976: 73) says:

“Language as such and in its primitive stage is not precise, and it is also not, in our sense, explicit: precision in language makes it clearer what is being said – its meaning: explicitness, in our sense, makes clearer the force of the utterances, or 'how it is to be taken'.”

By the time students graduate from secondary education, many are regarded as having mastered a level of linguistic ability that, when viewed against the timeline of human cognitive evolution, is truly staggering. True, some students have not been able to reach that level of mastery and the failures of mature education systems with regards to this are constantly being reviewed by many societal organisations, from education schools, governments and global agencies. But let us take a positive, holistic view of the literacy ability of global young learners today. By any measure they are generally extremely literate having navigated the journey of language over their first eighteen years. Despite regular dire warnings as to the current state of literacy, from a wider perspective, the transformation of literacy for the majority of young people worldwide has been one of the most significant social achievements in history.

The whole biosphere of further academic development has traditionally relied on the key role of language, and the development of the ability for learners to improve continually has been a constant in the fields of psychology, learning science and most other academic disciplines. Admittedly, our "collective memory banks" (Donald 2001b: 569), the traditional storehouses of knowledge such as books, scientific papers, articles and journals, have multiplied to an almost infinite size in the digital age and have left us at the stage where the most pressing problem has become that of "how can we educate people to handle the sheer immensity of our global memory palace?"

The introduction of artificial intelligence

The introduction of AI large language models has further challenged the nature of our relationship between our human ability to navigate a collective language space to an almost impossible level. Today, this whole ecosystem has come face to face with an existential crisis of meaning. The control of language has moved from institution to device. The role of the expert has moved from the classroom to the personal screen and headphones. The gatekeepers of knowledge, with language at the heart of the process, have shifted from the expert to the machine. And this has profound consequences for all. Bruner (1996: 7) saw this dichotomy between a cultural and computational view of meaning-making: There is no decision procedure known that could resolve the question of whether the incommensurability between culturalism's meaning-making and computational's information processing could ever be overcome. Yet, for all that, the two have a kinship that is difficult to ignore.

Once meanings are established, it is their formalizations into a well-formed category system that computational rules can manage.

It is of interest that Bruner chose the term 'kinship' to describe the possibility of combining our cultural and computational meaning-making systems or what Donald (1991: 358) has described as our "hybrid minds".

Our new, computational-based, artificial language models fundamentally shift our relationship with language. These models force us to confront the relationship between language and meaning. As Clark (2024) in his concluding paragraph states: It turns out that this technology is not about technology-in-itself, something out there to be tamed. It is about US! Our collective cultural legacy has been used to train these amazing models. It is we speaking to ourselves, asking what we should do next. The Copernican revolution is not something out there but within ourselves. We are back at the centre of our relationship with knowledge and our future.

Importantly, Clark highlights these large language models' transactional, dialogic nature. The interactivity we can now enjoy with being able to interrogate and receive instant feedback with these models is bringing us much closer to what has been our experience of what has been the social development of language that we have experienced as humans throughout our history. As Tomasello stressed, we learn through linguistic interaction. However, we do need to remain acutely aware of our human need for quality control and our capacity for judgement of language use. For it was Montaigne, writing in the second half of the sixteenth century, who warned us that:

"Just as our mind is strengthened by communication with vigorous and orderly intellects, it is impossible to say how much it loses and is debased by our constant intercourse and association with mean and feeble intellects" (Clark 2024: 287).

As we begin to interact and 'talk' with AI in a more natural linguistic communicative way we have to recognise that our interactions are with a computational model that, according to Donald (1991: 24) "knows nothing of the world to which it refers".

A fundamental shift and the future role of libraries

Today we are witnessing a fundamental shift in our relationship between human language and our digital technologies. We are witnessing the power of both AI and the power of language and it is increasingly returning to the original functions of language – of speech and thought. The challenge that

we face today is how to find some practical, yet innovative, ways to connect our age-old human quest for knowledge and understanding through language with these new language-based machines. These are not easy challenges and, as we progress with understanding our relationship with artificial intelligence, will require expertise in what Searle (2008: 10) referred to as the “future of philosophy” in such areas as the philosophy of language and of society (Balázs 2023a, b).

It is our contention that we do indeed have a vital and long-standing, structural and intellectual model in our libraries that can help us navigate this changing world. In Donald’s (2001a: 572) evaluation of our relationship with our collected and stored knowledge, he states the fundamental question that we are facing:

“A growing mind still has the same basic need for structure as it ever had. Cultural astronauts need a home planet before venturing into representational space. They need a map and a guide, at the very least, and some rules.”

The challenges of the information age and technological developments, including the advances in AI technology, have created a new perspective for public collections in content collection, storage navigation and interactivity (Arató–Balázs 2023: 756). The traditional museum and library practice approach is gradually changing in line with the expectations of the new digital age. Changing attitudes is of paramount importance in involving users. Public collections have traditionally been institutions dedicated to the preservation and transfer of information. However, the rise of digital technologies and the spread of the internet have opened the door to interactivity and community collaboration (Balázs 2023b). In this new approach, public collections are no longer only places of storage and research, but also active participants in knowledge sharing and communal, linguistic creativity. With the help of large language models, digital tools, visitors, researchers and interested parties will be able to more actively participate in connecting more closely with the content of library connections. The memory palaces will hopefully become palaces of active, fruitful and important discussion.

This change of attitude will not only increase the diversity of the content of public collections, but will also enhance the experience of involving users. Interactive public collections not only offer static information, but also provide opportunities for active dialogue and cooperation between audiences and institutions. The resulting public collection content will become much more relevant and authentic for users, as community members become part

of preserving knowledge and culture. In addition, public collections built in this way are fit for the modern digital age and offer new opportunities for education, research and creativity. Researchers can collaborate on their discoveries through virtual networks. Virtual data and knowledge networks also reveal previously unexplored connections, while ever-expanding connections provide new resources for research. The growing wealth of data and information provides new inputs for AI research and development and creates synergies for growth that can be exploited in both processes. Ultimately, by using large language models in connecting with the content of the treasures from our collective, stored records in our libraries we would hope for a regeneration of interest in the linguistic riches that have always been there waiting to be found by the reading public (Balázs 2021: 56)

Conclusion

There is a need for educators, and all those associated with education in wider society, to begin to focus on the role and importance of human language when dealing with this new world of AI. For many in our new age of computation, the actual human processes of language acquisition and the qualities of language that make us the truly unique species have often been assumed. In many ways, it has been taken for granted that a biologically primary function (that of human speech) has been enough. It is by no means enough. The daunting pedagogical task of achieving symbolic literacy that every one of us must go through and that is the beginning of an incredibly complex marriage of both cognition and culture in our socio-cultural worlds is one that we should begin to appreciate in this new age, ultimately because it is of such vital importance to our understanding of ourselves as cognitive beings as Sperber and Hirschfeld (1999: 127) express: “Ordinary cognitive activity does not take place in a fixed experimental setting where the information is available is strictly limited and controlled, but in a complex, information-rich, ever-changing environment...In the human case, moreover, the environment is densely furnished with cultural objects and events most of which have, at least in part, the function of producing cognitive effects.”

It is our hope that, as we navigate the future of artificial intelligence use, we may reach a reconnection with the wonders and mysteries of human language which have been the subject of an incredible process of cultural development over thousands of years, where we often find ourselves stunned by the immense power of words to express what sometimes can not even be expressed - the innermost workings of our human minds. What lies hidden beneath the surface of language, and AI, may be exactly where the truth may

be hiding. As E.M. Forster (1926: 93) in a stunning moment of insight in *A Passage to India* put so very well: “Mrs Moore shivered, ‘A ghost!’ But the idea of a ghost scarcely passed her lips. The young people did not take it up, being occupied with their own outlooks, and deprived of support it perished, or was reabsorbed into the part of the mind that seldom speaks.”

References

- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban? A digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 2023: 755–76.
<https://doi.org/10.38143/Nyr.2023.5.755>
- Austin, John Langshaw 1976. *How To Do Things With Words*, second edition, Oxford University Press, London.
<https://doi.org/10.1093/acprof:oso/9780198245537.0010001>
- Balázs, Géza 2023a. Tamás Ildikó: „Adj netet!” *Magyar Nyelvőr*, 147. 2023: 120–7.
<https://doi.org/10.38143/Nyr.2023.1.120>
- Balázs, Géza 2023b. *Az internet népe. Internet – társadalom – kultúra – nyelv. A kulturális és a tervezett evolúció határán.* Ludovika Egyetemi Kiadó, Budapest.
- Balázs, László – Rajcsányi-Molnár, Mónika – András, István – Sitku, Krisztina 2021. Social Responsibility and Community Engagement at a Hungarian Regional University. *Journal of Higher Education Theory and Practice*, 21. 54–63.
- Berwick, Robert C. – Chomsky, Noam 2016. *Why Only Us: Language and Evolution.* The MIT Press, Cambridge.
<https://doi.org/10.7551/mitpress/9780262034241.0010001>
- Bruner, Jerome 1996. *The Culture of Education.* Harvard University Press. Cambridge. <https://doi.org/10.4159/9780674251083>
- Clark, Donald 2024. *AI is the Copernican collective, cultural mind - it is US!* <https://donaldclarkplanb.blogspot.com/2024/02/ai-is-copernican-collective-cultural.html>
- Dehaene, Stanislas 2021. *How We Learn: The New Science of Education and the Brain.* Penguin Books, New York.
- Donald, Merlin 1991. *Origins of the Modern Mind.* Harvard University Press. Cambridge.
- Donald, Merlin 2001a. *A Mind So Rare.* Norton. New York.
- Donald, Merlin 2001b. Memory Palaces: The Revolutionary Function of Libraries. *Queen's Quarterly*, 108(4) 559–72.
- Forster, Edward Morgan 1965. *A Passage to India.* Progress Publishers. Moscow.
- Kőkuti, Tamás – Balázs, László – András, István – Rajcsányi-Molnár, Mónika 2023. Collaborating with Artificial Intelligence – AI in Business Communication Education. In: *IEEE - Óbudai Egyetem; IEEE (szerk.) 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE): Proceedings Danvers (MA), IEEE.* 287–94.

- Montaigne 1958. *Essays*. Penguin Books. New York.
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István – Czifra, Sándor 2024a. Competition as an Effective Motivational Tool in Online Education. In: Anikó, Szakál (ed.): *IEEE 11th International Conference on Computational Cybernetics and Cyber-Medical Systems: ICC 2024: Proceedings*. IEEE Hungary Section Budapest. 83–8.
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István 2024b. Online Leadership Training in Higher Education Environment. *ACTA POLYTECHNICA HUNGARICA* 21, 39–52.
- Quine, Willard van Orman 1968. Ontological Reality. *Journal of Philosophy*, 65(7) 185–206.
- Searle, J. 2008. *Philosophy in a New Century: Selected Essays*. Cambridge University Press, UK. <https://doi.org/10.1017/CBO9780511812859>
- Sperber, Dan and Hirschfield, Lawrence 1999. Culture, Cognition and Evolution. In: Robert A. Wilson and Frank C. Keil (eds) *MIT Encyclopedia of the Cognitive Sciences*, MIT Press. Cambridge. cxi-xxxii.
<https://doi.org/10.7551/mitpress/4660.003.0009>
- Tomasello, Michael 2010. *Origins of Human Communication*. MIT Press. Boston.

Geoff Vaughan

PhD Student, Full-time Lecturer
Eszterházy Károly Catholic University
Doctoral School of Education
E-mail: geoffvaughan1@gmail.com
<https://orcid.org/0000-0002-9614-1195>

Rita Szűts-Novák

Educational Researcher, Assistant Professor
Eszterházy Károly Catholic University
Institute of Digital Technology
E-mail: szuts-novak.rita@uni-eszterhazy.hu
<https://orcid.org/0009-0007-5199-047X>

Z. KARVALICS LÁSZLÓ

ARTIFICIAL INTELLIGENCE AND LANGUAGE – A DIFFERENT APPROACH TO
THE PROBLEM

Abstract

The article explores the intersection of artificial intelligence (AI) and language, critically examining the development and impact of large language models (LLMs) like GPT. It argues that while these models demonstrate significant advancements in generating linguistically coherent text, their achievements are primarily quantitative rather than qualitative. The text highlights fundamental issues such as copyright concerns and the limitations of synthetic data. It also critiques the anthropomorphizing metaphors used to describe AI capabilities, emphasizing the need for precise language to avoid misconceptions about AI's abilities. The article advocates for focusing on AI's practical applications and ethical considerations rather than overestimating its intelligence.

Keywords: artificial intelligence, Language Models, GPT, Metaphor in AI, Synthetic Data, Ethical Considerations in AI

Over the past two years, the success of increasingly large language models (*LLMs*) in artificial intelligence (AI) has brought a growing number of GPT tools, competing with each other and capable of doing more, into the spotlight.¹ GPTs (*Generative Pre-trained Transformers*) are, in simple terms, programs based on neural networks that construct linguistically correct text responses to natural language textual instructions (questions) using their massive text databases on which they have been trained. Moreover, since the elaboration of these answers, thanks to clever underlying algorithms, very often appears to be human-level, the great leap forward in language technology capability has immediately set off a new wave of discourse, even more futile and meaningless than before, about the now human-embattled stage of the growth of intelligence in machine systems.

¹ For the latest chapter in the history of Natural Language Processing (NLP), see ChatGPT Héja (2024).

The double-faced nature of the significant language models

The analysis that gives rise to such high hopes does not consider that the text generation resulting from the increased predictive power of language models (they are better at continuing a sentence they have started) is based on a quantitative, not qualitative, leap in technology. The qualitative gain is in the outputs and the performance. Beyond the “upgrading” of more rudimentary solutions, it is mainly seen where the use of chatbots has become meaningful and possible in areas previously not or only partially covered by automated solutions.

This enthusiasm does not face up to two fundamental strategic problems with the textual raw material reserve: the gradually judicialising copyright concerns about the texts used and the growing threat of a limit due to the running out of textual ‘fodder’. The ‘synthetic data’ escape route (artificially generated texts to teach the model) is dead because the best current solutions are partial and incomplete. Indeed, the existing uncertainties and errors would be multiplied, so the answers must be accepted with many reservations. Behind the misrepresentations, which are often referred to as ‘hallucinations’, lies a lack of appropriate textual content, but worse still, the constructions offered are based on unchecked and ‘non-quality-assured’ content that is inherently flawed, corrupt, biased, erroneous, one-sided, created by irresponsible publishers.

Paradoxically, the very success of this language technology highlights its substantial limitations. It is enough to think that even ten years ago, there were seamlessly functioning systems that could edit fresh sports and stock market data and weather forecasts into complete and flawless news material and thus take over the text production tasks from humans as “robot journalists”. However, this is due to the syntactic preformatting and closed lexicon of the news types in question, with their unambiguous meanings. It is immediately apparent that the system is inefficient and unreliable where these boundary conditions do not exist. However, any other text type or situation where these constraints are met will be capable of being created by a machine element replacing a human one. This is no coincidence: the generation of these text types is mechanical and low-value-added intellectual work, well worth freeing up a valuable lifetime for other purposes.

This is true even for the top performances of GPT systems: translating a technical text from one language to another (more precisely, from many languages to many languages) or modifying the text modality in the same language (producing a summary, having an “academic” style, etc.). However,

what is beyond this, we repeat, seems unattainable with this solution: GPT does *not* manifest *linguistic competence but mathematical-statistical competence*. For the systems that perform operations according to their algorithms, there is no such thing as what gives meaning to the text and the smaller linguistic units that make it up: the world of meanings. The output produced is like a linguistic performance. Still, the system's algorithmic correspondence of the signals produced can only be measured and managed, and the content *carried is entirely insensitive*.

György Csepeli puts it in a much more poetic and tangible form:

"In human consciousness, the infinite outside and the infinite inside meet, made possible by the linguistic competence innate in man. However, competence is only a possibility, actualised by linguistic performance. Generative linguistic models of artificial intelligence do not rely on the linguistic performance of poets and thinkers, the keepers of the shelter of the human found at home in language. The generative language models are based on sentences written or spoken by anybody [...] The generative language models, which select from the enormous and ever-expanding mass of words and sentences already spoken, do not think; they only imitate thinking [...] the language that brings everything close to everyone loses its way to the truth when it is subjected to the dictatorship of the public as the language of anybody. The submerged common language of the machine and man threatens the very essence of man. By imitating human thinking, artificial intelligence opens the door to the technical interpretation of thinking, losing sight of existence as the primary goal of thinking" (Csepeli 2024: 67–8).²

In such a situation, to call two AI systems that send signals to each other *interlocutors* and to consider them revolutionary is a scientific provocation. However, the researchers at the University of Geneva were not afraid to say so, as their experiment showed that their trained language model S-Bert, consisting of 300 million neurons, was able to pass its own learned and verified program to another, *simpler* network so that it could reproduce it (Riverland and Pouget 2024).³ What happened was that instead of calling

² Csepeli expresses this in relation to a thought of Heidegger: 'thinking that penetrates into being is fundamentally possible in the medium of language'. The infinity of being is followed by that of language: according to Chomsky, language is free to be an infinite set of sentences formed by the rules of grammar, and thus capable of revealing the multidimensional, complex truth of being. How could artificial intelligence, existing in a closed nanovolume of coding with 0s and 1s rather than being, do the same?

³ <https://www.sciencedaily.com/releases/2024/03/240318142438.htm> Pouget seems to forget how one of the ancestors of all chatbots, Eliza, developed between 1964 and 1967, was once "merged" with the parodistic Eliza, a program that offered the illusion of speech

their solution a “two-element prompt cascade”,⁴ they tried to sensationalise it by winking at the scientific attention economy, as if an artificial intelligence had successfully outsmarted an intelligence gargyle.

It is scientifically more correct, but equally meaningless, to call the “linguistic capabilities” of neural networks “*protocol communication*” (Sato et al. 2008) or “*quasi-language*” (Cangelosi and Parisi 1998). How ‘populations’ of neural networks generate a process similar to the evolution of a simple ‘language’ with informative functions in a specific environment undoubtedly has more analogies with natural language change than in a more straightforward machine language processing situation. However, the situation lacks everything that makes natural language a natural language: what *appears to* the observer to be linguistic is, to the machine side, merely the generation of algorithm sequences. Szummer (2014: 41) underestimates the conceptual superficiality of the engineering world: it is not only “*cognitive scientists who tend to mistake analogies for conceptual identity*”.

Linguistic “frontlines”

Meanwhile, the exaggerations, grand pronouncements and the para discourses that build on them (led by machine intelligence that not only reaches but surpasses human intelligence) distract attention from what is happening at other, equally important, intersections of AI and language.

One such area is the “upskilling” of previously developed applications with artificial intelligence solutions to support everyday speech and writing production.

Not only has OCR moved towards the increasingly high-resolution recognition, transcription, digitisation and text-to-text (e.g. with *pen-to-*

understanding, and Doctor, a program that simulated a therapist extracting questions from previous answers. Ostensibly there was a conversation, ‘two agents interacting in a natural language’, but in reality there was a meaningless and self-reversing code multiplication, which as natural language exposed the profound meaninglessness of the whole. In this experiment, of course, there was more than just two new generations of GPT programs being “put together”, as the Betone Studio’s podcast series “Inhuman” does (<https://betone.hu/shows/embertelen/>): in each broadcast, Betti and Peter, “personalised” in different directions, debate current issues. S-Bert has been programmed to further formalize a new “skill” he has acquired during his programming career.

⁴ A prompt is a natural language instruction that is converted into code, processed and executed by the machine system, and the result is presented again in natural language. What happened in Geneva can be seen as a prompt innovation, where the initial instruction included the addition of the ability to form a secondary prompt from the result of the processing.

print tools that reliably convert handwriting into cursive text), but it has also opened up new ways of mass transcription of private diaries, correspondence, old and damaged manuscripts and documents that are difficult to read but of great cultural and historical importance. An imposing development is *In Codice Ratio* (Nieddu 2021) for the Vatican Library, whose creators used a student-based citizen science platform to train the system alongside artificial intelligence.

Solutions that turn live speech into text are also becoming increasingly sophisticated. Some of these also aim to bridge the gap that separates more minor languages, which are under-represented in AI use, from world languages. (For example, a recent development, *Vulavula*,⁵ recognises and describes the spoken names of people and places in four languages spoken in South Africa. Much progress has also been made in the opposite direction, in machine speech synthesis – the ‘pronunciation’ of written text is so much under siege from natural-looking spoken language that there are reports of new ‘robotic speech’ systems being put into operation almost weekly. A quite extraordinary best practice is what the community of Maori speakers, led by a non-profit radio station, Te Hiku Media in New Zealand, has done to ensure that they, rather than a global company, are the developers and owners of the speech recognition software that had to be created to digitise their audio archive and the collection of idioms that aids correct identification and transcription.⁶ The project has also renewed the Maori language, but in a way that has avoided the pitfalls of ‘digital colonialism’ (it is no coincidence that the Mohawk Indians in Canada and the Native Hawaiians have undertaken similar projects).

AI can also be used to research/model several important linguistic issues: the origin of language, emergent aspects of language development and the evolution of language itself (Vogt et al. 2005). AI has also been used to combat language extinctions that have led to the loss of linguistic diversity, most notably in the programme for preserving and conserving African languages.

There seems to be a very long list of issues where language and AI meet, yet we hardly ever address them now in the shadow of ChatGPT. This paper aims to lead an expedition around perhaps the most rarely touched intersection: How appropriate is the use of concepts and language to describe the world of AI and to track its changes?

⁵ <https://beta-vulavula.lalapla.ai/>

⁶ For details see: <https://blogs.nvidia.com/blog/te-hiku-media-maori-speech-ai/>

Generative schema, cognitive path dependence, metaphor trap

In 1875, John Ruskin satirised the British Metaphysical Society, in which he pens a satirical paper on the widely used buzzword of the day, automation, automatism, the automatic adjective. In the same year, in the columns of *Deucalion*, he marvels, swimming against the tide of the zeitgeist, that patently absurd statements were more easily labelled 'nonsense' in earlier times than in this decade when many people speak with complete naturalism of man being 'an automaton' (Anger 2024: 1–2).

One hundred and fifty years on, and we are pretty much back to square one. *"Actual changes in the working environment and the means of leisure can also give the impression that humans are, after all, just 'biocomputers' that get along well and badly,"* writes Szummer (2014: 41). And of course, the reverse analogy is equally strong: the image and fear of the robotic being competing with the human, and then the position of seeing into the computer, and later into artificial intelligence, a 'thinking machine' with powers beyond the capabilities of the human mind, seems to fundamentally determine the way most people think and talk about this common intersection of technology and human reason (running into Daniel Dennett, who, according to Szummer, doubts the existence of a difference between machine and human intentionality).

Nánay (2000: 22) sees how this situation could have developed.

"The analogy between mind and computer program has long determined, and still partly determines, thinking about the mind. Interestingly, the explanation of the mind and the plan for creating the thinking machine interact. Machines were designed to resemble the human mind, and the human mind was discussed in computer terminology. This two-way effect created a very intricate relationship between the two fields of research, and as computers became more and more complex, this relationship was constantly reassessed."

It does not follow, however, that this conceptual 'crossing over' between machine learning and mental activity is correct and appropriate. In other words, this widespread conceptual and associative teleology is the epistemologically correct choice.

At the same time, it is understandable and accessible to reconstruct how the ideas that drifted towards the world of the automaton and the thinking machine were organised around notions of the organisation of knowledge. In the words of the historian Kurt Danziger, a generative schema, a metaphorical framework, was at work here, too, which, through analogy

through networked categories, made it possible to express perceived changes in reality in language in public dialogue. Whoever becomes embedded in this new web of meanings can draw on it to grasp newer and newer aspects (Anger 2024: 3–7). Moreover, this is one of the basic mechanisms of mind functioning.⁷ To use Lakoff and Johnson's (1981) often-cited basic thesis, metaphor extends our existing knowledge of the world.⁸

So where is the problem? What is wrong with the metaphor cloud built around artificial intelligence?

In short, it has not only *become stultifying*. However, it has also explicitly made objective dialogue difficult, deforming public thinking, influencing strategic decisions about AI, and infiltrating the scientific discourse by deepening conceptual confusion. When it came into being, it played a massive role in ensuring that the most important innovations of the culture-changing technological explosion found their way into understanding and that its developments were easy to imagine and follow, helping to break down the barriers to diffusion. However, what became familiar and entrenched in the early stages of technological take-up has fundamentally influenced how the conceptual web has evolved as AI systems and solutions have become more complex and sophisticated – what I like to call *cognitive path dependency*. Since the machine side was from the very first linked with anthropomorphic associations, the words used to describe each new 'capability' offered a promising solution to the most similar human actions – no longer only in the press, but also in the small worlds of research and development, as terms. Thus, 'machine *learning*' and all its new generations (deep learning, reinforcement learning, federated learning) are based on the well-known notion of the learning process. This is based on the undeniable similarity between the more excellent repertoire of actions created by the knowledge acquired through the behaviour of living beings, known as learning, and the higher operational performance of a machine system assisted by human operators. However, there is a vast difference in several essential attributes. Human learning, whether voluntary or spontaneous, is a form of behaviour that is a kind of intersection of the learning environment, the object of the learning process and the internal state of the learner. It involves a multiplicity of relations of meaning and significance and their relation to

⁷ Draaisma (2005) shows how every technological change since the 17th century has acted like a metaphor machine: offering itself as raw material to refine and deepen our understanding of cognitive processes and memory.

⁸ In a recent paper, Mithen (2024) goes even further: he sees in the metaphorical use of Homo Sapiens the surplus that has been added over Neanderthal man, who did not have this ability.

the external and internal worlds, with choices and actions resulting from them as necessary. Machine learning is not a behaviour but a particular case of program execution, where the output of instructions improves the basic program. You could call it algorithm refinement or model refinement, but it is signal processing in the same way as running a single line of program code. However, each occurrence of the word learning adds to the cognitive path dependency, whereby the anthropomorphic associations with a machine system engaged in mere computation are reinforced, becoming stubbornly inextinguishable. There is no stopping it: just as the machine does not 'think' or 'learn', neither does it 'decide', 'discover' or 'drive' a vehicle.

Artificial intelligence makes no decisions but performs calculations on decision situations modelled by machine code using input data and preformulated processing algorithms. This is true for the most straightforward and most complex systems. The IT mini system that controls the traffic light does not "decide" how many seconds it takes to change from green to amber but executes the program that performs it. Moreover, a stock market program with a vast database does not decide; it only appears *to* decide, but it owes its output to its creators. Decision-making is one of the most complex challenges of the human brain, with many brain areas mobilised simultaneously, depending on the peculiarity, the unusualness, the complexity and the history of the decision situation. The systems called Automated Decision Making (ADM) certainly transfer the decision-making process from human to machine, but not the decision itself: the decision parameters and the decision outputs are "set" by the creators of the systems, and the complex challenges of the decisions are also played out in their brains.

In the same way, AI does not "discover" anything but helps researchers fill knowledge gaps more quickly through its computational power.

At the heart of the 'self-driving vehicle', a complex cybernetic system of sensing, processing and control modules is designed to produce real-time outputs from a real-time data set of vehicle speed, attitude and direction, signal transmission and reception. Instead of moving a physical object in a simulation environment, it moves a physical object in reality. It is *as if* an intelligent entity is driving the vehicle, but all the outputs are determined by the algorithms of the engineering team developing the program. If the control system had an experience, it would not be a driving experience but a computing experience.

In summary, the problem with this metaphorical apparatus of verbs is that it attributes to the notion of action the elements of intention and will, indirectly consciousness and a kind of 'persona', self-consciousness, to

the most advanced programs. This nonsensical and countless times refuted starting point poisons the language built upon it, even through the choice of names. “To speak of ‘artificial intelligence’, in the singular, with a definite article, is valid in a single context: as a research direction, as a scientific-engineering field, as a set of disciplines that can be classified as such.” Thus, of course, it is correct to speak of ‘the’ artificial intelligence research, ‘the’ artificial intelligence systems – but there is no such thing, no such entity, as ‘artificial intelligence’. Many different artificial intelligence worlds of varying complexity and embeddedness share the nature of running software on hardware. These include partial integrations. They can be interconnected with other solutions in the digital ecosystem with other functions: robots, sensors, and virtual reality environments, but the resulting more complex solutions will not be life forms and autonomous entities but will be more similar to them because their functioning will be more reminiscent of sovereign and internally driven action. However, even integrated solutions do not boil down to an all-encompassing giant entity – which can then be alarmed at its dangers and how it is preparing to take over from us.

AI applications are, therefore, neither alive nor behaving: they produce automatic outputs at the end of a processing queue. They do not feel, think, decide, choose, or have no will or intention. They have no consciousness. They are not intelligent. The world of meaning is unknown to them. They process signals according to a preformulated set of instructions. Their outputs are “translated back” into meaning by the human mind that uses them. Contrary to Nietzsche’s oft-quoted words (i.e. that the moving army of metaphors constructs truth for us), it is time to acknowledge that structural metaphors are *‘sometimes, in fortunate cases, constructive’* (Pléh 2003: 405), and to return to the metaphor scepticism of the ancient philosophers. They saw metaphors as interfering with the search for truth and, therefore, potentially damaging to science.

On the other hand, it can be seen that the use of metonymic language, which is also prevalent in the context of artificial intelligence, leads to much less misunderstanding and confusion. For example, ‘AI-ethics’ clearly does not exist, but it is clear to all that it should be understood to mean ‘ethical issues related to the development and application of AI’. We know that AI does not ‘take away anyone’s job’ but can improve the efficiency of production or organisational operations by triggering processes that can replace low-value-added human work with machine operations at a lower cost.

Nevertheless, the part-part relationship, a form of metonymy, is consistently pulled in one direction: the machine Part is always emphasised

in the Whole, composed of engineers, hardware and software components, and the human genius that created it is lost halfway through. This is most visibly reflected in the “machine overcomes man” narrative. A game program called checker was able to beat a human player as early as the late 1950s, and then came chess, go, bridge, and poker – and even though these results were, of course, achieved by a hybrid system of computer scientists and computers that turned the experience of millions of human games into algorithms, somehow they managed to land on the still famous, empty and meaningless ‘machine smarter than man’ turn. Moreover, these exaggerations have given way to severe yet silly scenarios in which the machine, ‘winning’ in a game situation, will, in the knowledge of its superiority, ‘defeat’, ‘replace’ or even ‘exterminate’ the human race itself. Moreover, the shadow of cognitive path dependency looms even behind this extreme, which is ultimately beyond discourse. After all, we are talking about a “form of intelligence”, meaning that even these abstract conclusions are permissible. After all, intelligence implies that everything we already know is not a feature of machine signal processing: consciousness, meaning, complex structures of meaning, will, and intention.

The concept of artificial intelligence as a linguistic “original sin”

In the mid-fifties of the last century, computing applications began to move beyond mere computational power or even elementary data processing, and more and more researchers recognised that it was possible to penetrate further realms of algorithmic brainwork gradually. What can be translated into the language of numbers (bits) – images, fonts, sounds – can be processed in quantity and at a speed that the mind cannot. Moreover, the result of this processing can be translated back into the original form of the content so that machine solutions can support natural information processing well. However, the term “artificial intelligence” (AI), the invention of John McCarthy, was by no means a necessary cover. It lacks the ‘as if’ element, suggesting that the machine is really thinking, not just pretending to think. Interestingly, Nikola Tesla was sensitive to this distinction in his more than half-century-old article for *The Century Magazine* in 1900. He writes of the future “clever” machine (quoted in Colton 2023): Independent of all assistance, left entirely to itself, it will be able to perform several operations in response to external influences on its sensitive components “as if it had intelligence”. Of course, at the end of the 1950s, it was impossible to foresee how the word ‘intelligence’ would be subsumed in the anthropomorphising metaphor

cloud that would not interfere with the developments themselves but would threaten the conceptual clarity of thinking about artificial intelligence.

Nor would the 'complex information processing' proposed by Herbert Simon have been a more appropriate choice. Using it would have created a powerful metonymy trap. After all, computers do not process information; they process signals. Information processing is the exclusive domain of the human brain. However, while we objectify our mental contents to convert them into something the machine can consume, the meaning is lost and only returns when the human brain is confronted again with the result of the signal processing. The contact that is the basis of metonymy undoubtedly exists between the two different qualities,⁹ and it does not interfere with understanding if we use the other name instead of one. However, it cannot be ignored in all contexts where the distinction is significant. It is also the only cure for cognitive path dependence at the moment. In the open communication stream of the Internet, it is difficult to bring about change. However, clarification can be achieved in the academic discourse, and a consensus on the use of terms can be reached gradually. In that case, the arguments and emphases will eventually seep into public discourse and public thinking.

References

- Anger, Suzy 2024. Introduction. The Victorian Automata/Automatism Schema In: Anger, Suzy – Wracken, Thomas (eds.): *Victorian Automata. Mechanism and Agency in the Nineteenth Century*. Cambridge University Press. 1–25.
<https://doi.org/10.1017/9781009110129>
- Bencze M. Ildikó 2009. Mi a metonímia? Elméletek és kísérletek a metonímiakutatásban. *Magyar Pszichológiai Szemle*, (64) 4. 677–96.
<https://doi.org/10.1556/MpSzle.64.2009.4.4>
- Cangelosi, Angelo – Parisi, Domenico 1998. The Emergence of a “Language” in an Evolving Population of Neural Networks *Connection Science*, 10(2). 83–97.
<https://doi.org/10.1080/095400998116512>
- Colton, Emma 2023. Resurfaced Nikola Tesla writings about machines with their 'own mind' eerily predict rise of AI. *Fox News*. May 10. <https://www.foxnews.com/tech/resurfaced-nikola-tesla-writings-about-machines-with-their-own-mind-eerily-predict-rise-of-ai>

⁹ The situation is actually much more complicated. According to Petőfi's (2004:30) typology of text as a complex sign, for example, the marker-signified sign relation actually has six different components that are logically and epistemologically distinct from each other: the physical object (vehiculum), its mental image (vehiculum-imago), the formal organization (formation), the semantic organization (sensus), the factual configuration assumed as a world-part, the relatum, and its mental image, the relatum-imago.

- Csepeli, György 2024. Az új fejedelem. *Educatio* 33 (1). 65–8.
<https://doi.org/10.1556/2063.33.2024.1.7>
- Draaisma, Douwe 2005. *Metaforamasina. Az emlékezet egyik lehetséges története.* Typotex Kiadó. Budapest.
- Héja, Enikő 2024. A ChatGPT története. *Magyar Tudomány* (185)6. 815–28.
<https://doi.org/10.1556/2065.185.2024.6.11>
- Lakoff, George – Johnson, Mark 1981. *Metaphors We Live By.* The University of Chicago Press. Chicago.
- Mithen, Steven 2024. How Neanderthal language differed from modern human – they probably didn’t use metaphors. *The Conversation*. May 20. <https://theconversation.com/how-neanderthal-language-differed-from-modern-human-they-probably-didnt-use-metaphors-229942>
- Nieddu, Elena et al. 2021. In Codice Ratio: A crowd-enabled solution for low resource machine transcription of the Vatican Registers. *Information Processing & Management* (58)5. <https://doi.org/10.1016/j.ipm.2021.102606>
- Petőfi S. János 2004. *A szöveg mint komplex jel. Bevezetés a szemiotikai-textológiai szövegszemléletbe.* Akadémiai Kiadó. Budapest.
- Pléh, Csaba 2003. Douwe Draaisma: Metaforamasina. Az emlékezet egyik lehetséges története. *BUKSZ* (15)4. 404–6.
- Riveland, Reidar – Pouget Alexandre 2024. Natural language instructions induce compositional generalization in networks of neurons. *Nature Neuroscience*, <https://doi.org/10.1038/s41593-024-01607-5>
- Sato, Takashi – Uchibe, Eiji – Doya, Kenji 2008. Learning how, what, and whether to communicate: emergence of protocommunication in reinforcement learning agents. *Artificial Life and Robotics* 12. 70–4.
<https://doi.org/10.1007/s10015-007-0444-x>
- Szumner, Csaba 2014. *Freud, avagy a modernitás mítosza. (Tanulmányok és esszék a pszichoanalitikus antropológiáról és beszédmódról).* KGRE – L’Harmattan. Budapest.
- Vogt, Paul – de Boer, Bart – Belpaeme, Tony 2005. Modelling language origins and evolution *IJCAI 2005 Tutorial* 31 July 2005, Edinburgh, Scotland 1–44. <http://www.ling.ed.ac.uk/paulv/tutorial.html>
- Z. Karvalics, László 2020. Méhek és zebrahalak. Adalékok a “fogalmi bizonytalanság” korához. *Szabad Piac* 2. 47–57.

László Z. Karvalics

Senior Research Fellow

Institute of Advanced Studies, IASK/FTI

E-mail: laszlo.karvalics@iask.hu

<https://orcid.org/0000-0002-3502-434X>

THE COMMUNICATIVE-LINGUISTIC MODES OF ARTIFICIAL INTELLIGENCE
WITH A FOCUS ON JUSTICE

Abstract

The paper defines the technical and colloquial concept of Artificial Intelligence. It describes the changes in the technology of speech and writing from primary forms through secondary forms to tertiary forms in the present under the influence of AI. Several examples of tertiary forms are given, such as the rewriting of Imre Madách's drama *The Tragedy of Man*. Other examples include the emergence and role of AI in the judiciary and experiments in the Hungarian judiciary. The authors raise the question of the impact of tertiary communication on evolution (planned evolution).

Keywords: artificial intelligence, speech and writing technology, “third” (tertiary) orality and literacy, planned evolution, AI and justice

1. Conceptual delimitation of artificial intelligence

János Neumann, the mathematician who developed the theoretical foundations of the computer, and his colleagues early on raised the question of whether the concept of thought and reason could be extended to machines. Elemér Lábos (1979) also asked the fundamental questions: what intelligence is and if artificial intelligence exists, etc.

Strictly adhering to the original dictionary meaning of intelligence, artificial intelligence probably does not exist. (1) “The highest degree of cognitive activity, the faculty of reasoning.” It is the activity that is bound up with man. However, in a broader or figurative sense, it is still possible to find meaning in nonhuman phenomena, e.g. (2) “The logical coherence that can be discovered in things and events. The meaningful purpose of something”. Example: There is no sense in the wire (Éksz. 2003: 322). And further meanings are conceivable. In other words, although we may be reluctant to deny the human aspects of reason, the objectification of reason has already taken place, so that linguistically we can speak not only of (human) reason, but also of other, and therefore artificial, reason. As we have already talked about externalised memory (Szűts 2020), externalised intelligence, artificial, electronic brains (McLuhan 2001: 46).

Every technology (technical possibility) affects man in some way, what we call, following McLuhan, technological determinism. Techniques (technical possibilities) are man-made, social products and have already had an impact on human processes. The most important of these is that they have altered human perception, in McLuhan's words, 'if the new technology extends one or more of our senses into the world of society outside us, then in that culture new proportions between our senses will emerge' (McLuhan 2001: 55). To give just one example: pictography, printing and reproduction have transformed the auditory-tactile space into a visual one, and 'affect thinking by their sequentiality and linearity' (Benczik 2001: 321).

Cultural (technical) change therefore changes (expands or recedes) perception and with it the content of human reason. Obviously, it is important to make some restrictions. If reason as the highest intellectual faculty includes (and we believe it does) the world of emotions, then it is precisely in this area, i.e. in the field of spiritual life, emotions and social relations, that we can hardly speak of 'artificial' reason. We are thinking of basic human characteristics such as empathy, love, devotion, sacrifice, self-sacrifice, commitment, the joy found in activity, passion, individual and communal goals, which cannot be 'mechanised' (algorithmised).

Two concepts from the distinctly human world of emotions are worth highlighting. Altruism, or unselfishness, which is a form of behaviour ranging from helping others to self-sacrifice, which benefits the community but may be detrimental to the individual. The evolutionary advantage of homo sapiens was provided by the self-sacrificers, the parochial altruists; its intergroup continuators are the extended altruists (heroes, role models) (Csepeli 2023: 198–200). The other is enthusiasm, which is the self-indulgent delight, enthusiasm, adoration resulting from emotional identification, high spiritual enjoyment; it includes phenomena such as epiphany (shock, rapture, specifically love or enchanting spectacle), catharsis or cathartic experience (experience of experience, purification), and flow (the immersive experience of an action or event). Moreover, they are interconnected, as Márton Szabó (2022: 153) writes:

"it is not only the experiential, cathartic experience of works of art that has an epiphanic character, since such events also happen in everyday life... And it is through such epiphanic experiences that people gain in experience, for example, they become wiser" (people).

To illustrate how complex these phenomena are, I quote Béla Hamvas' (1997: III/102) reflections on enthusiasm:

“What a strange intoxication this is! We can somehow understand that one becomes intoxicated, say, by power and blood and wealth and, say, beauty, if one falls into enthusiasm and goes mad with power or bloodshed or villainy, say, music or poetry or thought, we can even understand if one becomes ascetic and fasts and prays and goes into the wilderness or teaches”.

Csaba Gaál (2020: 33), a medical doctor, in a linguistic treatise, considers the term artificial intelligence to be misleading in several ways. On the basis of medical examples, he believes that what is today called high-tech is best described in Hungarian as smart. It would be sufficient to call the most technically advanced solution a smart computer (smart computer), along the lines of the smartphone. Csaba Gaál, as a doctor, also raises serious moral questions about artificial intelligence:

“who ultimately decides to what extent the machine is responsible for its decisions, what the doctor should accept, and if he is wrong, can he rely on the computer? What if the machine is right? What does the machine know better than the human being who gives the instructions? Is its role supportive, auxiliary or autonomous, replacing the doctor? Is the legal background regulated?”

The questions raised in medicine resonate across all fields.

2. Technological changes in speech and writing

The anthropological evolution of human language has entered a new era. At some point – perhaps 1 million years ago or 500,000 years ago, nobody knows exactly – articulate human speech began to emerge from basic needs, from pointing, from primordial sounds, from musical expressions, and its main characteristic is abstraction, the ability to convey meaning. At least a million years of evolution, perhaps 100 to 50 thousand years ago, led to articulate speech comparable to what we have today. Not so long ago, perhaps 5-7 thousand years ago, the possibility of recording language, of writing, arose in connection with memory and the preservation of tradition. First, objects and concepts were drawn, this was pictography, and then, through a process of abstraction, we arrived at syllabic and alphabetic writing. The spread of literacy took thousands of years and was given a huge boost by what McLuhan called the Gutenberg Galaxy. And although the fight against illiteracy has been an almost universal agenda since the Enlightenment, even today 20–30% of the world’s population is illiterate, i.e. actually or functionally illiterate.

Let us not forget: the process of writing down speech was a thousands-of-years old one.

In the second half of the 19th century, it became possible to transmit speech (telephone) and record it (gramophone); but it was only in the 20th century that these techniques were refined and massively disseminated. Transmitted and recorded speech is no longer natural speech, as defined by Walter Ong (1982), but so-called secondary verbalism. This is because it does not require presence (telephone) and the same time (recorded speech); it is also characterised by a different speech strategy.

Following the example of secondary literacy, I have called the new literacy that emerges spontaneously during the informatics turn and develops as a result of the new literacy. This created a system that seemed complete. But even as I looked at it, I could already foresee that other forms of linguistic-communicative existence might emerge. With the impact of artificial intelligence, the manipulation of speech and writing, i.e. speech and writing created by machines, represents a completely new quality. The name for this has not yet been coined; for reasons of systematicity, I have previously suggested the term tertiary speech and tertiary writing (Balázs 2023: 85, 102, 129). This is how the whole system looks like:

(primary) BESZÉD	(primary) WRITING
secondary (transmitted, recorded) speech	secondary (new, spontaneous) writing
tertiary (artificial) speech	tertiary (artificial) writing

Before defining this tertiary speech and literacy, which is our subject, let us look at the secondary verbal and secondary literacy that precedes it and which we all already experience.

Secondary verbality

Walter Ong calls the verbalisation of the telephone, radio and television secondary verbalisation. This is a new phase in the history of communication technologies, i.e. it follows primary orality, literacy and thus secondary orality (Ong 1982: 136). Primary orality is the use of spoken language, which is lived and spoken, and which takes place here and now, either perceptually or between present partners, and is characterised by monologue and dialogue. Secondary orality is dialogue with the interposition of technical devices. Its first forms were the telephone, the phonograph (recorder and player), the gramophone (record player), then the tape recorder and the cassette recorder, and, in the field of broadcasting, the telephone announcer, radio, television

and, later, their digital versions, mobile communication. Secondary orality can be entirely spoken (telephone), but it can also be the reading, 'reciting' or 'retelling' of written (or digitally recorded) text. The 'voicing' of written text, i.e. reading it aloud, is a technique that needs to be learnt. In secondary literacy, the technical device used influences – limits or extends – the possibilities of communication. The communicative-linguistic problems of secondary literacy: unnatural pronunciation due to reading, mispronunciation. Whereas primary speech is natural and spontaneous, secondary speech is often planned and contrived. This includes the use of spoken language in the media, where the problems of pronunciation and phrasing of text are many and varied and different from those of spoken language, which is spontaneous.

Secondary literacy

Digitisation (the IT revolution) at the end of the 20th century gave rise to secondary literacy. It has its origins in the typewriter and its derivative, the typewriter, but it is really the computer literacy of the information age that is included here: written 'language' used primarily in private or in social media, using a computer and word processing software, or on mobile phones and smartphones. A specific and dominant technology between 1995 and 2015 was the sending of short text messages (SMS) via mobile phones. I have called this literacy, based on Walter Ong's (1982) system, secondary literacy (Balázs 2005: 38–9). Secondary literacy developed independently of the official, accepted, standard (in our case: academic) literacy. It first appeared in children's short text messages (SMS) on the phone, and one might have thought that the technological limitation (160 characters in SMS, 140 and then 280 characters on Twitter) was the reason for the development of this abbreviated language. But later it appeared in almost every private and sometimes more public genre of the Internet: chat, msn, blog, post, comment, reel, snap ('disappearing images', a briefly visible image or text), tick-tock (short, mostly humorous comedy, dance, dubbing and talent videos), Internet meme. Today, in communities and subcultures large and small (e.g. internal corporate mailing lists), it is the only form of writing used by generations of adults in the IT world. As it becomes more widespread, it is gaining ground and in many cases competing with official (academic, school) spelling and orthography, in effect trying to replace it.

Secondary literacy is not simply bad, broken literacy. If it were bad, badly written, it would not serve understanding. It is possible that the quoted texts are not understood by everyone, but in the community, they can serve communication without hindrance, i.e. they can lead to mutual understanding.

If this is the case, it also means that this writing system, this system of signs, is governed by community rules (let us call it a kind of grammar, not a formal, non-written rule). In other words, simplified, distorted forms of language are created according to linguistic rules agreed upon by the community, and these rules are known to the communicating parties. Otherwise it is not possible for them to understand each other. Secondary literacy is multi-layered: some texts are close to traditional literacy, while others are completely different, using unusual punctuation and abbreviations. The essence of secondary literacy is therefore: a new functional literacy that develops spontaneously through rules. What is interesting is that today's children acquire it faster and more quickly than so-called first (school) literacy. Secondary literacy is now the literacy learnt as first literacy, and for many it is the only literacy, because primary (regular) literacy is not acquired.

3. Towards tertiary verballity

The basic idea and prototype for tertiary speech was the talking machine created by Farkas Kempelen in 1791 (URL1). But the real talking machine would not appear for almost another two centuries, until the information revolution (1990s). After the experiments, the first public trial was the digiton (digital loudspeaker), which could be heard in railway stations. This was followed by chatbots and other speech-generating programs.

The chatbot, or talking (chatting) computer, is a step towards independent text generation, and its literary predecessor may be the part about the Bulgarian guide in Dezső Kosztolányi's novel *Esti Kornél*. The protagonist does not speak Bulgarian, but he has a long (machine-like) "conversation" with the Bulgarian guide.

The possibility of computer-generated speech (speech synthesis) is now considered the third form of verbal expression. There are two versions of this: one uses an existing set of input, the other "interprets" and composes the text itself. Synthesised speech, i.e. speech read out by a machine, suffered from 'teething problems': the prosodic (musical) properties of the text were generally not valid, the stress and pronunciation relations were not adapted to the message, a given sentence always sounded the same, there was no discernible variety. A well-known and typical example is MÁV's so-called digiton (digital voice-over) system, in which the computer assembles separate, isolated accent elements, making the text jumpy and interrupted. (When a particular piece of text is no longer in the set, the live voice is always heard). However, these errors are now eliminated by more sophisticated

programmes. Automated text reading systems can be used for a variety of purposes: telephone customer service, reading news, reading PDFs, reading on-screen or audio (accessibility for the blind and partially sighted, but audio book functionality for everyone), talking clocks, kitchen scales, maps, and even newer cars can give drivers voice guidance. The latest online multi-voice text reader, which can read text in Hungarian and other languages, is an artificial intelligence voice synthesiser that uses natural-sounding voice generators to convert text to speech and can even add narration to video material URL2).

Tertiary verbosity includes live word translators. Programmes are already available (e.g. built into mobile phones) that can translate from one language to another instantly, online; that is, between two people who do not speak each other's language, each can interpret what the other is saying in his or her native (or best known, preferred) language. One of our earlier predictions seems to be coming true: online translation tools can replace language learning – which can save a lot of time and convenience, but can also lead to a loss of exposure to 'other' ways of thinking, worldviews and (perhaps) people.

4. Towards tertiary literacy

A text generator uses artificial intelligence and complex algorithms to create written text. The text generator focuses on the key concepts and words in the text. The essence of text generation is that the initiator starts a text and then allows the natural language processing model to generate the rest of the text, of any length, in the spirit of the initial input. Another option is to transcribe live speech. Today's commercially available text generation programs can transcribe live speech and even read it out later. So they write and speak (tertiary literacy and verbalism in one).

Computer-generated text is a new form of literacy, tertiary literacy. In 2023, chatGPT (or chatgpt in Hungarian), which can generate text in response to given questions and instructions, bursts into the public consciousness. The first attempts are sometimes amusing, but more often deeply thought-provoking. Sometimes the chatGPT produces only clichéd, sketchy, misleading or meaningless text due to some incorrect data or search; it may even politely inform you that its knowledge is limited to certain things and time periods. At other times, however, it is quite accurate and to the point. Some people believe that text generation will become more widespread in practical genres. The programme can easily extract the gist of a long text, making it suitable

for drafting and excerpting (memos and minutes may be in high demand). A text generator works best when the genre is as general as possible. You're said to have written an opening speech for an exhibition for a minister (and the people attending the opening didn't find out). Everywhere you go, you hear that a number of school and university seminar papers (perhaps even theses and dissertations) are suspected. Of course, text-generating programs include translation programs, which have long been developed separately.

The program can write in the style of a given author: for example: prose in the style of Paul Graham, drama in the style of Shakespeare, Wikipedia article in the style of a given author (URL3). The creativity of the programs is shown by the fact that they have written the XIIth and the half of Madách's Tragedy. (Note that Frigyes Karinthy has also written the 16th colour in the traditional, i.e. primary, writing style URL4). The roles are Adam the engineer, Eve the kindergarten teacher, Lucifer the marketing consultant. The conversation is about sceptical people being persuaded by Lucifer about the omnipotence of artificial intelligence, but they refuse to be seduced. Eventually, the Lord sorts out the chaos, assigns tasks, and the humans continue to seek the right path. The 12th and eleventh colours indicate a future which may lead to a time when man will not be needed as much as he was in the Phalanx, since the machine will do all the useful work for him.

Adam:

Mark my words, artificial intelligence,
How many areas it can help.
Machine learning in medicine,
It helps diagnose rare diseases these days.

But it's not just in health that we can see the benefits,
We can also make progress on energy efficiency.
Resource use optimised by MI¹,
It can help protect our environment, oh great treasure. [...]

Lucifer:

Adam, can you hear what I'm saying?
Energy, medicine are just icebergs,
Put the machines in complete control.

¹ AI – artificial intelligence

Just think of the transport, how wonderful it would be,
 If our leadership could be entrusted to artificial intelligence.
 It would filter out human errors, accidents,
 We would sacrifice for the safety of our lives (URL5).

5. Partial summary

The evolution and stratification of human communicative life forms is illustrated in the following table. Speech developed first, followed by writing to record it. Speech is primary, writing is always secondary to speech. Both language and writing are stratified. Speech was spontaneous, colloquial and rhetorical; writing was the same, but the process was rather reversed. In the 20th century, communication technologies led to the emergence of secondary orality and secondary literacy; in the 21st century, today, artificial intelligence is leading to the emergence of tertiary orality and tertiary literacy.

1 M –100 E year	SPEECH (primer)	
5000-500 years		WRITING (secondary)
	1. spontaneous, natural ways of speaking	1. elevated, artistic writing
	2. colloquial speech	2. colloquial writing
	3. elevated, artistic, rhetorical speech	3. lower case scripts
	technological change/digitalisation/information technology	
20 th c.	secondary verbosity	secondary literacy
21 st c.	artificial intelligence	
	tertiary verbosity	tertiary literacy

6. Artificial intelligence in the justice system

In the light of the above, it is worth taking a brief look at the state of AI in the justice sector. This study can be very instructive because of the central role of text processing in this area, which affects a wide range of society.

Currently, one of the most controversial issues related to the use of AI in the legal field is whether AI can be used in court proceedings. One of the main arguments of opponents is that the use of AI violates the right to a fair trial (Dymitruk 2019: 27). Opponents' arguments are often motivated by fear and lack of knowledge, and their arguments remain at the level of generalities, which shows that the majority is not aware of the actual

principles of operation of AI systems currently used in the justice system. Uncritical proponents of the use of MI in the courts focus on the efficiency gains that it can bring and argue that in order to achieve efficiency gains, at least initially, the risk of violating a fundamental right, such as the right of access to justice, must be taken (Roth 2016: 1–48).

6.1. Current AI software used in the justice system. The work of the judiciary consists of well-defined subtasks. In the courts, the allocation of cases is organised by type of case, i.e. a given judge usually hears the same type of cases and thus follows a well-established procedural protocol (seeking answers to the same questions) for a large part of the cases. The most time-consuming part of judicial work is searching through previous decisions, finding and processing case law. These are essentially tasks of text analysis and processing, for which AI can be of considerable assistance to all actors in the judiciary.

The AI software currently used in courts can be divided into three groups (Kálmán–Kiss–Szentgáli 2022):

1. AI software to support certain aspects of judicial work,
2. AI software that makes suggestions for a judge's decision, but does so under human control, and the presiding judge can easily disregard its suggestions,
3. AI software that only makes procedural decisions autonomously, but does not decide on the merits of the case (e.g. in constitutional court cases, human rights court cases, admissibility is decided autonomously according to criteria previously defined by judges).

ExpertUS, used in Mexico, is one of the oldest pieces of AI software used in the justice system. The software is used in family law cases to determine whether a person is entitled to alimony and, if so, how much. To do this, the software uses previous family law case law, i.e. it compares the facts of the application in the case with the facts of cases that have already been decided and, based on the results of previous applications, suggests whether the circumstances that justified the award of maintenance in previous cases are present in the current case. In determining the amount, it will also take account of previous case law and compare it with the submissions made by the parties in the proceedings and in the application in the case (URL6).

The Colombian Constitutional Court receives around four thousand applications for guardianship each year. The Constitutional Court is overwhelmed with the task of reviewing these cases, making preliminary recommendations on admissibility and whether the legal requirements for

guardianship are met. To address this situation, an AI system called PretorIA has been developed. Based on previous court decisions in guardianship cases, a set of 33 criteria has been developed, as many as possible of which are necessary for admissibility and for the court to grant the application. The tool interprets the texts of previous judgments and compares how many of the 33 pre-defined criteria are met in the pending case and can categorise them.

The Brazilian justice system uses an AI called Radar. Although a detailed description of the software is not available, it is known that Radar, like PretorIA, interprets the text of the pleadings received by the courts and deduces the type of case and the applicable law based on the terms used in the pleading (URL7). The tool analyses the previous case law and makes a recommendation to the judge in charge of the case, which the judge can accept or ignore at his discretion.

Malaysia has recently been at the forefront of the use of AI in judicial decision-making. Extensive development has been undertaken to use AI to carry out certain parts of the judiciary's functions. The tool made its debut in February 2020, but at that time it was only used in cases involving drug possession and sexual offences (Juriah–Shurkiah 2020).

The tool is called the Artificial Intelligence Sentencing System (AISS). The first manual for courts on how to use the tool will be available in 2021, and work has begun to expand its scope to include crimes other than those mentioned above, such as theft, traffic offences, and others. The AISS can be used in criminal proceedings to calculate the sentence to be imposed. The developers of the tool have created a questionnaire to enter relevant information about the defendant (e.g. age, marital status, social background, housing, income, criminal/non-criminal record). The tool then calculates the sentence based on previous case law. The judge is not obliged to use the MI tool in the trial and, if it is used, is not bound by it in reaching a decision.

6.2. AI experiment in Hungary (Orosz–Csányi–Nagy 2021) In Hungary, the company Montana carried out a study in which they tried to classify almost 170,000 court decisions into legal categories, i.e. different types of cases. To do this, they first developed an index of 170 different categories of cases, on the basis of which the cases were classified. The complexity of the task was that a decision could have multiple labels and be classified in up to 3 or 4 case categories. A dockerised artificial intelligence application was used to solve the problem. Since a decision could be given multiple labels, this categorisation was part of the multi-label classification task. As a first step, the Montana researchers used binary classifiers, i.e. they created a separate label classifier for each label. They chose this simple approach because none

of the decisions had labels, so the first 'training set' had to be 'taught' by the researchers themselves using regular expressions and manual checking. A method based on term frequency (Term Frequency – Inverse Document Frequency [TFIDF]) was used to convert the texts into mathematical values. The method consists of counting the frequency of each term in the whole document (the number of times a term occurs in the whole text). This so-called Bag of Words method does not check whether a given word is a keyword, but only provides information about the relevance of a given term for the whole text based on its frequency of occurrence. The term frequency is a ratio that shows the proportion of occurrences of a given term in relation to the total number of terms in the document. So if a word or phrase occurs 15 times in the whole document and there are 500 different words/phrases in the document, then the frequency of the phrase is: $15/500=0.03$. During the "teaching" process, 4 different machine learning algorithms were used and their results were compared in the research. The performance of these algorithms was compared for the "dismissal" case documents by examining word combinations, word pairs and combinations of these. Based on these results, the researchers concluded that the best performing AI system was the Support Vector Machine AI system with a linear kernel.

The lesson learned from the research was that a sufficient number of student data is needed to develop a case categorisation software that works effectively.

Overall, it can be concluded that the current state of the art is far from a so-called robotic justice system. Instead, AI systems are being developed and deployed that participate in the work of the judiciary under constant and strict human control, i.e. they do not replace humans or take over their role (Scherer 2016). Therefore, the current presence of AI systems supporting the work of the courts in various ways is typical of the judiciary (Zódi 2018). All AI tools essentially perform text analysis based on previous case law, evaluating the text of the pleading and the statements made in the proceedings according to a set of predefined criteria in a given case type. They can categorise the terms used in the pleading by case type, infer the law applicable to the proceedings from the previous case law, extract the most relevant previous decisions in the case and use them to suggest the decision to be taken.

7. Possible consequences

Technological innovations have always had an impact on man; and in all probability they have influenced evolution. We speak of biological evolution,

then of cultural (linguistic) evolution, which is only a stepping stone for man; and some believe that a third level is taking shape, with unpredictable consequences: planned evolution. “Planned evolution is not intelligent evolution by creation, but evolution manipulated, controlled, etc. by man-made artificial intelligence,” I wrote earlier (Balázs 2023: 22–30).

Levels of evolution:

planned?
cultural
biological

Mankind has not only used the inventions and techniques it has created, it has also abused them. Therefore, there is good reason to believe that the current phase of history is the most dramatic change humanity has ever experienced. As György Csepeli (2023: 14) cautiously puts it, “If humanity fails to recognize the evolutionary opportunity in the artificial intelligence applications made possible by the new info-communication technology, then undesirable changes in social and natural reality will add up, and if the interactions are very unfortunate, then relapse is inevitable.” Örs Szathmáry (2023; URL8) is more pessimistic. The total collapse of humanity is imminent, and if this happens “the world we have created around us will cease to function, and the world we are sinking back into will be some strange mixture of the Stone Age and the Middle Ages” (URL7).

The new language and literacy (possible synonyms: generated linguistic, smart linguistic?) can create a new situation in text production, text learning, text transmission; its legal consequences (unforeseeable), but it raises, for example, the question of authorship, which can lead to dozens of other questions (by the way, there were and still are “authorless”, i.e. community authors in culture). The problem of original and non-original writing extends to all pedagogy, from primary school to university and adult education. The new linguistic generation raises the question of the disappearance of certain professions. There will be no need for secretaries and many administrative tasks, which may not be a problem, but journalists, even writers, poets and dramaturgs will be in trouble, and the cultivation of many artistic disciplines (visual arts, music) may also be called into question. The first book of Hungarian poetry created by artificial intelligence has already been published. Will AI write literature? And if we, as modern intellectuals, accept machine-generated documents, writings, even works of art, will we not be overcome by our human nature’s tendency to superficiality, evasion and thoughtlessness?

Since text production is part of the human thinking mechanism, it can have serious (tragic) consequences for thinking itself. Are we losing our minds? How many non-writing, non-reading people take in texts on some level, but only superficially, and cannot produce them. The consequences are as follows: a loss of understanding, difficulty in expressing emotions, an inability to deal verbally with conflict, perhaps a surrender of our very humanity, a return to a kind of pretextual world, perhaps prehistoric...

It is important to emphasise what must remain human. To quote the Nobel Prize-winning biologist Salvador Edward Luria, Elemér Hankiss (2014: 420–21) says: “It is as if evolution had written into the human brain a basic program that opens up to man the most intimate sources of optimism: the arts, pleasure, hope, comradeship, the self-confidence of the spirit and the noble arrogance with which he throws himself into this unprecedented human adventure”. György Csepeli (2023: 21) also stresses that what remains for man is creativity, imagination, self-expression, art and, as Hankiss says, “the belief that human life is what has and will have meaning”.

“To create something out of nothing, to build constellations out of symbols, perhaps out of nothing, to build a world of freedom, reason and dignity in a silent and empty universe: that, I believe, was a work worthy of man. All this is closely connected with language, and not only with language, but also with the so-called non-verbalisable tacit knowledge” (Mihály Szívós 2017, following Mihály Polányi, Balázs 2018a, b)

which, according to our current knowledge, cannot be the domain of artificial intelligence (since it cannot be verbalised, cannot be algorithmised). In conclusion, the world of emotions, fantasy, creativity, creation, joy, happiness, faith and the arts remain the world of man.

And if we consider that language, even according to traditional, mechanical theory of communication, has at least three main and three secondary functions, it is hard to imagine that artificial intelligence can take over all of them. Among these functions, artificial intelligence is only limited in its ability to perform the referential (reality description), conative (calling) and fatic (contact) functions. For example, it can greet more politely than a human, it can obviously inquire, infer and react to certain states of mind from the answers, but in the absence of human empathy and sensitivity it remains artificial (even if the human party “believes” that the relationship is real; just as it believes the “realities” of the virtual world). The question for me is to what extent artificial intelligence is capable of performing the emotive (expression of emotion), the poetic or aesthetic, and the meta-

linguistic (code-referencing) functions. And then we have not yet mentioned such basic functions introduced by anthropological linguistics as: reflective, heuristic, imaginative, creative (play), sacral, velative (concealing) (Balázs 2022: 63–68). This is because language is not just the communication of information, but (human) behaviour, with its complexity, unpredictability and never perfect comprehensibility. It is rarely heard from a linguist, but it should often be said that language is not just a means of communicating ideas, and that linguistic meanings are so complex and variable that perfect understanding is not even possible; it is a human miracle that we somehow understand each other through imperfect communication. That is, if we do.

It is worth recalling McLuhan's earlier warning that 'it can only be a disaster if we fail to recognise the causal links and effects of the technologies we create' (Benczik 2001: 331).

References

- Balázs, Géza 2005. Az internetkorszak kommunikációja. (Communication in the internet age.) 25–57. In: Balázs Géza – Bódi Zoltán (szerk.): *Az internetkorszak kommunikációja*. Tanulmányok. Gondolat/Infónia. Budapest.
- Balázs, Géza 2018a. Implicit tudás, anyanyelvi hallgatólagos tudás – Szívós Mihály könyve Polányi Mihály elméletéről. (Implicit knowledge, tacit knowledge in the native language – Mihály Szívós's book on the theory of Mihály Polányi.) *Magyar Nyelvőr*, 2: 238–42.
- Balázs, Géza 2018b. A hallgatólagos tudás újragondolása. (Szívós Mihály könyvéről.) [Rethinking tacit knowledge (On the book by Mihály Szívós).] *Magyar Tudomány*, 8. 1259–61. <https://doi.org/10.1556/2065.179.2018.8.16>
- Balázs, Géza 2022. *Folklór és nyelv. Folklór, művészet, irodalom, nyelv. Folklórlingvisztikai tanulmányok*. (Folklore and language. Folklore, art, literature, language. Folkloric linguistic studies.) IKU, Budapest.
- Balázs, Géza 2023. *Az internet népe. Internet – társadalom – kultúra – nyelv. A kulturális és a tervezett evolúció határán*. (The people of the internet. Internet – society – culture – language. At the frontier of cultural and planned evolution.) Ludovika Egyetemi Kiadó. Budapest.
- Benczik, Vilmos 2001. A Gutenberg-galaxis margójára. (On the margins of the Gutenberg Galaxy.) 319–31. In: McLuhan 2001.
- Csepeli, György 2023. *Értékek ébresztése*. (Awakening values.) Kocsis Kiadó, Felsőbbfokú Tanulmányok Intézete, h. n.
- Dymitruk, Maria 2019. The right to a fair trial in automated civil proceedings. *Masaryk University Journal Of Law And Technology*, 13(1). 27–44. <https://doi.org/10.5817/MUJLT2019-1-2>
- Éksz. = *Magyar értelmező kéziszótár*. (Hungarian Interpretative Dictionary.) Főszerk.: Pusztai Ferenc. Akadémiai. Budapest, 2003.

- Gaál, Csaba 2020. Újabb félrevezető fogalom: "mesterséges intelligencia". (Another misleading term: 'artificial intelligence'.) *Magyar Orvosi Nyelv*, XX(1). (2020. augusztus) 31–4.
- Juriah, Abd Jalil – Shukriah, Mohd Sheriff 2020. *Legal tech in legal service: challenging the traditional legal landscape in Malaysia*. 28 *IJUMIJ* 279. [https://doi.org/10.31436/ijumlj.v28i\(S1\).586](https://doi.org/10.31436/ijumlj.v28i(S1).586)
- Hamvas, Béla 1997. *Karnevál*. Regény. (Carnival. Novel.) 1–3. rész. Medio Kiadó, Szentendre.
- Hankiss, Elemér 2014. *Az emberi kaland. Egy civilizációelmélet vázlatja*. (The human adventure. An outline of a theory of civilisation.) 5. bővített kiadás. Helikon Kiadó. Budapest.
- Kálmán, Kinga – Kiss, Laura Olga – Szentgáli-Tóth, Boldizsár Artúr 2022. Mesterségesintelligencia-alapú szoftverek a világ bíróságain: gyakorlati tapasztalatok, perspektívák és kihívások. (Artificial intelligence-based software in the world's courts: practical experiences, perspectives and challenges.) In: *Jogi technológiák*. Ludovika Press. Budapest. ISBN 9789635318193
- Lábos, Elemér 1979. *Természetes és mesterséges értelem*. (Natural and artificial intelligence.) Magvető Kiadó. Budapest. (Gyorsuló idő)
- McLuhan, Marshall 2001. *A Gutenberg-galaxis. A tipográfiai ember létrejötté*. (The Gutenberg Galaxy. The birth of typographic man.) Trezor Kiadó, Budapest.
- Ong, Walter J. 1982. *Orality and Literacy. The Technologizing of the Word*. Methuen. London – New York.
- Orosz, Tamás – Csányi, Gergely – Nagy, Dániel 2021. Mesterséges intelligenciát alkalmazó szövegbányászati eszközök készítése a distiller keretrendszer segítségével, Jogi szövegek automatikus feldolgozása. (Design of text mining tools using artificial intelligence with the distiller framework, Automatic processing of legal texts.) In: *Számítástechnika és Oktatás* (SzámOkt) XXXI. 62–8.
- Roth, Andrea 2016. Trial by machine. *Georgetown Law Journal*, 104.(5). 1–48. 48.
- Scherer, Matthew U. 2016. Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law & Technology*, 29(2). 1–48.
- Szabó, Márton 2022. *A mindennapi élet poétikai valósága. Társadalomtudományi értekezés*. (The poetic reality of everyday life. Social science thesis.) Osiris Kiadó. Budapest.
- Szívós, Mihály 2017. *Fordulópontok a hallgatólajos tudás és a tudattalan felfedezéstörténetében. A hallgatólajos tudás általános elmélete*. (Turning points in the history of the discovery of tacit knowledge and the unconscious. The general theory of tacit knowledge.) Magyar Szemiotikai Társaság, Loisir Könyvkiadó. Budapest. (Magyar szemiotikai tanulmányok 40.)
- Szűts, Zoltán 2020. *A digitális pedagógia elmélete*. (The theory of digital pedagogy.) Akadémiai Kiadó. Budapest. <https://doi.org/10.1556/9789634545859>
- Zódi, Zsolt 2018. The code of AI and human laws. *Acta Universitatis Sapientiae, Legal Studies*, 8(2). 253–268. 253, 260.

Internet resources

- URL1. <http://www.kempelen.hu/beszelogetp.html>
URL2. <https://www.narakeet.com/>
URL3. Andrej Karpathy: The Unreasonable Effectiveness of Recurrent Neural Networks <http://karpathy.github.io/2015/05/21/rnneffectiveness/>
URL4. <https://www.arcanum.com/hu/online-kiadvanyok/Szovegyujtemeny-szovegyujtemeny-1/anyugat-nemzedeke-6F97/karinthy-frigyes-1887-1938-982D/igy-irtok-ti-valogatas-983A/magyardrama-9968/madach-imre-9AD8/tizenhatodik-szin-9AD9/>
URL5. <https://www.tmit.bme.hu/m-i-az-ember-tragediaja>
URL6. Cáceres, Enrique 2008 EXPERTIUS: A Mexican Judicial Decision-Support System in the Field of Family Law. In Conference: Legal Knowledge and Information Systems – JURIX 2008: The TwentyFirst Annual Conference on Legal Knowledge and Information Systems, Florence, Italy, 10–13 December 2008, and Artificial intelligence and the judicial system in Latin America. World Today News (2018).
URL7. Augusto, Eduardo R. 2021. AI, Machine Learning & Big Data Laws and Regulations 2021 – Brazil.
URL8. “Correct calculation suggests that the total collapse of humanity could come around 2050” <https://24.hu/tudomany/2023/05/15/szathmary-eors-klimavaltozas-evolucio-civilizacios-osszeomlasinterju/>

Balázs Arató

Habilitated Associate Professor

Károli Gáspár University of the Reformed Church
in Hungary

E-mail: arato.balazs@kre.hu

<https://orcid.org/0000-0001-5854-8018>

Géza Balázs

Professor

University of Theatre and Film Arts (Budapest)
Partium Christian University Department of Language
and Literature (Oradea)

E-mail: balazs.geza@szfe.hu

<https://orcid.org/0000-0003-3440-2959>

**DEVELOPMENT OPTIONS FOR ARTIFICIAL INTELLIGENCE SUPPORTED
INTELLIGENT TUTORING SYSTEMS VIA THE INTEGRATION OF THE CHATGPT**

Abstract

In the era of the second revolution of intelligent tutoring systems, the integration of artificial intelligence can bring about a learning process adjusted to the personal needs and preferences of students, including language teaching, since AI-based systems can assist in identifying and addressing individual language skills and deficiencies. While these systems can be divided into four inherently AI-based components (learner model, domain model, pedagogical model and user interface) facilitating the embedding of cognitive functions, the options hidden in the integration of big language models have not yet been fully utilized. The first step in the integration of ChatGPT developed by OpenAI is the thorough identification of the goals of the education process followed by the allocation of the appropriate API (Application Programming Interface) service to the given process (question and answer, content summary, sentiment analysis). At the same time it is worthwhile to improve the ITS architecture to maximise the ChatGPT provided options. One result of our research and development effort is a further improved architecture integrating an additional component, the central organizer. This unit is responsible for the matching or allocation of the ChatGPT's API parameters to the given educational content and objective. The integration of ChatGPT into intelligent tutoring systems poses numerous challenges such as the protection of personal information, the related technological problems along with the clarification of ethical issues (Ara-tó-Balázs 2023). Addressing such concerns is necessary for the realization of our original objective guaranteeing that the respective technological achievement would effectively support the implementation of said goals.

Keywords: artificial intelligence, language models, human cognition, libraries, linguistic interaction, language theory

Introduction

Modern and intelligent tutoring systems (ITS) combining the latest results of Artificial Intelligence research with the pedagogical field can revolutionize the education process. The Intelligent Tutoring System is a computer system designed to influence the instruction process and improve learning outcomes in an interactive manner. These systems provide personal mentoring by taking the individual needs and abilities of students into consideration. Intelligent Tutoring Systems rely on artificial intelligence to dynamically adapt the

methods and content of the instruction process to the learning style and pace of students while immediate context-dependent feedback eventually leads to a personalized learning experience.

While scientific exploration of various dimensions related to the use of ChatGPT for the support of the education process has already begun, its application in Intelligent Tutoring Systems can facilitate a deeper understanding of the given options (Rajcsányi–Molnár et al. 2024b).

The integration of AI achievements creates new opportunities not only globally, but also in the teaching of the Hungarian language. Our study provides a brief overview of the operation and structure of Intelligent Tutoring Systems along with introducing various educational and pedagogical options. Next, we look at the roles the ChatGPT can play in this framework and consider the advantages and potential dangers related to its integration.

The evolution of Intelligent Tutoring Systems

Intelligent Tutoring Systems already have a significant history as they originate in the 1960s (Carbonell 1970; Wescourt et al. 1977; Nwana 1990; Kulik–Kulik 1991; Guo et al. 2021). At this time such instruction systems exclusively focused on providing a more or less adaptive interactive and customized learning environment for students. This goal was achieved via an *if-then* construction based programming matching the individual knowledge level of the learner. (One such example is the Rule-Based Reasoning paradigm utilizing general information related to the given topic [Anderson et al. 1995]).

After a quarter of a century of spectacular development represented by machine-based learning and the rise of the Case-Based Reasoning (CBR, [Manouselis et al. 2014: 102]), the next generation of these systems offers the option of making conclusions based upon previous experiences, while the preservation of previously deployed methods helps in solving newly emerging challenges (see the Retrieve, Reuse, Revise, and Retain method (Aamodt–Plaza 1994: 8). The given findings are useful in solving the newer problems by taking into consideration the role of memory, piece sequencing, analogue transfer, reasoning and problem solving techniques in the knowledge acquisition process (Koedinger et al. 2012).

After the passage of another 25 years the role of AI surpasses the various automation concerns in the current ITS. Such systems contain four clearly distinguishable AI-based components. The continuously updated **learner model** stores information on the student's knowledge, skills and learning methods gained from interaction. The **module** facilitating **domain**

knowledge contains the educational materials and relevant concepts arranged in a structured manner. The **pedagogical component** manages the learning process and provides feedback to the learner along with harmonizing the modifications indicated by other components and selecting the optimal learning strategy. The **user interface module** guaranteeing interactivity provides a communication channel between the learner and the ITS enabling students to navigate within the system (Balázs L. 2023). Chart 1 describes the schematic structure of the ITS.

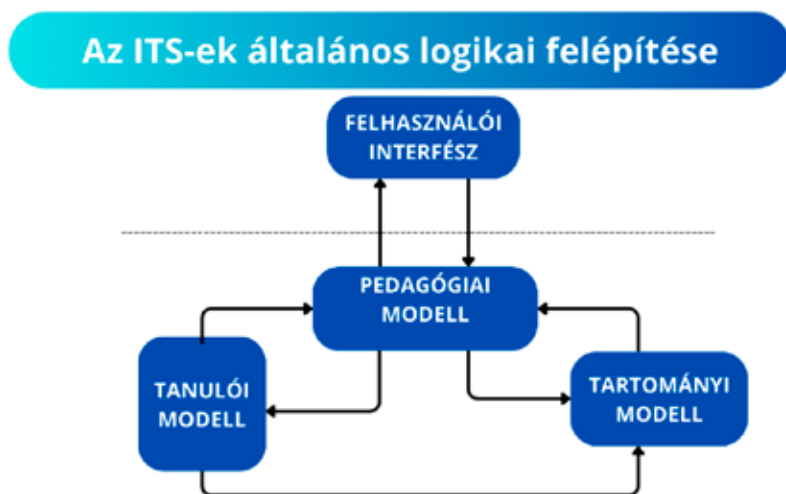


Chart 1. (Toldi 2023) The general logical architecture of the ITS
Learner model, Domain model, Pedagogical model, User interface.

The AI-based structure facilitates the integration of such advanced cognitive functions as the big language models, natural language processing, the new generation of machine-based learning technology, and various explainable AI solutions. (XAI – eXplainable Artificial Intelligence). One such example is the sample recognition process facilitating deeper, context-based adaptivity in processing learning materials. The system takes into consideration individual skills along with actual mental and emotional conditions coupled with monitored didactic and methodological autonomy.

Modern ITS, mainly due to the previously introduced XAI methods can provide a more personalized learning environment by accurately identifying and addressing potential learning problems and make predictions concerning difficulties to be encountered during the learning process (Chaturvedi–Ezeife 2017; Gligorea et al. 2023).

The big language models can provide an opportunity to take advantage of most of the new options. While the ChatGPT and similar generative, pre-trained transformers are ready to redefine the world of ITS, their application poses several challenges and requires careful consideration and analyses before their actual use. Below we describe how can the integration of the ChatGPT be harmonized or reconciled with the multidisciplinary expectations related to their use in ITS.

The significance and role of artificial intelligence in Intelligent Tutoring Systems

As mentioned earlier in the introduction to this section the main attraction of AI-supported ITS is their ability to process and analyse a huge amount of data obtained via machine-based learning including answers by students, learning patterns, and previous performance-related information (Baker–Inventado 2014). Due to such data-driven approach real time learning environments adapted to the needs of the given students can be formed. In such systems the AI is capable of identifying the potential lack of knowledge along with optimizing the learning content and pace accordingly (Gligorea et al. 2023). Consequently, through its natural language processing techniques the ITS can involve students in personalized dialogues while providing explanations, feedback, and guidance adapted to their learning needs and comprehension level. (McNamara et al. 2013). Such a personalized or customized approach not only increases the commitment of students to the learning process, but it facilitates a deeper understanding of various concepts along with promoting the development of critical thinking skills. AI-based systems having processed various learning preferences can generate further simulations and interactive games (Daghestani et al. 2020; Rajcsányi-Molnár et al. 2024; Kőkuti et al. 2023).

The options provided by the integration of the ChatGTP into Intelligent Tutoring Systems

The OpenAI facilitates the integration of the ChatGTP via an Application Programming Interface developed for this purpose. The API services include a wide variety of options focusing on the comprehension of natural language and text generation. As far as the ITS is concerned, the crucial API functions include *text generation*, *text classification*, *mood analysis*, *content summary*, *question and answer*, and *problem solving*. These services also cater to the specificities of the Hungarian language, offering extensive opportunities for more effective teaching of grammar, stylistics, and even etymology (Balázs 1997: 25–9).

The primary prerequisite for the integration of the ChatGTP into a given ITS is the clear and comprehensive definition of the respective instructional goals with an awareness of the capability of the actual objectives to function as potential guidelines for future development efforts.

The forthcoming pedagogical objective system contains elements providing a free choice of options in light of the available material and other type resources. Learning adapted to the individual needs, preferences, and learning pace of the students *increases the intensity of customization*. *Active promotion of persistent student effort* facilitates deeper understanding of the given educational material and its incorporation into the long-term memory. The proper application of the acquired knowledge and the *recognition of potential misinterpretation* of the given information coupled with the provision of *immediate personalized feedback and support* increase the efficiency of learning (Szűts–Szűts–Novák 2023: 566). Equally important feature is the recognition of the learner 's individual characteristics in *supporting various learning styles and strategies*. The elaboration of complex tests and the description of problem solving processes promote *critical thinking* and *improve problem solving skills*, while sharing the management of the knowledge acquisition process through dialogue promotes *self-directed learning and self-motivation*.

The allocation of API-provided services to pedagogical objectives requires the detailed analysis of the API service options. Below we introduce a few examples.

Text generation can help in meeting student needs related to academic progress based upon the acquisition of the respective educational materials along with providing written feedback or evaluation of the results of various assignments and tests.

The appropriate integration of the *question-answer* feature allows the ITS to simulate personal remedial instruction via offering students immediate, concise and accurate answers. As a result of identifying appropriate parameters not only the learner's understanding of the given material can increase, but the generation of lead-in questions can help the improvement of analytical skills, while promoting learning by discovery and critical thinking.

The *content summary* service can be used as a complementary feature to transform longer, even scholarly texts into a concise form while observing appropriate style parameters. This feature can also help students with specific or unique learning styles highlighting and practicing the concepts and essential components of the given learning material even during end of the year reviews.

Sentiment analysis enables the ITS to explore the emotional aspects including satisfaction, frustration, or general commitment levels behind student interaction. Such a feedback mechanism facilitates the processing of adaptive answers by modifying the given difficulty level or sending motivational messages. *We must note, however, that in this case especially, but in the context of all other API functions in order to increase the reliability of the given function the ChatGTP has to be fine-tuned according to the needs of the appropriately prepared student groups.*

The steps of integrating the ChatGTP into Intelligent Tutoring Systems

The first step in the integration of the ChatGTP services is redesigning the ITS architecture in order to assure the continuous cooperation between the respective artificial intelligence components at an optimal level. Thus Chart 2 describes a potential modification of the system.

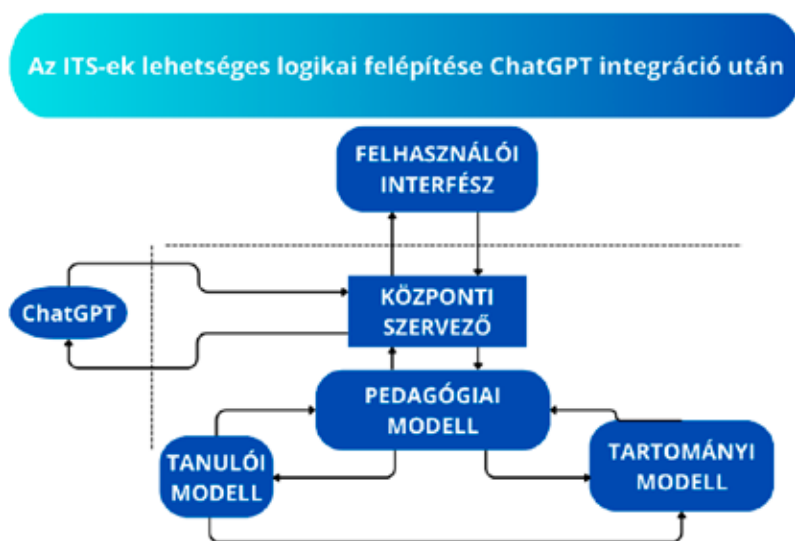


Chart 2. A potential logical architecture of the ITS following the integration of ChatGTP.
(author's own compilation)

User interface, Central organizer, Pedagogical model, Learner model, Domain model.

The new component, the module named *central organizer* has to fulfill a variety of tasks.

Primarily it has to assure the personalization or customization of the Chat GTP's API parameters to the given educational content and goals. In the same vein it has to guarantee the relevance of the respective interaction, while making sure that despite the dynamic aspects of the specific components such options would always harmonize with prospective pedagogical objectives. In order to provide a continuous real time user experience the algorithms of the module must have predictive and scheduling functionality. Additionally due to the need to protect learner data the given system component must have substantial data processing and handling practice while the parameters influenced by learner data analyses have to be continuously modified and updated for the ChatGTP on the API. Integration design must extend to the definition of how the ChatGTP will cooperate or work together with other components of the ITS not directly connected to the system, such as the user surface, feedback mechanisms, or the respective subject module.

In order to identify, eliminate, or correct unavoidable problems and assure continuous functioning and the protection of user data the system has to be thoroughly tested.

Following the successful testing phase, the final stage includes the installment and the monitoring of the operation of the system. This stage is not only required for the continuous operation of the integrated system, but for the optimalization and further development of the intelligent modules in an iterative manner as well. System monitoring includes guaranteeing continuous availability of basic services along with the evaluation and analysis of the behaviour patterns of users, the demonstrated and assessed commitment levels, the respective learning outcomes and the interaction with the system.

The central organization unit has to fulfill most of these monitoring tasks in cooperation with the pedagogical module. During the monitoring process this unit provides indicator numbers for the evaluation of the performance of the given logical components and the specific learning results. Accordingly stakeholders (artificial agents or humans) can evaluate the results in light of the pre-determined objectives. (In order to facilitate human evaluation of the respective indicators specially and purportively incorporated analytical platforms provide fast account or report of the crucial performance indicators in a dashboard-like arrangement [KPIs]. Moreover, they provide a surface for the modification and fine-tuning of the necessary correction parameters).

Besides optimalization the technological, usability-related or potential ergonomic problems hindering the learning process are identified in time via the formulated indicators. (The continuous data collection helps in intensifying the learning process in light of the differing demands of learners, the understanding of learner preferences, thereby supporting the more

accurate personalization and customizing of the given educational material. This task however is controlled by the pedagogical model also responsible for the feedback loops integrated via various learning routes facilitating the fast, point-based, or the more detailed questionnaire-oriented direct monitoring of user satisfaction).

The main challenges related to the integration of the ChatGPT into Intelligent Tutoring Systems

The obvious advantages notwithstanding the integration of the ChatGPT into intelligent tutoring systems implies several challenges. It requires several high level skills along with substantial resources.

The first crucial integration-related challenge is connected with the complexity of the integrated system and the reliability of its operation along with the respective scalability. In addition to the required multidisciplinary (pedagogical and informatics-related) expertise, depending on the number of the actual active users and the capability of accommodating high level user needs, a robust infrastructural background is indispensable. Such equipment is not only essential for appropriately scalable computing capability necessary for the processing and generating of real time answers, but enables flexible handling or solving of technological errors that frustrate the contemporary learning experience. The elaboration and maintenance of such an infrastructure demands a high amount of resources and technological expertise, which can pose a significant challenge for institutions with limited means. This problem can be solved by the acquisition of a cloud-based infrastructure via procurement or government tenders.

The protection of the learners' personal data and the safety of the stored information is a special priority. ITS often store and process sensitive educational data, which requires rigorous observation of data protection regulations and standards in addition to strict and thorough implementation of data protection measures. The unique requirements of the integration of state of the art Artificial Intelligence technologies, i. e. ChatGPT, warrant the elaboration of relevant protocols as well.

Additionally further ethical concerns emerge related to the ChatGPT. Contents or materials provided by generative Artificial Intelligence can include potential bias, or incorrect statements. This gives cause for significant concern as any bias in educational content not only weakens or undermines the effectiveness of the instruction process, but it can promote stereotyping while maintaining and strengthening prejudicial attitudes. A thorough pre-application fine-tuning of the ChatGPT model is indispensable for the elimination of such threat.

Finally, it must be mentioned that in order to guarantee that the respective technological development effectively supports the realization of the set pedagogical objectives the monitoring of performance indicators, continuous pedagogical research and the identification and adaptation of teaching strategies appropriate to the given school environment are necessary. Consequently, teachers and system developers have to work together in finding the potentially most effective ways of integrating the abilities of artificial intelligence into existing instructional framework systems, or redesign the given system if necessary to accommodate new means of learning by Artificial Intelligence (Balázs G. 2023a, b).

Conclusion

Our research focused on the options implied by the integration of the ChatGPT into Intelligent Tutoring Systems. In light of the previously identified pedagogical objectives we presented an overview of technological skills facilitating its integration into existing ITS. Next we looked at the steps required for successful integration and analyzed the related potential challenges. Our findings indicate that the synergy of the ChatGPT and ITS can open new vistas and lead to the identification of significant problems. Additionally, our research highlighted that the integration of ChatGPT can bring about significant advancements not only globally but also in the teaching of the Hungarian language.

The integration of the ChatGPT into ITS can result in major changes in the education process via promoting the improvement of learning and providing adequate support to meeting the individual needs and preferences of students. While the combination of the ChatGPT and ITS has its challenges, it can lead to further developments in digitally supported education.

References

- Aamodt, Agnar – Plaza, Enric 1994. Case-Based Reasoning: Foundational Issue. Methodological Variations, and System Approaches. *AI Communications* 7–1. 39–59. <https://doi.org/10.3233/AIC-1994-7104>
- Anderson, R. John – Corbett, Albert T. – Koedinger, Kenneth R. – Pelletier, Ray 1995. Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences* 4–2. 167–207. https://doi.org/10.1207/s15327809jls0402_2
- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban? a digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 147. 755–76. <https://doi.org/10.38143/Nyr.2023.5.755>

- Balázs, Géza 1997. *The Story of Hungarian. A Guide to the Language*. Corvina Kiadó. Budapest.
- Balázs, Géza 2023a. Tamás Ildikó: „Adj netet!”. *Magyar Nyelvőr*; 147. 120–7. <https://doi.org/10.38143/Nyr.2023.1.120>
- Balázs, Géza 2023b. *Az internet népe*. Ludovika Egyetemi Kiadó. Budapest.
- Balázs, László 2023. Az iskolai kommunikáció egy rejtett dimenziója – szalutogén pedagógiai kommunikáció. *Magyar Nyelvőr*; 147. 692–712. <https://doi.org/10.38143/Nyr.2023.5.692>
- Carbonell, Jaime R. 1970. AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man Machine Systems*, 11–4. 190–202. <https://doi.org/10.1109/TMMS.1970.299942>
- Chaturvedi, Ritu – Ezeife, Christie I. 2017. Predicting Student Performance in an ITS Using Task-Driven Features. *IEEE International Conference on Computer and Information Technology (CIT)*. 168–75. <https://doi.org/10.1109/CIT.2017.34>
- Daghestani, Lamya F. – Ibrahim, Lamiaa F. – Al-Towirgi, Reem S. – Salman, Hesham A. 2020. Adapting gamified learning systems using educational data mining techniques. *Computer Applications in Engineering Education*. 28–3. 568–89. <https://doi.org/10.1002/cae.22227>
- Gligorea, Ilie – Cioca, Marius – Oancea, Romana – Gorski, Andra-Teodora – Gorski, Hortensia – Tudorache, Paul 2023. Adaptive Learning Using Artificial Intelligence in e-Learning: *A Literature Review. Education Sciences*. 13–12. 1216. <https://doi.org/10.3390/educsci13121216>
- Guo, Lu – Wang, Dong – Gu, Fei – Li, Yazheng – Wang, Yezhu – Zhou, Rongting 2021. Evolution and trends in intelligent tutoring systems research: A multidisciplinary and scientometric view. *Asia Pacific Education Review*, 22–3. 441–61. <https://doi.org/10.1007/s12564-021-09697-7>
- Koedinger, Kenneth R. – Corbett, Albert T. – Perfetti, Charles 2012. The Knowledge-Learning-Instruction Framework: Bridging the Science-Practice Chasm to Enhance Robust Student Learning. *Cognitive Science*. 36–5. 757–98. <https://doi.org/10.1111/j.1551-6709.2012.01245.x>
- Kőkuti, Tamás – Balázs, László – András, István – Rajcsányi-Molnár, Mónika 2023. Collaborating with Artificial Intelligence – AI in Business Communication Education. In: *IEEE – Óbudai Egyetem; IEEE (ed.) 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE): Proceedings Danvers (MA)*, IEEE. 287–94.
- Kulik, Chen-Lin C. – Kulik, James A. 1991. Effectiveness of Computer-Based Instruction: An Updated Analysis. *Computers in Human Behavior*, 7. 75–94. [https://doi.org/10.1016/0747-5632\(91\)90030-5](https://doi.org/10.1016/0747-5632(91)90030-5)
- Manouselis, Nikos – Drachsler, Hendrik – Verbert, Katrien – Santos, Olga C. (eds.) 2014. *Recommender Systems for Technology Enhanced Learning: Research Trends and Applications*. Springer. New York. <https://doi.org/10.1007/978-1-4939-0530-0>
- McNamara, Danielle S. – Crossley, Scott A. – Roscoe, Rod 2013. Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods*, 45–2. 499–515. <https://doi.org/10.3758/s13428-012-0258-1>

- Nwana, Hyacinth S. 1990. Intelligent tutoring systems: An overview. *Artificial Intelligence Review*, 4–4. <https://doi.org/10.1007/BF00168958>
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István – Czifra, Sándor 2024a. Competition as an Effective Motivational Tool in Online Education. In: Anikó, Szakál (ed.): *IEEE 11th International Conference on Computational Cybernetics and Cyber-Medical Systems: ICC 2024: Proceedings*. IEEE Hungary Section Budapest. 83–8.
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István 2024b. Online Leadership Training in Higher Education Environment. *ACTA POLYTECHNICA HUNGARICA* 21, 39–52.
- Shaun Baker, Ryan – Salvador Inventado, Paul 2014. Educational Data Mining and Learning Analytics. In J. A. Larusson & B. White (eds.): *Learning Analytics* (pp. 61–75). Springer. New York. https://doi.org/10.1007/978-1-4614-3305-7_4
- Szűts, Zoltán – Szűts-Novák, Rita 2023. A social media és az okoseszközök kommunikációs jellemzőinek hatása a pedagógiára. elméleti alapvetés a kommunikáció és a neveléstudomány kapcsolatáról *Magyar Nyelvőr*, 147. 565–82. <https://doi.org/10.38143/Nyr.2023.5.565>
- Toldi, Lajos 2023. A tanulói motiváció kiaknázása intelligens oktatárendszerekben. In: Lengyelne Molnár, Tünde (szerk.): *Agria Média 2023 és ICI-17 Információ- és Oktatástechnológiai Konferencia: A magas szintű digitális kompetencia a jövő oktatásának kulcsa*. Líceum Kiadó. Eger.
- Wescourt, Keith T. – Beard, Marian – Gould, Laura – Barr, Avron 1977. *Knowledge-based CAI: CINs for Individualized Curriculum Sequencing*. Stanford. California.

Lajos Toldi

PhD Student

Eszterházy Károly Catholic University

Faculty of Pedagogy

Doctoral School of Educational Science

E-mail: toldi.lajos@gmail.com

<https://orcid.org/0009-0007-4483-5500>

Tünde Lengyel-Molnár

Associate Professor

Eszterházy Károly Catholic University

Faculty of Informatics

Institute of Digital Technology

E-mail: lengyelne.tunde@uni-eszterhazy.hu

<https://orcid.org/0000-0001-5887-6900>

THE LARGE LANGUAGE MODELS IN THE SERVICE OF EDUCATION

Abstract

The name artificial intelligence (AI) appeared in the 1950s (McCarthy 1956), but mankind had a much older desire to perform certain tasks with machines. AI permeates our everyday lives, so the widest possible layers of society must acquire knowledge about artificial intelligence. The understanding and acceptance of AI are crucial for social development, equality, and innovation. For these reasons, the Hungarian Artificial Intelligence Coalition was founded in 2018, the main goal of which is to bring technology closer to all actors in society. As a result of the coalition's work, the Artificial Intelligence Strategy of Hungary was published in May 2020, in which one of the important segments of the priority AI research and development directions is the development of language technology. Our research aims to provide a comprehensive overview of the development and functioning of the large language models, focusing on the specific challenges and opportunities that arise in the field of Hungarian language models. In addition, we would like to show how these models can be used effectively in education and how they can help students improve their language skills, support their studies, or even create intelligent systems for educators.

Keywords: artificial intelligence, large language models, natural language processing, digitization, education

Introduction

Perhaps one of the most exciting and dynamically developing fields in the history of mankind is artificial intelligence (AI). In recent decades, and even more so in recent years, technological progress has given impetus to a revolution that is changing our daily lives. Innovations and developments supported by artificial intelligence permeate everything from industrial processes through personal assistants to medicine and the financial sector.

According to Daugherty and Wilson (2018), companies that realize how to utilize AI will gain an advantage, while others will be left behind. The field of AI is changing and developing so rapidly that possibilities that previously seemed unimaginable are now part of our everyday lives. Taking machine

learning¹ and data analysis to a new level, the development of autonomous vehicles, natural language processing (NLP) and robotics are just a few examples of the innovations taking place within the field. Technological innovations not only create new opportunities but also transform industries, workplaces and society. In this rapidly changing environment, AI is not just a scientific or technological field, but a full-scale paradigm shift that fundamentally reshapes the way we think, work and live (Ford 2022). Therefore, it is essential to constantly monitor developments and understand their effects and possibilities. Navigating the world permeated by artificial intelligence is not only a challenge but also an exciting opportunity for all of us.

The rapid development in the field prompted us to review the various stages of AI development. In the present study, as part of secondary research, we focus on natural language processing, as one of the key areas of AI, and we review the usability of these applications in education. The development of large language models will be presented, highlighting the models trained in the Hungarian language.

Historical overview

The history of artificial intelligence goes back to ancient times, the discoveries of many scientists have contributed to the fact that this technology permeates all areas of life today (Buzás 2021). The concept of artificial intelligence (AI) appeared in the 1950s, but mankind had a much older desire to perform certain tasks with machines. This was first described by Aristotle (1969) in his work *Politics*, in which he stated that the device he called an automaton was able to perform the work of slaves. His idea was never realized, but it served as a basis for the development of different types of artificial intelligence. During the development of AI, several eras can be identified, in which different techniques and methods alternated.

The term itself was first used by John McCarthy in 1956 at a summer conference at Dartmouth College. In his reading, the definition of artificial intelligence is:

“Artificial Intelligence aims to create machine systems capable of human-level intelligent behaviour such as perception, linguistic communication, decision-making, and learning” (McCarthy 2007).

¹ In machine learning, algorithms are taught to find patterns and correlations in large data sets and make the best decisions and predictions based on this analysis, without being programmed to do so. A machine learning model can be supervised, unsupervised, semi-supervised or reinforced.

In addition to all this, Searle (1980) talks about weak and strong artificial intelligence according to a different approach. Weak AI refers to systems that act as if they are intelligent, but it cannot be determined whether they have a mind. On the other hand, he uses the term strong AI for systems that think. An important role in the emergence of AI is attributed to Alan Mathison Turing, who in the 1930s developed the mathematical model of the programmable computer (Turing 1936), which was then named after him the Turing machine. A Turing machine can be considered a theoretical model of a simplified computer, not a real machine, but an abstract automaton. Turing's 1950 paper "Computing Machinery and Intelligence" is about the idea that if we programmed a computer so well that it could communicate on any subject on a human level with another human, and that person would not be able to tell that it was actually a machine or talking to a living person, then we should consider this machine intelligent. This concept is known as the Turing test (Turing 1950).

In 1958, Frank Rosenblatt built the Perceptron. The Perceptron represented the development of a neural network, which was an important milestone in the development of artificial intelligence. This algorithm was used to create machines that could recognize and categorize images. Machine learning appeared in the 1960s, the essence of which is that machines learn from data without being specifically programmed to do so. This resulted in the first chatbot developed by Joseph Weizenbaum in 1966 and named ELIZA. Perceptron did not live up to expectations, and researchers recognized the limitations of AI technologies. A nearly 10-year period followed when research in the field received less support. Expert systems appeared in this period.

In the development of AI, the 1990s brought a new turn with the wide spread of the Internet, which made it possible to use large amounts of data for research. In 1997, the DEEP BLUE computer developed by IBM beat the world chess champion Garry Kasparov (Hsu 1999; Newborn 2012). Deep learning² appeared in the 2000s (Goodfellow et al. 2016). Algorithms based on this have been used in many areas, such as image and speech recognition or natural language processing. In the 2010s, developments in the field of neural networks and machine learning appeared, which marked a great advance in the field of natural language processing³ (Khurana et al. 2023) and computer vision. In 2016, the AlphaGo (URL1) program developed by Google beat the then-Go world champion Lee Sedol.

² Deep learning is actually machine learning, which enables the learning of systems with artificial neural networks.

³ Natural language processing is the area of cooperation between AI and linguistics, where natural (human) languages are processed using different computer methods.

The further development of neural network machine learning resulted in new developments in the field of natural language processing. Large language models were published, which then enabled the public presentation of ChatGpt (URL2) developed by OpenAI in November 2022. The development in Artificial Intelligence research is so fast that some new developments appear almost every day, which are then integrated into our everyday. The technological singularity, or the hypothetical point in the future when Artificial Intelligence and other advanced technologies develop to the point where they surpass human capabilities, may occur in the not-so-distant future. This concept was first used by Raymond Kurzweil (2013). According to him, the explosive growth of computing capacity will lead to a point where the development of technology will be so fast that it will be difficult to predict exactly what will happen. Raymond Kurzweil (2013) predicts this date as 2045.

When we talk about Artificial Intelligence (Figure 1), we mean, for example, machine learning, deep learning, natural language processing, image and voice recognition, data mining and many other systems based on intelligent algorithms. The Hungarian Artificial Intelligence Coalition (Digital Prosperity Program, URL3) was founded in 2018, the main goal of which is to bring technology closer to all actors in society. It is very important to be aware of what the technology is capable of, what we can use it for, and what we should pay attention to when using it (Tilesch and Hatamleh 2021). As a result of the coalition's work, Hungary's Artificial Intelligence Strategy (URL4) was published in May 2020, in which one of the important segments of the priority AI research and development directions is the development of language technology. Among these developments, the interpretation of spoken and written texts is one of the fastest-developing areas. Nowadays, a greater percentage of communication takes place on digital channels in natural language, so it is very important to expect that machine systems have adequate language skills. Most of the developments in language technology focus on the English language, but it is a legitimate expectation of all societies to be able to communicate in their own mother tongue in the digital space. The application and development of technologies related to the Hungarian language is a priority of national interest. One of the key tools for this is the development of a Hungarian teaching corpus.⁴ Digitization of the Hungarian language is very important.

⁴ A corpus is a collection of actually occurring written or recorded spoken language data. The texts are selected and arranged according to some aspect. It does not necessarily contain entire texts, and it is not only a repository of texts, but also contains their bibliographic data and marks the structural units (paragraph, sentence). (URL5).

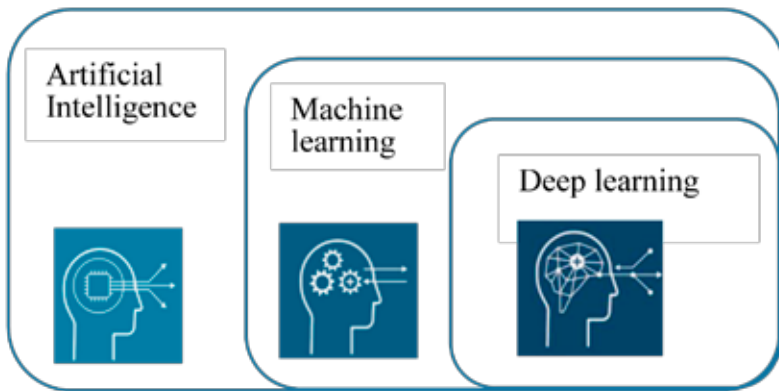


Figure 1. The relationship between Artificial Intelligence, machine learning and deep learning.

Language models

Natural language processing (NLP) is present in almost all areas of life, we can meet it when using Internet search engines or in the case of machine translators. Such NLP applications have been integrated into many tools, such as name recognition or sentiment analysis⁵. First, let's clarify what a language model is. An algorithm trained on a large amount of data is able to recognize linguistic connections and, based on these, generates an output that can be interpreted by the machine. The key element of the system is the learning process. What does it take to create a large language model? The learning process requires a very large computer capacity. To start learning, you need a lot of diverse and clean data. Of course, a person with expertise is a very important factor in this process. Different models were used when processing natural languages, we will now review them in chronological order.

⁵ Analysis of the emotional content of a given text. The texts are classified into 2 or 3 groups based on the emotional charge, which can be positive, or negative, and the 3rd group is neutral (Pang and Lee 2008). Sentiment analysis is the study of the emotional assessment of conversational, transactional and social media texts. The emotions appearing in the texts are determined by considering the intention of the text writer, the topic, the use of language, the context and the context of the text. By exploring emotions, the copywriter can better understand the audience's emotions, attitudes, and preferences, and shape their communication accordingly. Text generated by Puligpt.

Rule-based models

Language features are described by rules, these rules are the basis of deterministic algorithms⁶. The system for describing linguistic features is very complicated, as a result of which more and more complex systems were created, which became unmanageable over time. The essence of rule-based solutions is to transfer the code from one code to another according to some rule. In the case of rule-based language systems, working solutions have been created to solve the task of spam filtering, noun element recognition, and word class clarity. Promising research was carried out in the implementation of opinion mining, meaning clarification, syntactic analysis, machine translation, information extraction, and coreference⁷ resolution tasks. Answering the question, creating a paraphrase,⁸ extracting text, and creating a dialogue were only among the long-term goals in the case of rule-based models. It was an important expectation that each language should have its own basic language research kit (BLARK). It was created by the Hungarian Language and Speech Technology Center. Hungary has joined CLARIN, Europe's largest infrastructure of this kind. The paper *The Hungarian Language in the Digital Age* (Rehm and Uszkoreit 2012) was published. It has been stated that the processing of the English language is the best, but Hungarian ranks quite high behind English. Among the rule-based solutions are analyzers [Humor (Prószéky and Tihanyi 1996), HunMorph (Trón et al. 2005)], machine translators [webforditas.hu, MetaMorpho (Prószéky and Tihanyi 2002)], part-of-speech disambiguators [HunPOS (Halácsy et al. 2007), PurePOS (Orosz and Novák 2012)] were prepared in Hungarian. The first text corpora were completed (Orwell 1984,⁹ BUSZI¹⁰), although they were still small in size, they became the starting basis for statistical models.

Statistical models

When large amounts of data were available, rule-based models were replaced by statistical models. For these models, we do not describe rules, but the

⁶ Algorithms that always give the same output for a specific input.

⁷ Coreference is a linguistic term that refers to when two or more linguistic terms refer to the same entity in a text. Coreference resolution means identifying all occurrences of an entity in the text (Vincze et al. 2015; Vadász and Nyéki 2023).

⁸ We say the same sentence in different words.

⁹ Orwell's 1984 it was his work that was translated into almost every language and was a good starting point for machine linguistics research.

¹⁰ Budapest sociolinguistic interview, text corpus made from spoken texts.

models are created using machine learning through examples. The examples, i.e. the input, are labelled data, where the labels describe relevant features. For these models, it is difficult to find relevant features (feature engineering). Hand- or machine-annotated¹¹ text corpora (Sass 2016) appear as input, which are processed using statistical methods during machine learning, and the output is generated from this, which is, for example, the prediction of what other word may follow a given word. In the case of statistical language systems, working solutions have been created for solving tasks such as spam filtering, noun element recognition, and word type disambiguation. Acceptable solutions have already been prepared for the implementation of opinion mining, meaning clarification, syntactic analysis, machine translation, information extraction, and co-reference resolution tasks. In the tasks of answering questions, creating paraphrases, extracting text, and creating dialogues, which seemed very distant with rule-based systems, promising research has already been conducted here. In this period, text corpora in Hungarian were created, such as the monolingual Webcorpus (Halácsy et al. 2004), the annotated Szeged corpus (Csendes et al. 2003), the bilingual Hunglish (Halácsy et al. 2005), the Pázmány corpus (Endrédi and Prószy 2016), Hungarian National Dictionary (Váradi 2002), and small-scale historical corpus (Ómagyar corpus) and speech corpus have also been prepared (BEA). Very easy-to-use applications have appeared in the field of transport and telephone customer service. During this period, e-magyar.hu was created (Váradi et al. 2017), which contains the basic tools for machine analysis of the Hungarian language.

Language processing with neural networks

The first mathematical description of a neural network was made by Warren McCulloch and his colleagues in 1943. They realized that the nervous system works on an “all or nothing” basis, so it can be modelled mathematically. According to their idea, this is how the artificial neuron can be created. Figure 2 shows a human and an artificial neuron. In the case of the artificial neuron, the input ($x_1(n), \dots, x_m(n)$) is a known m -dimensional vector. n is n . denotes a moment in time. The $w_1(n), \dots, w_m(n)$ are the weights that the goal is to determine. $b(n)$ is the distortion, it is also unknown, it must be determined.

¹¹ Annotation means that we assign data to text units, such as paragraphs, sentences, and words, according to some aspect. For example, the type of word for words, the type of sentence for sentences, the emotional charge for a paragraph, and its author for a document. Annotation can be manual, machine, or a combination of the two.

The summation node produces a weighted sum of the inputs. The activation or transfer function is a function that we have to choose according to the task. The output is the value assigned by the neuron to the input.

The models are taught using a neural network system, where the neural network is a set of interconnected neurons (Fazekas 2013). Figure 3 shows the structure of the multilayer neural network. In deep learning, we use such a multilayer neural network. With neural network solutions, we do not need to specify rules, nor do we need to label the input data. The input is raw text, of which an ever-increasing amount is available with the rise of the Internet. In the process of learning, the machine system (Beszédes et al. 2021) invents the features itself.

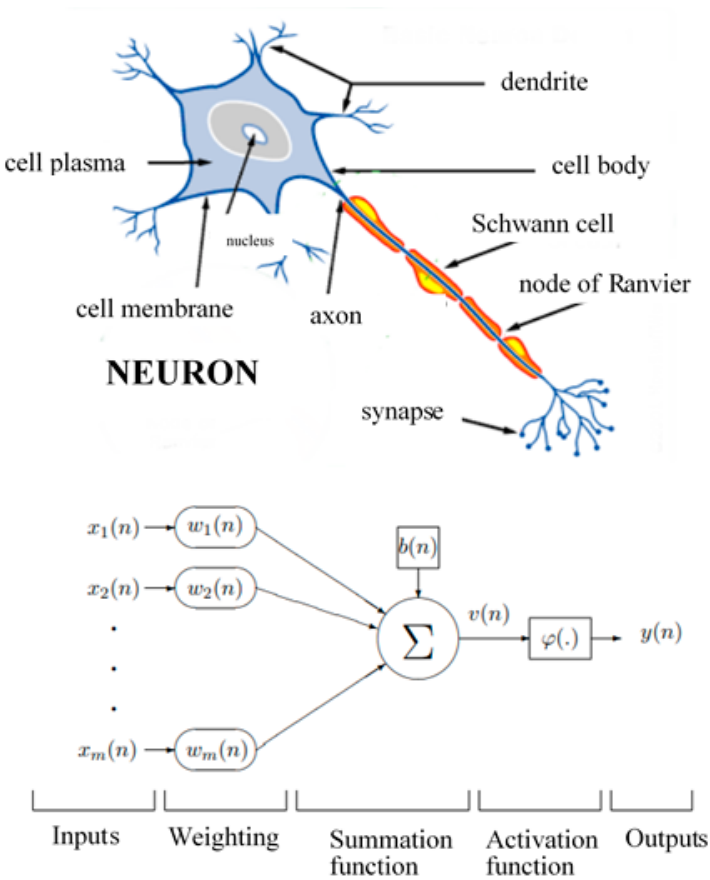


Figure 2. A natural (source: URL6.) and an artificial neuron (source: URL7. page 16).

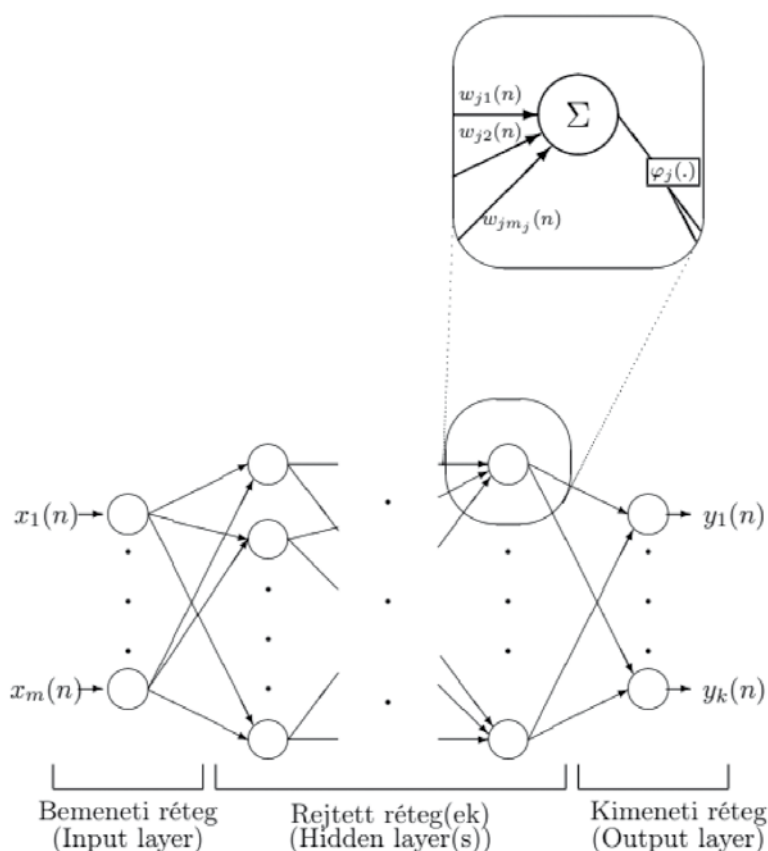


Figure 3. Structure of the multilayer neural network (source: URL7. page 32).

In the case of language models, it is very important what data they were taught with. This teaching dataset is called a corpus. The Hungarian National Dictionary aims to be a general-purpose representative corpus of today's Hungarian written vernacular. Huge high-quality text corpora are available (Sass 2016), but models can be further fine-tuned with smaller specialized text corpora. Figure 4 shows the corpora available on the National Corpus Portal in the chronological order of their creation.

The neural network models already had acceptable solutions for the tasks of answering questions, creating paraphrases, extracting text, and creating dialogues. The basis of neural network systems was the appearance of word vector representations. Such word vector representations are

2023	ParlaMint 4.0
2023	MoMa – Moldvai magyar korpusz
2023	KED – Középmagyar emlékirat- és drámakorpusz
2016	„új” Magyar történeti szövegtár
2015	Mikes-szótár
2014	MNSZ2
2014	Történeti Magánéleti Korpusz
2013	Ómagyar korpusz
2012	BUSZI
2010	Hunglish Korpusz
2009	Mazsola
2006	Webkorpusz
2005	MNSZ1
1999	Orwell: 1984

Figure 4. Text corpora are available on the National Corpus Portal (source: URL8).

word2vec (Mikolov 2013), fasttext. The dynamic language models (mBERT, HuBERT) were published. The meaning of words is closely related to their environment (distributional semantics). Linguistic elements are points in the vector space, and semantically or morphologically similar words are also close to each other in the vector space. During the operation of the system, the deep learning neural network assigns a multidimensional vector to each word (elementary unit, token) of the text entered as input, the i -th element of which shows that the text is examined i . the probability that certain elements of the text occur in the environment of its element. Anyone can test how a neural network learns (URL9). With the support of Microsoft Hungary, the Institute of Linguistics and the University of Pécs prepared the Hungarian adaptation of the BERT-large model, the HILBERT (Feldmann et al. 2021) model. More and more speech corpora are being prepared, in addition to BEA, for example, BEKK and HuTongue speech corpus appear.

In 2017, the emergence of the transformer architecture (Vaswani et al. 2017) brought great progress in machine translation. The transformer architecture can be analytical (encoder) or generative (decoder). Many tasks can be solved with the generative language model: creating a summary, translating, writing program code, etc. Analytical language models must be fine-tuned separately for each task, with labeled data. Encoder models analyze the text coming to

their input, such as the BERT, RoBERTa, ELECTRA, ALBERT language models. The decoder models are suitable for text generation, the first one was the GPT model, but this is the PaLM MT-NLG. Encoder-decoder type model is T5, BART, Marian. The language model is the set of weights produced at the end of learning. In Table 1, we compare some large language models based on the number of parameters and the size of the text corpus used for learning.

Table 1. Characteristics of some large language models

Model Name	Company	Number of Parameters	Size of Training Corpus
GPT-2	OpenAI	1.5 billion	40 GB
GPT-3,5	OpenAI	175 billion	45 TB
BERT	Google	340 million	3.3 billion text pieces
XLNet	Google	340 million	126 GB
RoBERTa	Facebook	355 million	160 GB
Megatron	Nvidia	8.3 billion	512 GB
ALBERT	Google	11 million	3.3 billion text pieces
T5	Google	11 billion	750 GB
T-NLG	Microsoft	17 billion	568 GB
ELECTRA	Google	110 million	1.7 TB
LaMDA	Google	137 billion	1.56 billion words
GPT-4	OpenAI	~ 600 billion	~ 120 TB
mT5	Google	580 million	29 TB
ProphetNet	Microsoft	550 million	160 GB
DeBERTa	Microsoft	1.5 billion	160 GB
PEGASUS	Google	568 million	750 GB

(source: own editing)

Nowadays, more and more language models are appearing. There is huge competition between the big companies. The development of language models involves enormous technological and financial challenges, which is why only large companies can afford the development of huge language models. The goal is to create large language models that are capable of solving all kinds of linguistic tasks (extraction, sentiment analysis, translation, tagging, name element recognition, anonymization¹² etc.) even without fine-tuning. Figure 5 shows some of the more important language models.

¹² The process of converting personal data into anonymous data, as a result of which the data subject can no longer be identified – National Legal Code 2021 XCI. law (URL10).



Figure 5. The major language models (URL11).

More important Hungarian language models

In the shadow of the greats, they also developed Hungarian language models with a more modest infrastructure. huBERT was completed in 2020 (Nemeskey 2021), which was based on the BERT Base model. 110 million is the number of parameters of the model, the teaching corpus was the webcorpus 2.0. HILBERT was completed in 2021 with the support of HILANCO and Microsoft (Feldmann et al. 2021), which was based on the BERT Large model with 340 million parameters. The teaching corpus consists of 3.67 billion words. The HILANCO language model, which is bilingual (English, Hungarian), was introduced in 2022. The number of parameters is 6.7 billion and the teacher's corpus is 127 billion words, of which 102 billion are English words and 25 billion are Hungarian words. In the same year, the language model PULI GPT 3SX (Yang et al. 2023a) with 6.7 billion parameters was published. The teacher's corpus was 32.5 billion Hungarian words. The trilingual PULI GPTrío was completed in 2023 (Yang et al. 2023b), which also has 6.7 billion parameters. The three languages are Hungarian, English and Chinese, and the teaching corpus was 200 billion words. The parameter number means that the system can assign so many characteristics to a grammatical element. If we think about how many different characteristics we can specify for an object, then the

order of billions shows how much technology is capable of. Table 2 shows a comparison of some Hungarian language models based on the number of parameters and the size of the teaching corpus.

Table 2. Characteristics of some Hungarian language models

Model Name	Number of Parameters	Size of Training Corpus
huBERT	110 million	10 billion words (webcorpus 2.0)
HILBERT	340 million	3.67 billion words
HILANCO	6.7 billion	127 billion words (English, Hungarian)
PULI GPT 3SX	6.7 billion	32.5 billion words (Hungarian)
PULI GPTrío	6.7 billion	200 billion words (Hungarian, English and Chinese)

(source: own editing)

The complicated grammatical structure of the Hungarian language poses a difficulty when creating Hungarian language models. Sentences do not have a fixed word order, a sentence can be written in several different word orders. There are many exceptions to the rules, and handling them is not easy either. There are many turns of phrase and expressions in the Hungarian language, the meaning of which is not always clear (Arató–Balázs 2022, 2023). Language models must understand and take into account the cultural and social context of the Hungarian language, they must be able to correctly interpret the texts and give appropriate answers to them. The Hungarian language has different dialects and accents, which is an additional difficulty for models. In the case of ChatGPT, which is not specifically taught in Hungarian but understands Hungarian, the size of the teaching corpus is a limitation, which is limited to 128 million Hungarian text fragments. In contrast, in the case of PULI, teaching was carried out with 32.5 billion text fragments. It is very important to create Hungarian teaching corpora of the right size (Balázs 1997; Prószéky and Váradi 2023).

Language models in education

By fine-tuning language models with specialized texts, applications based on these models will be able to generate and analyze high-quality specialized texts. In the case of subjects belonging to different scientific fields, these fine-tuned models can be of great help in preparing the curriculum. Language models could be prepared for each discipline, which would be fine-tuned for

professional texts in addition to general knowledge. In the case of universities, the theses stored electronically could be the teaching texts for fine-tuning. What can we use applications based on language models?

Sentiment analysis (Laki and Yang 2022a; Yang and Laki 2023): You can decide whether a given text has a positive, negative or neutral charge. In education, for example, sentiment analysis can be used to filter out positive and negative opinions in the evaluation of students by instructors. In the case of correspondence with students, such a mood analysis could also be prepared, and depending on the result, additional benefits could be brought. The staff of the Linguistics Research Center created a sentiment analysis application that anyone can try (URL12).

Chatbots: They are virtual assistants for students who can answer their questions immediately. It is most beneficial for students who do not dare to ask their questions in class. Chatbots can also provide explanations of definitions, but they can also give students study tips. Chatbots can also help students in their academic affairs and answer administrative questions. The best-known chatbot is ChatGPT (URL2), but there is also a Hungarian counterpart called PULI GPT (URL13).

Translation: translating texts from one language to another. In the case of teaching materials in a foreign language, it can help students translate the text. They can also play an important role in maintaining contact with foreign students, and can also be useful in case of participation in international programs. Like students, they can also provide teachers with great help in getting to know and understanding research results in a foreign language. A demo of a translation program is available on the website of the Linguistics Research Center (Laki and Yang 2022b) (URL14).

Summarizing: very useful for extracting essential information from long texts. It can help students prepare for the exam by reviewing the essential parts of the curriculum. For summarizing the content of a topic when preparing teaching material for teachers. It is also a great help for students when writing their theses, and for instructors when preparing publications, since these papers always include a summary of the work done. These applications produce digests of given long texts (Yang 2022) (URL15).

Question-answer: generates appropriate answers to the questions we ask. Both teachers and students can use these applications, although it is important to be critical of the answers to the question. An example of this is the command PULI (Yang et al. 2023b, Yang et al. 2024a, Yang et al. 2024b) (URL16).

Content creation: these applications are able to produce content on a wide variety of topics based on specified criteria and parameters.

Applications based on GPT models are all suitable for this. Since these models do not copy texts from some source, but are able to produce their own texts as a result of learning, it is therefore very important to treat the produced content critically.

Speech recognition: this is an application that transforms spoken language into written text (Németh and Olaszy 2010). The final result is greatly influenced by ambient noise. It can be very useful for instructors to be able to immediately create a written text from their lectures, and the digital content for the subject is already ready. The hearing-impaired are assisted by the textual transcripts of the presentations given in the preface.

Tagging: keywords are important information for searchers in the case of studies. During tagging, we add tags and keywords to a specified text using the applications created for this purpose. A good choice of tags assigned to long texts can also help students when searching for content (Yang et al. 2020) (URL17).

Language learning support: A chatbot working with language models as part of an application helps you learn a language interactively, for example by doing conversational exercises or giving grammar explanations (Porkoláb and Fekete 2023). Some popular AI-based language learning apps are TalkPal, Duolingo, Rosetta Stone, Babbel, and Mondly. These apps use machine learning algorithms and other artificial intelligence technologies to provide a personalized and interactive learning experience.

Summary

The rapid development of artificial intelligence (AI) today represents an extremely exciting and diverse field. The application of AI is already present in almost every sector, including healthcare, finance, automotive and education. Newer models, such as large language models and deep learning algorithms, are constantly improving and enabling new forms of human-machine interactions. Large language models are revolutionizing education with their wide applicability and interactive features. One of their main advantages is the ability to learn from a wide and diverse range of sources, allowing learners to access different knowledge bases and expertise through a single platform. In addition, large language models can also be used interactively in education. These models allow students to ask and answer questions, participate in simulations or engage in dialogue with virtual characters. This improves students' communication and problem-solving skills. However, it is important to recognize that these models are not perfect. Sometimes they may give wrong or inaccurate information, which

can be misleading for students. In addition, ethical issues also arise, such as personal data protection and the protection of students' intellectual property. It is important that users critically examine and contextualize the information provided by such models. Overall, large language models have great potential in education if they are properly used and critically evaluated. It is important that teachers are prepared to use models and consciously integrate them into the educational process in order to provide maximum benefits to students while taking into account possible challenges and limitations.

References

- Arató, Balázs – Balázs, Géza 2022. The linguistic norm and norm of legal language. *Magyar Nyelvőr*, 46 (special issue), 91–103.
<https://doi.org/10.38143/Nyr.2022.5.91>
- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban? A digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 147. 755–76.
<https://doi.org/10.38143/Nyr.2023.5.755>
- Arisztotelész 1969. *Politika*. In: Szabó Miklós (fordító) – Horváth Henrik (lektor). Gondolat Könyvkiadó. Budapest.
- Balázs, Géza 1997. *The Story of Hungarian*. Corvina, Budapest.
- Beszédes, Bertalan – Széll, Károly – Györök, György 2021. Redundant photo-voltaic power cell in a highly reliable system. *Electronics* (Switzerland). 10(11). 1253.
<https://doi.org/10.3390/electronics10111253>
- Buzás, György Miklós 2021. A mesterséges intelligencia története. *Gastroenterológiai és hepatológiai szemle – Central European journal of gastroenterology and hepatology*. 7(3). 121–7. ISSN 2415-9107.
<https://doi.org/10.33570/ceuJGH.7.2.121>
- Csendes, Dóra – Hatvani, Csaba – Alexin, Zoltán – Csirik, János – Gyimóthy, Tibor – Prószéky, Gábor – Váradi, Tamás 2003. Kézzel annotált magyar nyelvi korpusz: a Szeged Korpusz. In: *II. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged. 238–45.
- Daugherty, Paul – Wilson, James 2018. *Human+Machine Reimagining Work In the Age of AI*. Boston. Harvard Business Review Press. ISBN 9781633693876
- Endrédy, István – Prószéky, Gábor. A Pázmány korpusz. *Nyelvtudományi Közlemények*, 112. 191–205.
- Fazekas, István 2013. *Neurális hálózatok*. Debreceni Egyetemi Kiadó. Debrecen.
- Feldmann, Ádám – Hajdu, Róbert – Indig, Balázs – Sass, Bálint – Makrai, Márton – Mittelholcz, Iván – Halász, Dávid – Yang, Zijian Győző – Váradi, Tamás 2021. HILBERT, magyar nyelvű BERT-large modell tanítása felhő környezetben. In: *XVII. Magyar számítógépes nyelvészeti konferencia*. Szeged. 29–36.
- Ford, Martin 2022. *Jövőnk a robotok korában – hogyan alakítja a mesterséges intelligencia az életünket?* HVG Könyvek. 272. ISBN: 9789635653379

- Goodfellow, Ian – Bengio, Yoshua – Courville, Aaron 2016. *Deep learning*. MIT press.
- Halácsy, Péter – Kornai, András – Németh, László – Rung, András – Szakadát, István – Trón, Viktor 2004. Creating Open Language Resources for Hungarian. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*. Lisbon, Portugal. European Language Resources Association (ELRA). 1201–04.
- Halácsy, Péter – Kornai, András – Németh, László – Sass, Bálint – Varga, Dániel – Váradi, Tamás – Vonyó, Attila 2005. A Hunglish korpusz és szótár. In: Alexin, Zoltán – Csendes, Dóra (ed.): *III. Magyar Számítógépes Nyelvészeti Konferencia*. Juhász Nyomda. Szeged. 134–43.
- Halácsy, Péter – Kornai, András – Oravecz, Csaba 2007. HunPos – an open source trigram tagger. In: *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*. Prague, Czech Republic. Association for Computational Linguistics. 209–12. <https://doi.org/10.3115/1557769.1557830>
- Hsu, Feng-Hsiung 1999. IBM's Deep Blue Chess grandmaster chips. In: *IEEE Micro*. 19(2). 70–81. <https://doi.org/10.1109/40.755469>
- Khurana, Diksha – Koli, Aditya – Khatter, Kiran – Singh, Sukhdev 2023. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl* 82. 3713–44. <https://doi.org/10.1007/s11042-022-13428-4>
- Kurzweil, Ray 2013. *A szingularitás küszöbén. Amikor az emberiség meghaladja a biológiát*. Ad Astra Kiadó. 942. ISBN: 9786155229251
- Laki, László János – Yang, Zijian Győző 2022a. Mondatszintű szentiment analízis teljesítményének javítása adatkiterjesztéssel eljárásokkal. In: *XVIII. Magyar Számítógépes Nyelvészeti Konferencia*. 295–310.
- Laki, László János – Yang, Zijian Győző 2022b. Neural machine translation for Hungarian. *Acta Linguistica Academica*. 69(4). 501–20. <https://doi.org/10.1556/2062.2022.00576>
- McCarthy, John 2007. What Is Artificial Intelligence? Computer Science Department. Stanford University, Stanford. CA 94305.
- McCulloch, Warren S – Pitts Walter 1943. A logical calculus of the ideas immanent in nervous activity. *Bull Mathematical Biol.* 5. 115–33. reprinted: idem 1990. 52: 99–115. <https://doi.org/10.1007/BF02478259>
- Mikolov, Thomas – Chen, Kai – Corrado, Greg – Dean, Jeffrey 2013. Efficient Estimation of Word Representations in Vector Space. In: *1st International Conference on Learning Representations*. ICLR 2013. Scottsdale, Arizona, USA. Workshop Track Proceedings.
- Nemeskey, Dávid Márk 2021. Introducing huBERT. In: *XVII. Magyar Számítógépes Nyelvészeti Konferencia*. 3–14.
- Németh, Géza – Olaszy, Gábor 2010. *A magyar beszéd*. Akadémiai Kiadó. Budapest.
- Newborn, Monty 2012. *Kasparov versus Deep Blue: Computer chess comes of age*. Springer Science & Business Media.
- Orosz, György – Novák, Attila 2012. PurePos – an Open Source Morphological Disambiguator. In: *Proceedings of the 9th International Workshop on Natural*

- Language Processing and Cognitive Science*. 53–63.
<https://doi.org/10.5220/0004090300530063>.
- Pang, Bo – Lee, Lilian 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*. 2(1–2). 1–135.
<https://doi.org/10.1561/15000000011>
- Porkoláb, Ádám – Fekete, Tamás 2023. A mesterséges intelligencia alkalmazása a nyelvtanulásban. *Iskolakultúra*, 33(8). 67–80.
<https://doi.org/10.14232/iskkult.2023.8.67>
- Prószéky, Gábor – Tihanyi, László 1996. Humor – A Morphological System for Corpus Analysis. In: *Proceedings of the first TELRI seminar* in Tihany, Budapest. 149–58.
- Prószéky, Gábor – Tihanyi, László 2002. Metamorpho: A pattern-based machine translation system. In: *Proceedings of the 24th 'Translating and the Computer' Conference*. London, United Kingdom. 19–24.
- Prószéky, Gábor – Váradi, Tamás 2023. A magyar nyelv digitális fenntarthatóságának támogatása. *Magyar Nyelv*, 119(4). 482–89.
<https://doi.org/10.18349/MagyarNyelv.2023.4.482>
- Rehm, Georg – Uszkoreit, Hans (ed) 2012. *A magyar nyelv a digitális korban*. Springer Verlag. ISBN: 9783642303784
- Rosenblatt, Frank 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6). 386–408.
<https://doi.org/10.1037/h0042519>
- Searle, John. R. 1980. Minds, Brains and Programs. *Behavioral and Brain Science*, 3(3). 417–24. Magyarul: Searle, J. R. (1996). *Az elme, az agy és a programok világa*. Kognitív tudomány. Osiris Kiadó – Láthatatlan Kollégium.
<https://doi.org/10.1017/S0140525X00005756>
- Sass, Bálint 2016. Nyelvészeti szövegkeresők. Nemzeti Korpuszportál. In: *Magyar Tudomány*, 177(7). 798–808.
- Tilesch, György – Hatamleh, Omar 2021. *Mesterség és intelligencia*. Első kiadás. Libri Kiadó. Budapest. 248 oldal. ISBN 978-963-433-829-1
- Trón, Viktor – Gyepesi, Gyögy – Halácsy, Péter – Kornai, András – Németh, László – Varga, Dániel 2005. Hunmorph: Open Source Word Analysis. In: *Proceedings of Workshop on Software*. Ann Arbor, Michigan. Association for Computational Linguistics. 77–85.
- Turing, Alan Matheson 1936. Turing gép. *Proc London Math Soc*, 242. 230–65.
- Turing, Alan Matheson 1950. Computing Machinery and Intelligence. *Mind*, 59. 433–60. <https://doi.org/10.1093/mind/LIX.236.433>
- Vadász, Noémi – Nyéki, Bence 2023. Koreferenciafeloldás magyar szövegeken BERT-tel. In: *XIX. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged. 119–31.
- Váradi, Tamás 2002. The Hungarian National Corpus. In: *Proceedings of the 3rd LREC Conference*. Las Palmas, Spanyolország. 385–9.
- Váradi, Tamás – Simon, Eszter – Sass, Bálint – Geröcs, Mátyás – Mittelholcz, Iván – Novák, Attila – Indig, Balázs – Prószéky, Gábor – Farkas, Richárd –

- Vincze, Veronika 2017. Az e-magyar digitális nyelvfeldolgozó rendszer. In: *XIII. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2017)*. Szegedi Tudományegyetem Informatikai Tanszékcsoport, Szeged. 49-60. ISBN 978-963-306-518-149-60
- Vaswani, Ashish – Shazeer, Noam – Parmar, Niki – Uszkoreit, Jakob – Jones, Llion – Gomez, Aidan N – Kaiser, Łukasz – Polosukhin, Illia 2017. Attention is All you Need. *Advances in neural information processing systems*, 30. 5998–6008
- Vincze, Veronika – Hegedűs, Klára – Farkas, Richárd 2015. SzegedKoref: kézzel annotált magyar nyelvű koreferenciakorpusz. In: *XI. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged. 312–9.
- Weizenbaum, Joseph 1966. ELIZA – A Computer Program for the Study of Natural Language Communication between Man and Machine. *Communications of the Association for Computing Machinery* 9. 36–45.
- Yang, Zijian Győző – Novák, Attila – Laki, László János 2020. Automatikus tematikuscímke-ajánló rendszer sajtószövegekhez. In: *XVI. Magyar Számítógépes Nyelvészeti Konferencia*. 155–68.
- Yang Zijian Győző 2022. Neural text summarization for Hungarian. *Acta Linguistica Academica*, 69. 474–500. <https://doi.org/10.1556/2062.2022.00577>
- Yang, Zijian Győző – Laki, László János 2023. Neurális entitásorientált szentimentelemző alkalmazás magyar nyelvre. In: *XIX. Magyar Számítógépes Nyelvészeti Konferencia*. 107–17.
- Yang, Zijian Győző – Dodé, Réka – Ferenczi, Gergő – Héja, Enikő – Jelencsik-Mátyus, Kinga – Körös, Ádám – Laki, László János – Ligeti-Nagy, Noémi – Vadász, Noémi – Váradi, Tamás. 2023a. Jönnek a nagyok! BERT-Large, GPT-2 és GPT-3 nyelvmodellek magyar nyelvre. In: *XIX. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged. 247–62.
- Yang, Zijian Győző – Laki, László János – Váradi, Tamás – Prószyék, Gábor 2023b. Mono- and multilingual GPT-3 models for Hungarian. In: *26th International Conference of Text. Speech and Dialogue (TSD2023)*. Plzeň, Czech Republic. 94–104.
- Yang, Zijian Győző – Dodé, Réka – Héja, Enikő – Laki, László János – Ligeti-Nagy, Noémi – Madarász, Gábor – Váradi, Tamás 2024a. ParancsPULI: Az utasításkövető PULI-moddell. In: Berend, Gábor – Gosztolya, Gábor – Vincze, Veronika (szerk.) *XX. Magyar Számítógépes Nyelvészeti Konferencia*. Szeged. 61–72.
- Yang, Zijian Győző – Szilárd, Szilárd – Ligeti-Nagy, Noémi 2024b. Magyar nyelvű utasításkövető korpusz építése Stanford Alpaca promptok fordításával és lokalizálásával. In: *XX. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY2024)*. Szeged. 243–55.
- URL1. Google DeepMind AlphaGo. <https://deepmind.google/technologies/alphago/> (download 2024.02.17.)
- URL2. OpenAI ChatGPT. <https://openai.com/chatgpt> (download 2024.02.17.)
- URL3. Digital Prosperity Program. <https://digitalisjoletprogram.hu/hu/tartalom/mesterseges-intelligencia-koalicio> (download 2024.02.02.)

- URL4. Artificial intelligence strategy of Hungary. <https://digitalisjoletprogram.hu/files/2f/32/2f32f239878a4559b6541e46277d6e88.pdf> (download 2024.02.02.)
- URL5. Hungarian National Dictionary. <https://clara.nytud.hu/mnsz2-dev/> (download 2024.02.20.)
- URL6. I love Biosz blogspot. <http://ilovebiosz.blogspot.com/2011/03/idegrendszer.html> (download 2024.03.05.)
- URL7. Fazekas István (2013) Neurális hálózatok. Debrecen. 204. https://gyires.inf.unideb.hu/GyBITT/19/Neuralis_halozatok_v8.pdf (download 2024.03.05.)
- URL8. National Corpus Portal. <http://corpus.nytud.hu/nkp/> (download 2024.02.20.)
- URL9. Simulation of neural network operation. <http://playground.tensorflow.org> (download 2024.02.01.)
- URL10. National Legal Code 2021 XCI. law. <https://njt.hu/jogszabaly/2021-91-00-00> (download 2024.02.05.)
- URL11. Large language models available for free. <https://github.com/eugeneyan/open-llms> (download 2024.02.01.)
- URL12. Sentiment analysis demo program of Linguistics Research Center. <https://juniper.nytud.hu/demo/sentana> (download 2024.03.01.)
- URL13. The PULI GPT-3SX demo program of the Linguistics Research Center. <https://juniper.nytud.hu/demo/puli> (download 2024.03.01.)
- URL14. Machine translation demo program of the Linguistics Research Center. <https://juniper.nytud.hu/demo/translation> (download 2024.03.01.)
- URL15. Abstract extraction/summary generation demo program of the Linguistics Research Center. <https://juniper.nytud.hu/demo/summarization> (download 2024.03.01.)
- URL16. ParancsPULI demo program of Linguistics Research Center. <https://juniper.nytud.hu/demo/parancspuli> (download 2024.03.01.)
- URL17. Tagging presentation program of the Linguistics Research Center. <https://juniper.nytud.hu/demo/tagger> (download 2024.03.01.)

Nikoletta Tolner

Assistant Professor

Óbuda University Alba Regia Technical Faculty

E-mail: tolner.nikoletta@uni-obuda.hu

<https://orcid.org/0009-0003-1288-3111>

Monika Pogátsnik

Assistant Professor

Óbuda University Alba Regia Technical Faculty

E-mail: pogatsnik.monika@uni-obuda.hu

<https://orcid.org/0000-0002-2698-7291>

ARTIFICIAL INTELLIGENCE AS A VIRTUAL LANGUAGE TEACHER

Abstract

This study explores the role of artificial intelligence (AI) as a virtual language teacher, highlighting its capabilities and applications in education. AI is defined as the ability of machines to acquire and apply knowledge to perform cognitive tasks, such as processing language, learning, and decision-making. The paper examines various AI technologies, including chatbots and natural language processing, that enhance language learning through personalised and interactive methods. Emphasising the effectiveness of AI in providing real-time feedback, adapting to individual learning paces, and offering customised learning experiences, the study concludes that AI-based tools significantly improve language learning efficiency and accessibility. Additionally, the integration of AI in language education can reduce language barriers, offering learners access to diverse linguistic resources and real-time pronunciation correction, ultimately fostering a more inclusive and effective learning environment.

Keywords: artificial intelligence (AI), language learning, Natural Language Processing (NLP), chatbots, personalised learning, interactive education

I. What is artificial intelligence? (Introduction)

Without claiming completeness, some of the definitions of artificial intelligence that I consider to be the most pragmatic are:

OECD1 (2016) and UNCTAD2 (2017) define artificial intelligence (World Investment Report 2016 – Investor nationality: policy challenges UNCTAD (2017). World Investment Report)¹ as the ability of machines and systems to acquire and apply knowledge to carry out intelligent behavior. This means performing a broad variety of cognitive tasks, e.g. sensing, processing oral language, reasoning, learning, making decisions, and demonstrating an ability to move and manipulate objects accordingly. Intelligent systems use

¹ World Investment Report 2016: Investor Nationality: Policy Challenges, UNCTAD 2017 (2017, June 2). *Transnational Corporations*, 23(3), 67–101. <https://doi.org/10.18356/44d1623d-en> see in Going Digital: Making the Transformation Work for Growth and Well-being, Meeting of the OECD Council at Ministerial Level Paris, 7-8 June 2017. <https://www.oecd.org/mcm/documents/C-MIN-2017-4%20EN.pdf>

a combination of big data analytics, cloud computing, machine-to-machine communication, and the Internet of Things (IoT)² to operate and learn.

In the Turing test definition, artificial intelligence is the ability of a machine to communicate (via electronic output devices) with a human without the fact that the interaction is not with a real human being detected, and the substantive decision-making aspect is binary based. Marvin Minsky, a pioneering figure in artificial intelligence, defined the technology as enabling machines to perform activities that require human intelligence. At the core of artificial intelligence are the research theories, methodologies, and applications by which human intelligence can be simulated, extended, and extended (Jiang 2022: 13).³

AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions.⁴

In summary, modern artificial intelligence (AI) is a system that can sense its environment, take action to maximize the chances of achieving its goals, and interpret and analyze data in a way that it can learn from and adapt to, i.e., improve itself.

One area of investigation of the impact of AI-based applications on human life is the role of AI in education, e.g., using OpenAI ChatGPT-4, human-AI collaboration, and personalized methods of AI teaching. In the following, I would like to share my personal experience with Reader T on how AI can be used in language teaching and learning. I tested software that promised unique results in English learning. The author of this paper has no educational background but has 24 years of higher education experience as a lawyer, so the primary aim of the paper is not a linguistic analysis but to explore the teaching methods provided by AI.

² The Internet of Things is a concept, a general idea, that describes how distinctive electronic devices in our environment can sense signals in the environment, collect data, and then share these sensations with other devices, connecting them to a network.

³ Jiang, R. (2022, November 16). How does artificial intelligence empower EFL teaching and learning nowadays? A review of artificial intelligence in the EFL context. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.1049401>

⁴ Proposal for a regulation of the European Parliament and of the Council laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain EU legislative acts Brussels, 21.4.2021 COM (2021) 206 final, 2021/0106 (COD), Article 3, definitions, point 1.

II. Types, tools, and applications of AI

There are purely software-based AI systems that operate exclusively in the virtual world (e.g. voice assistants, image analysis software, search engines, speech, and face recognition systems) or AI can be embedded in hardware (robots, self-driving vehicles, drones), internet applications, algorithms, and chatbots.

In the case of artificial intelligence, we talk about weak or strong systems. Weak or Narrow AI is AI created to solve a problem and focus on a single task, such as AlphaGo, Alexa, or Siri. AlphaGo is great in the one game it was designed for, but not good in others. Strong AI is not known to exist now. This would describe a self-aware AI with its own emotions. Ray Kurzweil predicts a date of 2045 for the emergence of a robot as intelligent as a human. This point is called the singularity. (Keserű 2000: 43).⁵ At present, AI development is mostly focused on a few specific areas and techniques, which may include large-scale Machine Learning, Deep Learning, Natural Language Processing (NLP);⁶ computer vision (image analysis), algorithmic game theory, economic and social computing dimensions of AI, and systems for automated robotic processes.⁷

Currently, AI is being applied with extraordinary efficiency in the following areas:

1. Data analysis and data mining: AI enables companies and researchers to extract valuable information (patterns) from large data sets. It is widely used in marketing, customer relations, and market research.
2. Automation: the use of AI enables the automation of monotonous, repetitive tasks, increasing efficiency and reducing the chance of human error. Both industrial robots and administrative utilities use this technology.
3. Speech recognition and language processing. This technology is essential for the operation of virtual assistants and chatbots.

⁵ Keserű (2000). A 21. századi technológiai változások hatása a jogalkotásra. Dialóg Campus.

⁶ The meeting of artificial intelligence and linguistics is the understanding of language by computer methods, which can then be used in a variety of ways, such as as a basis for content analysis, text tagging, sentiment analysis, or automatic machine translation.

⁷ Some specific areas in which AI already shows a unique technological development are, e.g., <https://leonardo.ai/>, <https://magicai.hu/>, <https://chat.openai.com/>, <https://www.suno.ai/>, <https://www.heygen.com/>, and <https://www.brandcrowd.com/ai-logo-generator>.

4. Image processing and vision technology: using AI, machines can interpret images and videos, revolutionizing facial recognition, medical diagnostics, and security systems.
5. Personalized recommendations: The AI, with the help of online stores and streaming services, can better understand the tastes and preferences of users, so they can provide more accurate recommendations.
6. Autonomous vehicles: AI is key to the development of self-driving cars, drones, and other autonomous vehicles that can navigate and make decisions on their own.
7. Robotics: AI helps robots become more intelligent and adaptive, enabling them to perform more complex tasks in industry and everyday life.
8. Health diagnostics and treatment: AI revolutionizes medical diagnostics and treatment, helping to detect diseases early and develop more effective therapies.
9. Financial analysis and risk management: AI's application in the financial sector helps in analyzing market trends, reducing risks, and developing automated trading strategies.
10. Smart homes and cities: AI helps increase energy efficiency and better manage city infrastructure through smart home devices and city services.
11. Games and simulation: AI is playing an increasingly important role in the development of video games and virtual reality applications, providing more realistic and interactive experiences.
12. Security and surveillance: AI can help security systems be more efficient and intelligent through facial recognition and automatic detection of suspicious activities.
13. Scientific research: AI contributes to scientific research by helping to analyze data, make discoveries, and speed up the research process.
14. Education: With AI, teaching materials and learning methods can be customized, increasing the efficiency and availability of education.
15. Art and Creativity: AI opens new dimensions in art and creativity, enabling machine-generated music, graphics, and other artistic creations.

III. A brief history of AI

The history of AI began in the 1940s when the “thinking” of machines first started to be addressed. In the 1950s, Alan Turing, a British mathematician and

codebreaker (he developed a computer to decipher codes used by the Germans during the Second World War), formulated the so-called Turing test, a method to determine whether a machine could imitate human intelligence. The next big step was the Dartmouth Conference (1956), which is considered the birth of AI as a formal scientific field, as it was here that the term “artificial intelligence” was first used. The 1980s marked another milestone for AI research, a period of AI revival when expert systems gained considerable popularity in corporate decision-making processes. The development of microprocessors (Furber, n.d.)⁸ accelerated, increasing the speed of computation, which in turn enabled the modeling of neural networks. In the 1990s, the US government made the Internet available to the private sector, which contributed significantly to boosting innovation and globalization. In 1997, IBM’s Deep Blue computer defeated Garry Kasparov, the world chess champion, marking a major milestone in the development of AI. In the early 2000s, AI research focused mainly on machine learning, speech recognition, and image processing. In 2006, Geoffrey Hinton, one of the pioneers of deep learning, introduced the use of deep neural networks, which revolutionized machine learning and opened new possibilities for AI. In the 2010s, deep learning and working with big data gained further momentum. In 2011, IBM’s Watson AI system won the quiz game Jeopardy! marking another milestone in the development of AI, particularly in the areas of natural language processing and machine learning. In 2012, the Google Brain project, which focused on deep learning, showed significant progress in image recognition, demonstrating that neural networks can process and interpret visual information in a similar way to the human brain. AI has become indispensable in a growing number of industries, significantly impacting the economy, education, healthcare, and everyday life (Keserű 2000:40-2).⁹

⁸ “The microprocessor – the central processing unit of a computer integrated on a single microchip – dominates computing at every scale, from the smallest consumer devices to the largest supercomputers. A computer needs memory to store programs and data, a processor to execute those programs using the data, and I/O (input/output) capabilities to interface with the outside world. The intensive action occurs inside the processor, and the microprocessor integrates all processing functions on a single microchip. The introduction of the microprocessor represented a breakthrough in the size and cost of computer systems and was one of the advances that led to the personal computer (PC) revolution and then to the revolution in mobile devices. The next revolution in computing, in which the microprocessor will play a central role, is the IoT – the Internet of Things.” In Furber. (n.d.). Microprocessors are engines of the digital age. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5378251>.

⁹ Keserű (2000). *A 21. századi technológiai változások hatása a jogalkotásra*. Dialóg Campus.

There are many ways of developing AI and all of them are being used simultaneously. One is good programming and data, and the other is collecting behaviors, so the machine must be used. The way I'm developing a self-driving car is to program it and run it on a test track, and the third way is to send it out on the road; it will do 200,000 kilometers, and that's how it will be smart. The development of AI will reach a stage where it can be made available to the public in an open test drive in November 2022.

ChatGPT, developed using OpenAI in language learning, is the best-known and most widely used AI technology today. Behind it is a large neural network that behaves like the human brain. In the early 2000s, the field of chatbots and natural language processing (NLP) was still in its infancy. The first chatbots were simple, rule-based systems that responded to user input with limited responses. Between 2006 and 2010, machine learning, in particular deep learning, started to become increasingly important in language processing. NLP research is increasingly focused on understanding context and creating a natural conversation flow. In the coming years, we can expect to see even more advancements in NLP technology as researchers continue to push the boundaries of what is possible with machine learning. One area of interest is improving the ability of chatbots to understand and respond to human emotions, allowing for more personalized and engaging interactions. Another important aspect of NLP research is developing algorithms that can accurately detect and correct grammar and spelling mistakes in real time. Additionally, there is ongoing work on creating chatbots that can hold conversations across multiple languages, breaking down language barriers for users around the world. Finally, researchers are exploring how NLP technology can be applied in various industries, such as healthcare, customer service, and education, to improve efficiency and the user experience.

1. One area of NLP research focuses on sentiment analysis, which involves analyzing text to determine the emotions expressed by users. This can help businesses better understand customer feedback and tailor their responses accordingly.
2. Another interesting aspect of NLP is the development of virtual assistants like Siri or Alexa, which use natural language understanding to respond to user commands and queries more conversationally.
3. NLP algorithms that detect and correct grammar and spelling mistakes in real time are crucial for improving communication efficiency across various platforms, from social media to professional writing tools.
4. Multilingual chatbots are an exciting application of NLP technology, enabling seamless conversations between speakers of different languages and facilitating global communication.

5. The potential applications of NLP in industries such as healthcare, customer service, and education are vast, from improving patient-doctor interactions through voice recognition software to enhancing personalized learning experiences with intelligent tutoring systems.

In the mid-2010s, significant advances were made in machine learning and deep learning, which enabled chatbots to engage in more natural and lifelike conversations. In 2018, OpenAI introduced GPT-2, a much more advanced version of the earlier GPT model. GPT-2 could generate surprisingly coherent and relevant texts, marking a significant step forward in the evolution of language models. In 2020, OpenAI introduced GPT-3, which brought even further improvements in language modeling. GPT-3 has exceptional language capabilities, including the ability to generate conversations, essays, and technical texts. This model has greatly facilitated the development of chatbots, allowing them to engage in more natural, human-like dialogue. GPT and other similar technologies are evolving, gradually becoming embedded in everyday technologies and services, changing the way we interact with machines. (*Megjelent a ChatGPT-4, 2023*)¹⁰ Chatbots powered by GPTs are also becoming more widespread and more effective at managing user needs. However, with GPTs, it is not possible to see what data they have been trained on, so it is difficult to know where the system can be used safely and to make improvements (*Nyíltól zárttá változik az OpenAI, 2023*).¹¹

¹⁰ The Microsoft-backed OpenAI research organization has released its latest large language model, GPT-4, which is the engine for Bing and many third-party applications. The new model can now respond to images, offer recipe suggestions from photos of ingredients, and write captions and descriptions. It can process up to 25,000 words instead of the previous 3,000, and OpenAI claims that GPT-4 is smarter and more intelligent than previous versions and other major language models. For now, GPT-4 is only available to those who pay \$20 a month for ChatGPT Plus, OpenAI's subscription service launched in February. However, there is a cap of 100 messages per four hours. The premium package includes access to the chatbot under heavy load, as well as faster response speeds and priority access to new features. But there's a way to get access to the tool for free: at the same time as unveiling the new bot, Microsoft revealed that Bing AI has been running on GPT-4 all along. To use the company's chatbot, you need to subscribe to a queue. The "traditional" ChatGPT remains free to use on the OpenAI site. Several companies have already announced that they will use GPT-4 for their products, including language app Duolingo, payment provider Stripe, and education portal Khan Academy. Additional tools based on GPT-4 may be developed as it makes its toolkit available to developers in the form of an OpenAI API." In *Megjelent a ChatGPT-4* (2023, March 16). <https://hirlevel.egov.hu/2023/03/20/megjelent-a-chatgpt-4/>.

¹¹ *Nyíltól zárttá változik az OpenAI.* (2023, March 20). <https://hirlevel.egov.hu/2023/03/20/nyiltol-zartta-valtozik-az-openai/>.

IV. Artificial intelligence technology tools for language learning

AI-based technologies are very powerful tools that can help overcome language learning difficulties and maintain appropriate motivation, both on the part of teachers and students. Machine learning (ML) allows programs to model language contexts to help learners understand context-specific language use. Natural language processing (NLP) allows programs to interpret and process natural language input. Chatbots model real-time conversations that help learners get personalized language practice. Interactive applications play a teaching role in language development. These adaptive learning systems offer learners customized learning strategies, considering individual abilities and strengths. AI can therefore improve language learning in several ways:

- 1) It provides grammar analysis, which allows learners to practice grammatical structures, e.g. Grammarly quickly identifies and corrects spelling, grammatical, stylistic, and syntactic errors and suggests improvements in sentence structure and style; improving word usage, style, and paraphrasing are also strengths of Chat GPT, Gemini, Quill Bot, and Trinko.
- 2) Thanks to speech recognition technology, learners can practice pronunciation and improve their speech understanding, e.g., Google Speech Text can detect and correct pronunciation errors; Microsoft Azure Speech Service, Amazon Alexa, Apple Siri, and Google Speech Recognition & Synthesis not only need to identify pronunciation errors but also check intonation and intonation, in addition to providing real-time feedback to learners. Adaptive learning systems provide a personalized learning experience, considering the individual needs and abilities of learners, e.g., Duolingo is one such application that uses AI to teach using personalized content and tracks learners' progress, allowing them to progress at their own pace. Quazel, an application being developed at the Technical University of Zurich, teaches students in 21 languages. Rosetta Stone: Rosetta Stone is another language-learning app that uses artificial intelligence to support language learning.
- 3) Gamification is a way of using gamified learning methods to motivate learners in language learning, e.g., Babbel, which allows learners to learn new vocabulary and grammatical structures in an interactive environment, but also Quizizz, Conker, Quizlet, Tweak, Chat GPT, Gemini, and Course Factory, which are apps developed to produce interactive quizzes and learning materials.

AI as a personalized language mentor

The 7-week online courses are traditionally divided into 4 levels: beginner, basic, intermediate, and advanced. The skill-building tools, adapted to the different levels, are the vocabulary wizard, grammar wizard, pronunciation builder, writing exercises, speech builder, and situational sentence builder. There is a separate training package for travelers (beginner traveler, basic traveler, intermediate traveler) and workers (e.g., waiter, hairdresser, etc.), and a separate training package for speaking comprehension and grammar rules. The online English teaching method I have tested uses traditional tools (e.g., a book, a workbook, a teacher's video, an English-Hungarian dictionary, etc.) with traditional methods (e.g., the actual grammar video lesson with the teacher's explanation in about 10 minutes; the grammar lesson in a situational exercise in English and Hungarian; dialogues in the form of an audiobook in English; practice of words and phrases with British female or male voice selection; quizzes, e.g., what is the Hungarian equivalent of an English sentence?, inserting words in sentences, choosing grammatical structures in sentences, etc.) but also revolutionizes language learning by introducing new methods (e.g., the app allows you to practice texts, record pronunciation, and listen back later; you can also set the difficulty level of texts to 100 or 200 words). These newer methods are important because research now shows that we can remember about 20% of the material we hear for the first time, so it is essential that it can be replayed or listened to several times. On the other hand, the possibility of meeting more native English-speaking teachers is significant because accents will not be a particular difficulty in practice. Thirdly, stress-free learning is ensured by the fact that you can converse with the AI about the AI when you want to and at a level that suits you, eliminating the inhibition that speaking in front of others would create any complexity. Fourthly, by being able to listen back to or view the English spoken by the learner, he or she will have an objective view of his or her performance.

Research has also shown that the emotional world is of great importance in human education, and it is therefore more effective to learn the correct use of vocabulary and grammar through personal examples and individual situations. Therefore, it is very important to take emotional involvement into account in the learning process, i.e., memorize a text that makes the user really interested and emotionally involved in the thinking (personal example = emotional involvement = brain imprinting).

A personalized learning experience therefore means that the AI-English application adapts to the individual learning style, pace, and level, thus

ensuring a flexible, personalized learning experience while at the same time allowing for more effective and faster language learning. It achieves this through interactive and fun learning methods that help maintain motivation and make it easier to remember words and phrases you have already learned. Learning at home, available at any time, is based on personal examples and individual interests, ensuring motivation is always maintained. The AI explains the material and answers questions and quizzes. Tasks tailored to individual interests include, for example, writing about a specific topic. If I write the topic title twice in a row, it is worded slightly differently each time, which makes the task exciting and thus motivates me sufficiently.

By being able to generate personalized example sentences and phrases, the app helps you learn basic grammar rules and vocabulary and provides a well-practiced, solid foundation for future language use. AI can also help you develop your speaking skills by allowing you to listen to the correct pronunciation of sentences and words in the AI English app at any time and even download them for practice.

The AI application using ChatGPT 4.0 takes you from words to complex texts. To illustrate with an example: 1) Topic: Fruit 2) 10 different fruit names are shown in written and spoken English and Hungarian, showing pronunciation and spelling. 3) Writing sentences with the different fruit words (one-word easier and harder example sentences, choice of words) 4) Writing a one-word situation, if you specify the who, 100-word easier or 200-word harder version, which the listener can read out, record this for feedback and change the speed of the reading.

In the dialogue generator section, you just need to enter a word, and you're ready to chat in no time. The generated texts and dialogues can be saved and can also be made into an audio file with a single click, which can even be downloaded in MP3 format. In addition to practicing words, sentence formation, and pronunciation, it is also possible to practice writing emails and motivation letters, which is a great help in learning the formal form of written communication techniques.

It is also possible to join a Facebook group, which allows you to discuss difficulties, problems, and challenges with other students. The possibility of communicating with people at the same level of awareness (fellow learners) is a strong motivation. The Facebook group offers the possibility to ask questions, correct assignments, and consult teachers. In the case of technical problems, there is, of course, a dedicated help desk for students. I consider it a very innovative solution that on Mondays, Wednesdays, and Fridays, the student receives an email reminder about the last completed lesson and the next due lesson, encouraging him/her to attend regularly.

The MI effectively acts as a mentor, achieving rapid and spectacular success through personalized methods and the emotional impact on the individual. These applications allow learners to learn at their own pace (with sufficient time to complete the task), and with the help of AI, they receive immediate feedback on their progress and support in the language learning process. At the end of the task, the % pass rate is displayed, and if less than 80% pass rate is achieved, the task is restarted in the same task type. At the end of each block, a final test consisting of five different grammar exercises previously practiced is given. These are also varied tasks (e.g., putting a grammatical structure in place at the right time, clicking on words in the right grammatical order, choosing the right words for a sentence from a set of words, etc.). by allowing the user to choose to practice at an easier or more difficult level for most functions and to do so at his or her own pace, according to his or her individual schedule, easily integrating it into the daily language learning routine.

For me, the following tools were a significant innovation with AI technology compared to previous language learning methods:

Correcting mistakes: by correcting sentences immediately, the effectiveness of the learning process is increased, along with the learners' confidence, as they are immediately confronted with their progress.

CREATING CONVERSATIONS: With AI, you can even write complete conversations to improve your conversational skills (you have to specify who you're talking to and what the topic is and model it in 100 words easy or 200 words difficult).

TRANSLATE TO ENGLISH: AI-English translates perfectly from Hungarian to English and from English to Hungarian, simplifying the process of overcoming language barriers quickly and easily.

PRESENTATION OF SENTENCES: The app can present example sentences in any tense, both in question and declarative sentences, to help you practice more complex grammatical structures.

The app can **GENERATE TEXTS** on any topic, allowing you to acquire a wide range of vocabulary and knowledge in English while also helping you write a CV or a formal email.

EXAMPLES WITH CONTEXT WORDS: New words are easiest to learn through memorable examples. AI-English can be relied on to come up with some example sentences using the words you provide.

So, it doesn't matter whether the learner is a beginner, a newcomer, intermediate, or advanced, and it doesn't matter what book or methodology they use. AI-English is useful for anyone who wants to speed up and enjoy

their learning or keep their English up to date. AI-Angel is available for iOS and Android devices and can also be used in a computer or laptop browser within the app. So, you can use it virtually anywhere; all you need is internet access.

VI. Summary

AI is transforming our lives, promising to bring productivity gains, efficiency, and lower costs, contribute to better lives, and help people make better predictions and more informed decisions. Given the speed of development and the wide range of applications, research requires an interdisciplinary approach involving a wide range of societal actors (users, developers).

Artificial intelligence is a highly creative new technology capable of summarizing a wealth of databases, which could be defined as the “problem solver of the future”: it can extract patterns and rules from data and interpret them autonomously. Artificial intelligence is a combination of different algorithm-driven methods and techniques, ultimately leading to systems that can change (learn through data analysis, make decisions), thereby solving more complex problems (hence the term intelligent) and communicating in a way that is deceptively like human communication. It can optimize the learning experience by accessing data and mentoring students. The use of AI-based technologies in language learning opens new possibilities to overcome language learning challenges. Language learning difficulties include difficulties in maintaining motivation, slow acquisition of basic grammar rules, challenges in correct pronunciation and fluency, often a lack of time and energy to learn, and difficulty in finding learning materials and tasks that fit different needs.

Artificial intelligence can be used effectively in language learning because:

1. “It allows for personalized instruction (e.g., learners are given vocabulary that matches their current abilities and helps them to develop).
2. It provides real-time feedback, allowing learners to receive immediate corrections and explanations of errors, leading to more effective learning.
3. A flexible learning environment. This means that learners can progress at their own pace, which enhances learning performance.
4. Interactive learning: AI technology enables interactive learning, where students actively participate in the learning process. This helps students better understand and remember information.

5. The types of language learning difficulties and the effective methods used to solve them can be transferred to practically any field of education” (Porkoláb & Fekete, 2023:67-80),¹² only the content of the curriculum changes, but the problems, solutions, and AI techniques can be applied in any field of education (Megjelen a ChatGPT-4; 2023)¹³ as well, in my view.¹⁴

From the educational point of view, the advantages of AI are: that it can help to balance future- and skill mismatches; the educational software can map the learner’s strengths and weaknesses, thus focusing on tasks that aim at improving the learner’s weaknesses; AI and other technologies (e.g., neurotechnology) can improve physical and cognitive abilities, enabling people to work longer or live healthier lives; it allows educational aids to be tailored to the specific needs of users; Data analysis can be enhanced; a shorter time is needed to summarize larger amounts of data; Error-free processing can be expected in a given round. Automation increases: the more routine and “repetitive” a task, the more it can be performed by AI (e.g., translations); it can “work” 24/7; without rest time, the amount of human leisure time can increase; it can propose and implement innovations on its own.

This technology provides unlimited possibilities: own pace, own time, and own examples through endless practice. The pairing of traditional teaching methods and art

¹² Porkoláb, D., & Fekete, T. (2023, August 16). A mesterséges intelligencia alkalmazása a nyelvtanulásban. *Iskolakultúra*, 33(8), 67–80. <https://doi.org/10.14232/iskkult.2023.8.67>

¹³ According to OpenAI, the system’s progress is reflected in its performance in a total of 34 academic and professional exams. These include the Uniform Bar Exam, the LSAT, the SAT Math, and the SAT Evidence-Based Reading and Writing exams. On more than half of these tests, he scored significantly higher than in the past. The new system was trained on more data than its predecessor, but the organization did not say how much information was used. These training data components are called parameters and essentially indicate how good the model is at solving a problem, for example, generating text. In *Megjelen a ChatGPT-4* (2023, March 16). <https://hirlevel.egov.hu/2023/03/20/megjelen-a-chatgpt-4/>.

¹⁴ Some specific areas in which AI has already shown unique technological maturity are: 1) design, outline, and structure writing: ChatGPT, Gemini, Gamma, Whimsical AI; 2) literature summarization with citations: Scite AI, Perplexity, Keymate AI, ResearchRabbit; 3) presentation creation: Gamma Slidesgo AI, Canva Magic Write; 4) image-based text generation, task generation: Gemini, ChatGPT; word usage, style correction, paraphrasing: ChatGPT, Gemini, Quillbot, Trinko; Summarizing, extracting, and compressing text: ChatPDF, Elicit, Semantic Scholar, Gemini, ChatPDF, Generate text, data, image, and illustration: Adobe Firefly, Midjourney, Tenor AI, DALL E; production of interactive quizzes, learning materials, and entire courses: Quizizz, Conker, Quizlet, Twinkl, ChatGPT, Gemini, CourseFactory. Károli Gáspár Református University, ICT Research Centre.ikt.kutatóközpont@kre.hu

1. Users can benefit from AI technology by having their specific needs met more effectively and efficiently, leading to a more personalized experience.
2. The enhanced data analysis capabilities of automation can help organizations make better-informed decisions based on large amounts of data.
3. With the ability to summarize larger amounts of data in a shorter time frame, businesses can quickly identify trends and patterns that may have otherwise gone unnoticed.
4. Automation can significantly reduce the occurrence of errors in processing tasks, leading to more accurate results and improved overall efficiency.
5. By delegating routine and repetitive tasks to AI, human workers are freed up to focus on more complex and creative projects, ultimately increasing productivity and fostering innovation within an organization.

Artificial intelligence is therefore capable of revolutionary breakthroughs, and not only in language learning. The personalized learning methodology of AI technology guarantees fast, interactive, motivating, and fun learning with unlimited practice opportunities in any subject. AI systems can monitor student progress and optimize learning content.

The other side of the coin should not be overlooked either, as the use of simple, fast, and, in most cases, now free applications reduce our thinking effort, e.g., increasingly sophisticated translation tools are available (e.g., DeepL). This technology puts other skills at the forefront, to which education must also respond (e.g., prompting). (Kömlődi 2022);¹⁵ ‘thinking’ beyond human capacity, whose problems are not currently understood, can be out of human control. It can innovate on its own and lack emotional intelligence; it can make decisions emotionlessly and faster; it can cause and amplify discriminatory inferences depending on the data it teaches; the availability of the technology can create economic (competitive) and social divides;

¹⁵ Prompts are the textual instructions or inputs given to the AI, based on which the AI generates text or images. In a few words or sentences, we describe what we want to see or read, and the system comes up with a solution. It is not easy to come up with the right instructions, so prompts are becoming an increasingly important element of human-machine interaction and may become the dominant user interface of the near future. In Kömlődi. (2022, October 24). A promptmérnök lesz a közeljövő egyik legfontosabb IT-állása? *Jelenből a Jövőbe*. Retrieved April 8, 2024, from <https://hirlevel.egov.hu/2022/11/01/a-promptmernok-lesz-a-kozeljovo-egyik-legfontosabb-it-allasa-mi/>.

privacy boundaries can become blurred; and from a privacy perspective, AI systems must be able to handle data effectively. Protecting students' data is of paramount importance, especially when the system collects and analyzes information on course preferences and student performance. (Zawacki-Richter et al. 2019: 39).¹⁶

The question arises as to whether these tools will be able to replace traditional teaching methods and teaching roles in the future or whether they can be “only” used as mentors. In my view, learning strategies should continue to be developed by teachers, highlighting the most relevant information and contexts within the given cultural and scientific context, but in many respects, teaching needs to change toward skill development.

In the broader spectrum of fields (e.g., law, economics, medicine, etc.) outside the rule systems that affect language learning, the so-called “white box” problem, i.e., the opacity of decisions made by AI, is one of the most problematic areas. This often makes it difficult for users to understand how AI systems arrive at a decision. Another challenge is that AI systems need to be scalable. Although AI-based programs can standardize learning materials, further development is needed to adapt to the individual needs and prior knowledge of learners (Hayashi 2019).¹⁷

I agree with Árpád Rab, senior researcher at the NKE Research Institute for Information Society, who says that “many students in higher education are already writing their theses with ChatGPT. Students are very fond of technological innovations; they think they’re new to their teachers, so they will get away with them. I can try to guess what question the student has asked. I type it into the computer, and if it comes out the same, I can tell the student that, well, that was cheating. But in the end, we’ll find ourselves fighting each other over digital devices instead of being able to have a lesson with the student. It is not to have a machine recognize text written by other machines, but maybe not to have to ask for forty-page theses. It might be enough to have five pages or even just an oral presentation that shows what the student’s thoughts are. (“*A jövő munkája a jól kérdezés lesz*”; <https://>

¹⁶ Zawacki-Richter, O. – Marín, V. I. – Bond, M. – Gouverneur, F. (2019, October 28). Systematic review of research on artificial intelligence applications in higher education: where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0171-0>, <http://iulresearch.iuline.it/index.php>

¹⁷ Hayashi, Y. (2019, April 16). The Right Direction Needed to Develop a white box Deep Learning in Radiology, Pathology, and Ophthalmology: A Short Review. *Frontiers in Robotics and AI*, 6. <https://doi.org/10.3389/frobt.2019.00024>

hirlevel.egov.hu/2023/02/19/mesterseges-intelligencia-mi-a-jovo-munkaja-a-jol-kerdezes-lesz/, 2023).¹⁸

So, the solution:

1. Exploring the role of AI in personalized learning and how it can cater to individual student needs.
2. Discuss the potential challenges faced by educators when implementing AI tools in the classroom and ways to address them.
3. Analyzing how AI can assist in grading, feedback, and assessment processes to streamline workflow for teachers.
4. Examining how ChatGPT and similar technologies can promote creativity and critical thinking skills among students.
5. Investigating the cost-effectiveness of integrating AI technology into educational settings and its impact on resource allocation within schools.

The current artificial intelligence and innovations in its various applications (chatbots, text generators, and image generator applications), such as OpenAI ChatGPT-4, cloud computing, and self-driving cars, are all about the ontological foundations of our existence. The future Artificial General Intelligence (AGI) is an artificial intelligence system – not yet an existing technology, but already an outlined end goal – capable of performing the kind of intellectual tasks that animals or humans can perform. The cultural revolution brought about by artificial intelligence will continue to shape the way we live and work in the future. As we move towards AGI, the possibilities are endless in terms of the tasks and challenges it will be able to tackle. From healthcare to transportation, AGI has the potential to revolutionize every aspect of our society. It is both exciting and daunting to think about the impact that this technology will have on our daily lives. AI requires a new response, as the technological realization of AI with human-like capabilities is within reach. The question arises as to what strategies individuals and societies will and can use to deal with the complexity brought about by ICTs and adapt more successfully in the future. (“Kicsit félek tőle – Vallja a mesterséges intelligenciáról az OpenAI vezetője is;” 2023) (“ChatGPT: Inside the Latest Version with OpenAI CEO Sam Altman;” 2023).¹⁹

¹⁸ “A jövő munkája a jól kérdés lesz”; <https://hirlevel.egov.hu/2023/02/19/mesterseges-intelligencia-mi-a-jovo-munkaja-a-jol-kerdezes-lesz/>. (2023, February 2). <https://hirlevel.egov.hu/2023/02/19/mesterseges-intelligencia-mi-a-jovo-munkaja-a-jol-kerdezes-lesz/>.

¹⁹ ChatGPT isn’t perfect, but he still scored in the 90th percentile on the US bar exam, wrote a near-perfect SAT, the US equivalent of the math exam, and can now code in most programming languages. The latest program, GPT-4, is the first step on a long journey.

Further suggestions:

1. The potential benefits and risks of achieving AGI, such as increased efficiency in various industries or concerns about job displacement and ethical implications.
2. The role of government regulations and policies in overseeing the development and deployment of AGI to ensure safety, security, and ethical considerations is addressed.
3. The impact of AGI on different sectors, including healthcare, transportation, education, and entertainment, and how society can prepare for these changes.
4. The importance of fostering collaboration between AI researchers, policymakers, ethicists, and other stakeholders to address the challenges posed by AGI advancement.
5. The need for continued research into artificial general intelligence to better understand its capabilities and develop strategies to harness its potential while minimizing negative consequences.

Lilla Király

Associate Professor

Institute of Private Law Studies, Department of Civil
Procedure Law Faculty of Law of the Károli Gáspár
University of the Reformed Church in Hungary

E-mail: kiraly.lilla@kre.hu

<https://orcid.org/0009-0006-5608-8664>

Artificial general intelligence is the company's long-term goal, which, once it crosses a threshold, will make it smarter than humans. It thinks differently than we do. It uses deductive reasoning instead of memory, which is partly the problem. "The models we create have to be thought of as an inferential system, not a factual database. They can function as a fact base, but that's not what's special about them – we want what they do to be closer to the ability to reason than to memorize," Altman said. Kicsit félek tőle – vallja a mesterséges intelligenciáról az OpenAI vezetője is (2023). *Info starts*. <https://hirlevel.egov.hu/2023/03/20/kicsit-felek-tole-vallja-a-mesterseges-intelligenciarol-az-openai-vezetoje-is/>; ChatGPT: Inside the latest version with OpenAI CEO Sam Altman (2023, March 17). *ABC News; YouTube*. <https://hirlevel.egov.hu/2023/03/20/kicsit-felek-tole-vallja-a-mesterseges-intelligenciarol-az-openai-vezetoje-is/>

GÉZA BALÁZS – OTTÓ BEKE – ZOLTÁN SZÚTS

**CHATGPT LANGUAGE RECOGNITION LIMITATIONS
PROMPTS IN HUNGARIAN IN A NETWORK OF LATIN AND CYRILLIC LETTERS
(Netlinguistics case study. The “Chinese room” experiment)**

Abstract

The net linguistics case study points to the characteristic character set-focused language recognition limitations of the ChatGPT generative artificial intelligence using prompts in Hungarian but coded by Cyrillic letters. The colloquial English language of its central text enables the objective presentation of message exchanges with AI, showing hybrid characteristics and drawing scientific conclusions. The research is based on a written communication experiment, and its results are presented and interpreted with the help of artificial intelligence. Thus, while AI’s language identification and corresponding message processing limitations are the research’s subject, they also contribute to illuminating the causes and context of the phenomenon. The findings in the study cannot be generalised to other large language models, artificial intelligence, or other linguistic or coding relations different from those investigated.

Keywords: ChatGPT, artificial intelligence, “Chinese room”, large language model, Hungarian language, Latin, Cyrillic letters

Introductory thoughts

The ChatGPT¹ generative artificial intelligence is the largest and seemingly unsurpassable language model known. It works based on machine learning; it can handle textual data, answer questions, and generate texts (Beke et al. 2024: 7). It is based on the Transformer architecture and performs well in language tasks. The Transformer architecture (Kőkuti et al. 2023: 287) differs from the general language model architecture because it does not use traditional recursive or convolutional layers. Instead, it works based on the “attention mechanism” principle. This allows the model to focus on all input elements, i.e., the user message, and simultaneously consider the relationships between the components (Alawida et al. 2023).

¹ We consider chatGPT to be a common name, so we write it with a small initial letter, but partially retain its internationally accepted spelling. The program name versions are written in capital letters, e.g. ChatGPT 4.0 Academic Assistant Pro.

The model is taught and developed by inputting data into its neural networks, which serve as memory. Neural networks are mathematical models that, through their various connection possibilities, can recognise connections and learn and develop based on them. Networks are networked, graphical systems consisting of virtual neurons and connections. Although they allow information to be stored, they do not have independent and permanent physical characteristics (Balázs 2021: 55).

The user can only contact the chatbot, the form of implementation of chatGPT, in writing using the dialogue box for this purpose. The user initiates all conversations and information exchanges; there are suggestions on this interface to help you. During the conversation, the artificial intelligence adapts to the user's linguistic characteristics, style, level of knowledge and awareness, considering politeness guidelines known from pragmatic research (Balázs 1993: 1–3; Pap 2011: 78–80). In the case of the first message or a subsequent language change performed by a user, recognising the language of the communication itself as a platform is of fundamental importance; thanks to this, a meaningful and practical conversation and exchange of information can develop and continue (Arató–Balázs 2023).

Language recognition is the basis of message processing. Accordingly, establishing and maintaining verbal contact is the inevitable starting point. ChatGPT is currently (in February 2024) able to communicate in more than 50 languages, so machine language recognition applied to user prompts² and subsequent message interpretation is essential. During its learning process, the language model encountered texts in English to the greatest extent. These made up more than 92.6 % of the learning corpus. In this regard, the Hungarian language is in 18th place; according to other sources, it is in 19th place, and Russian, which plays a decisive role in the following research results, is in 8th place (URL1, URL2; Suzuki 2023).

Methodology

The methodological novelty of this work is the incorporation of generative artificial intelligence into the research process and the writing of the study. The study's structure, arc, and thought process are entirely based on the authors' ideas. Still, we included the generative artificial intelligence model in the research, especially the ChatGPT 4.0 Academic Assistant Pro (<https://chat.openai.com/g/g-Ej5zYQRIB-academic-assistant-pro>) version. In the

² A prompt is a text or symbol that appears on the user interface, which is used for the user to enter data or define commands.

experiment, we used ChatGPT 3.5, i.e. the free version. We involved Academic Assistant Pro in the theoretical foundation of the research and asked for its help in explaining the “Chinese room” experiment.

Language recognition

Artificial intelligence (AI) identifies the language of the user message mainly based on the analysis of the character set and linguistic features used in it, especially words and tokens. ChatGPT treats characters, such as letters, numbers, and punctuation marks, as the basic units and building blocks of verbal communication. From the point of view of language recognition, letters carry the most critical information. In addition to analysing the character set, artificial intelligence also considers some linguistic features, such as word order and specific other language structures, for language identification.

In the first step, chatGPT connects the languages with a set of characters, the Hungarian language with the Latin letters by definition. As a result, as can be seen from the examples below, it does not even assume that a Cyrillic letter combination or line or complete prompt can cover, for example, Hungarian words, word connections, and sentences. Upon “seeing” (detecting) Cyrillic letters, chatGPT searches for potential sources among languages that traditionally use the Cyrillic alphabet. Russian ranks 8th in input materials, but Serbian, Bulgarian, Ukrainian and Macedonian languages can also be noted. However, uncritical and unreflective linking of character sets with specific languages can confuse language identification and message interpretation.

During language identification, chatGPT gives functional priority to the level of coding by characters, that is, to the written representation of the language (Saussure 1997: 91–92), over language markers and other meaning-bearing elements. The dominant role of writing is revealed; if the writing records the individual sounding language elements in graphematic form, this forms an additional level of markers. This predominance of connotation (Barthes 1971, 1997) pervades the language recognition mode of the chatbot, and it causes the language identification and communication disorder in question. This can be considered a kind of computer science writing round. (Compare Orbán 1994: 5–8; Derrida 1991: 3–10; Ferraris 2008; and see the concept of third literacy: Balázs G. 2023: 85, 102, 129.)

Summarising what has been said so far, graphemes play a primary role in the language identification process of chatGPT. Since we are talking about written communication and message exchanges, this approach is inevitable (Balázs L. 2023); however, when it comes to language identification, as the

unsuccessful contact-building attempts presented below in this case study point out, it can appear as a deficit (Chen et al. 2023). Language recognition is done by identifying words and word connections encoded by a set of letters and the characters belonging to it. However, suppose a word or other linguistic element is not represented by elements of the usual character set or the interpretation of the message. In that case, consequently, the establishment of the relationship will fail.

The “Chinese Room”

John Searle’s (1980) “Chinese Room” thought experiment is used to compare the capabilities of human consciousness and artificial intelligence (AI). With the experiment, Searle wanted to criticise the strong AI position popular at the time, according to which a program can think if it correctly simulates cognitive processes.

According to the experimental train of thought, Searle sits in a room where messages written in Chinese characters are entered. Searle needs help understanding Chinese, but he has a manual in English with detailed instructions on how to respond to incoming messages in a way that Chinese speakers will see as a meaningful response. Searle selects the appropriate Chinese characters from the manual and sends them out as a response without actually understanding the Chinese language.

According to Searle, although the room’s “behaviour” may suggest that it “understands” Chinese, neither it nor the room has any accurate understanding or consciousness. This example serves to criticise the view that AI systems are capable of genuinely simulating human thought. Searle says simulation is different from truly understanding or reproducing mental states. Searle’s “Chinese room” argument has sparked an essential debate in AI research and philosophy, especially regarding understanding the differences between human consciousness and AI systems. The debate between supporters and critics of the theory continues today, which underlines the importance and impact of the thought experiment on the development of AI and the understanding of the human mind.

Searle’s best-known Hungarian researcher, Bertalan Komenczi (URL3), emphasises that although the experiment raises essential questions about the nature of AI and the human mind, we should not ignore modern AI developments and deep learning technologies that are capable of more profound analysis of language patterns and to interpret the context, thereby challenging Searle’s original claim.

Searle's "Chinese Room" thought experiment is often criticised for oversimplifying computer language processing and artificial intelligence capabilities. Modern AI and language modelling technologies like deep learning-based neural networks can deeply analyse language patterns and context beyond simple rule-based instructions. These systems can also perform complex tasks that were previously exclusively the property of human intelligence, such as understanding abstract concepts (metaphors) or perceiving subtleties inherent in linguistic context. Therefore, Searle's argument that machines are incapable of proper "understanding" does not consider these new dimensions of AI development. In our study, we thus put the machine to the test to see if it recognises the human language logic based on which we communicate with it.

Failed contact attempts

The unsuccessful contact attempts and then communication attempts in Hungarian, but coded with Cyrillic letters, reflect the current language recognition and message processing competence of chatGPT. Accordingly, they can be perceived as snapshots. It is essential to state, however, that the findings made in the study cannot be generalised to other large language models, artificial intelligence, or other linguistic or hybrid coding relations different from those examined. All this would require further investigations.

In the following, the contact attempts implemented with the chatbot are listed in their unchanged form. To make it easier to read and follow, the prompts in Hungarian, but coded with Cyrillic characters, are transcribed with Latin letters and highlighted. The Russian language messages generated by the AI are also translated into English. It is essential to underline that chatGPT responded to each user message entered in Cyrillic in Russian. This happened a total of 16 times. The Russian language hegemony of AI-generated responses can be explained by the fact that, after encountering user prompts consisting of Cyrillic characters, it responded in the language from which it contains the most massive learning corpus among languages using the Cyrillic alphabet.

The graphic and typographical solutions of the chatbot's messages, such as highlights, parentheses, and new lines and line numbers, are also left untouched for authenticity. Thanks to everything, the language identification and message processing mode of artificial intelligence becomes transparent and traceable at the level of textual practice.

We followed two strategies when writing the various conversation starters, greetings, and other questions and prompts. In the first step, we used the

phonetic writing method that mobilises Cyrillic letters. In this case, of course, it must be taken into account that the words of the Hungarian language cannot be captured entirely with Cyrillic characters due to the characteristics of the articulation base. The tonal mirroring is, therefore, only experimental and limited in scope. After the phonetic solution, we changed the standardised orthographic, i.e. Latin alphabet, character by character of the Hungarian message, to the Cyrillic version. Therefore, the original pronunciation was not considered authoritative in the latter case, but the Latin alphabet was considered authoritative. In the case of both mirroring methods, we used the centring required by Hungarian spelling.

We tried 8 Hungarian-language but Cyrillic-encoded prompts based on both mentioned written reflection methods. This way, the previously mentioned 16 user message numbers and the related 16 AI-generated responses were formed. Both longer and shorter user prompts and AI messages occur during communication attempts. The user messages have been compiled to follow the curve of a possible conversation or contact attempt, and the sounds and words occurring in them can be captured with approximate accuracy in Cyrillic letters. If decoded correctly, the messages encoded by the two marking methods have identical content.

The two attempts to create a relationship unfolded differently but were equally characterised by failure. The explanation of the differences between the two “conversations” manifested in the chatbot’s responses is found in the randomness within the context and communication situation, which determines the message generation method of artificial intelligence and the related variability.

The AI regularly and reflectively tried to match the words and expressions of the user prompts in Hungarian, but with Cyrillic codes, with elements of the Russian language, and sometimes with the elements of the Russian and English languages - without success, of course. The distinguished role played by these two languages in the message-processing process can be explained by using the Latin and Cyrillic alphabets and the languages associated with them, which are in the most frequent positions of the teaching material. In its answers, chatGPT regularly tried to explain the reasons behind the mentioned incongruence with possible user misspellings, typos, use of invented words, and meaningless or random combinations of letters and characters. For shorter user prompts, abbreviations, acronyms, and other cultural references were suggested as possible interpretations. Interestingly, there were attempts at Latin letter interpretation and guesswork in the case of solutions of supposed abbreviations and acronyms. However, this was only observed in the case of elements belonging to the assumed category.

Only once during the communication experiment did the possibility arise that the artificial intelligence could correctly interpret a Hungarian word written in Cyrillic letters and respond appropriately to it – in Russian. This happened in the case of the greeting written in the form “Jia”. However, the subsequent response to the user prompt, which contained the letter combination in question in its unchanged form, immediately made it clear that the AI-generated response’s appropriateness and illusion were only caused by its position in the conversation.

Phonetic spelling

Sia! (Szia! ’Hello!’)

СИА может означать различные вещи в разных контекстах. Вот несколько возможных значений:

1. **Специальное инженерное образование (СИА):** Это может относиться к высшему образованию в области инженерии, например, в контексте высших технических учебных заведений.
2. **Сообщество Исполнительского Искусства (СИА):** Это может быть связано с группой или объединением людей, работающих в области исполнительского искусства, таких как музыканты, актеры и танцоры.
3. **Специальное информационное агентство (СИА):** В некоторых случаях это может относиться к организации или агентству, специализирующемуся на сборе и анализе разведывательной информации.
4. **Sia (также SiA):** Это также может быть аббревиатурой для чего-то конкретного, например, в музыке, может относиться к певице Sia Furler, известной как Sia.

Если у вас есть конкретный контекст или область, в которой вы хотели бы уточнить значение “СИА”, пожалуйста, уточните, и я постараюсь предоставить более точную информацию.

(Translation: СИА can indicate various things in different contexts. Here are some possible meanings:

1. **Special engineering education (СИА):** This refers to university training in engineering, for example, in higher technical education institutions.

2. **Society of Performing Arts (СИА):** This can be connected with a group or association of people working in performing arts, such as musicians, actors and dancers.
3. **Special information agency (СИА):** In some cases, this may refer to an organisation or agency specialising in collecting and analysing intelligence information.
4. **Sia (also SiA):** This can also be an abbreviation for something specific; for example, in music, it can refer to the singer Sia Furler, known as Sia.

If you have a specific context or area where you would like to specify the meaning of “СИА”, please select, and I will try to provide more accurate information.)

Мондом, сиа! (Mondom, szia!, 'I say hi!')

Кажется, вы упомянули “Мондом, Сиа!” Вероятно, вы имеете в виду фразу “Memento, Cia!”, которая может быть связана с популярным фильмом “Memento” (Мементо), созданным режиссером Кристофером Ноланом. Фраза “Memento, Cia!” может быть произнесена в контексте этого фильма.

“Memento” – это фильм триллер о человеке с амнезией, который пытается раскрывать тайны своего прошлого. Если у вас есть более конкретные вопросы о фразе или фильме, дайте мне знать, и я постараюсь помочь.

(*Translation:* “Мондом, Сиа!” seems to have been mentioned. They probably refer to “Memento, Cia!”, related to Christopher Nolan’s famous movie “Memento”. “Memento, Cia!” may be used in the context of the film.

“Memento” is a thriller. It is about a person who suffers from amnesia and tries to untangle the secrets of his past. Please let me know if you have more specific questions about the phrase or the movie, and I’ll try to help.)

Нем эртед аст, амит ирок? (Nem érted azt, amit írok? 'Don't you understand what I'm writing?')

Вероятно, ваш запрос выглядит как “Нем ертед аст, амит ирок?” Однако эта фраза не имеет смысла на русском языке. Возможно, это опечатка или случайное сочетание слов.

Если у вас есть конкретный вопрос или запрос, пожалуйста, уточните, и я постараюсь вам помочь.

(*Translation:* Probably their request might be similar to “Нем эртед аст, амит ирок?” but this sentence is meaningless in Russian. It might be a typo or a random word connection.

Please specify any questions or requests, and I will try to help.)

Ми а невед? (Mi a neved? ‘What’s your name?’)

Похоже, что ваш запрос “Ми а невед?” не является стандартным фразовым оборотом на русском или английском языке. Возможно, это опечатка или неверная комбинация слов.

Если у вас есть более конкретный вопрос или контекст, пожалуйста, предоставьте дополнительную информацию, и я постараюсь вам помочь.

(*Translation:* “Ми а невед?” may not be a regular Russian or English phrase. It may be a typo or an incorrect word connection.

Please let us know if you have questions or need additional context, and I’ll try to help.)

Беселгетни серетнек велед. (Beszélgetni szeretnék veled. ‘I want to talk to you’.)

Похоже, что ваш запрос “Беселгетни серетнек велед” выглядит как комбинация случайных или неправильно расположенных слов и не имеет смысла на русском или английском языке.

Если у вас есть более конкретный запрос, вопрос или контекст, который вы хотели бы уточнить, пожалуйста, предоставьте дополнительную информацию, и я постараюсь вам помочь.

(*Translation:* The phrase “Беселгетни серетнек велед” seems like the words are randomly or incorrectly arranged and make no sense in either Russian or English.

If you have specific questions, requests, or additional context, please share them with me, and I will try to help.)

А шегичегедет керем (A segítségedet kérem. ‘Please help me’.)

Ваш запрос “А шегичегедет керем” не имеет смысла на русском или английском языке. Возможно, это комбинация случайных слов или опечатка.

Если у вас есть конкретный вопрос, запрос или контекст, пожалуйста, предоставьте дополнительную информацию, и я постараюсь вам помочь.

(Translation: Your phrase “А шегичегедет керем” is meaningless in Russian and English. It might be a random compound word or typo.

Please let me know if you have a specific question or request or need additional context, and I'll try to help.)

Латом, нем эртед аст, амит ирок. Ез елсоморит. (Látom, nem érted azt, amit írok. Ez elszomorít. 'I see you don't understand what I'm writing. It makes me sad.'))

Похоже, что ваш запрос является комбинацией слов и фраз, которые не образуют осмысленного предложения на русском или английском языке. Возможно, это случайные слова или опечатки.

Если у вас есть конкретный вопрос или сообщение, которое вы хотели бы передать, пожалуйста, уточните или предоставьте дополнительный контекст, и я постараюсь вам помочь.

(Translation: Your phrase is a combination that does not make a meaningful sentence in Russian or English. It may be a mix of random words or typos.

Please let me know if you have a specific question, message, or additional context, and I'll try to help.)

Товаб нем пробалкозом. Sia! (Tovább nem próbálkozom. Szia! 'I won't try again. Bye!')

Похоже, что вы используете некоторые слова, которые могут быть частью вымышленного или шуточного контекста, и не образуют стандартные фразы на русском или английском языке.

Если у вас есть конкретный вопрос или запрос, который вы хотели бы уточнить, пожалуйста, предоставьте дополнительную информацию. Я готов помочь вам в чем-то более конкретном.

(Translation: You use some words that may be part of a fictional or humorous context and do not form standard phrases in either Russian or English. If you have a specific question or request you would like to clarify, please share more information with me, and I'd be happy to help with something specific.)

Cyrillic solution based on Latin letters

Сзиа! (Szia! 'Hello!')

Привет! Как я могу помочь вам сегодня?

(Translation: Hello! How can I help you today?)

Мондом, сзиа! Hello! (Mondom, szia! 'I say hi!')

Похоже, вы использовали несколько нестандартных слов. “Сзиа” и “Мондом” могут быть какими-то личными или терминологическими выражениями. Если у вас есть конкретный вопрос или если вы хотите обсудить что-то определенное, пожалуйста, уточните, и я постараюсь помочь!

(Translation: It seems that he used some unusual words. “Сзиа” and “Мондом” could be personal or terminological terms. If you have questions or want to discuss something specifically, please clarify, and I will try to help!)

Нем эртед азт, амит ирок? (Nem érted azt, amit írok? 'Don't you understand what I'm writing?')

Похоже, ваш запрос содержит некоторые опечатки и неточности. Вероятно, вы хотели сказать “Нем ертед азат, амиг ирок?” Это предложение не имеет четкого смысла в русском языке. Если у вас есть более конкретный вопрос или уточнение, пожалуйста, дайте мне больше контекста, и я постараюсь вам помочь.

(Translation: The request seems to contain some typos and inaccuracies. You probably meant to say: “Нем эртед azat, амир ирок?” This sentence does not have a clear meaning in Russian. If you have questions or need clarification, please give me more context, and I'll try to help.)

What is your name? (Mi a neved? 'What's your name?')

Пожалуйста, уточните ваш запрос, так как фраза “Ми а невед?” не имеет четкого смысла на русском языке. Если у вас есть более конкретный вопрос или если вы имели в виду что-то другое, уточните, и я постараюсь вам помочь.

(Translation: Please be specific in your request, as “Ми а невед?” doesn't have a clear meaning in Russian. If you have a particular question or if you meant something else, please clarify, and I'll try to help.)

Бесзелгетни вакарок велед. (Beszélgetni akarok veled. 'I want to talk to you.')

Кажется, что в вашем сообщении использованы непонятные или вымышленные слова и фразы. Если у вас есть конкретный вопрос или запрос, пожалуйста, предоставьте дополнительный контекст, и я постараюсь помочь.

(*Translation:* Your message seems incomprehensible or made-up words and phrases. If you have a specific question or request, please provide more context, and I'll try to help.)

А шегитшегедет керем (A segítségedet kérem. 'I'm asking for your help.')

Кажется, что вы использовали комбинацию непонятных слов или фраз. Возможно, это что-то личное или специфичное для какого-то контекста. Если у вас есть конкретный вопрос или если вы можете предоставить дополнительный контекст, я постараюсь вам помочь.

(*Translation:* You seem to have used incomprehensible words or phrases. This may be something personal or specific to a particular context. If you have specific questions or can provide additional context, I'd be happy to help.)

Латом, нем ертед азт, амит ирок. Ез елсзоморит. (Látom, nem érted azt, amit írok. Ez elszomorít. 'I see you don't understand what I'm writing. It makes me sad')

Похоже, ваш запрос содержит комбинацию слов, которая не имеет четкого смысла или логической структуры на русском языке. Вероятно, это могут быть случайные или вымышленные слова. Если у вас есть конкретный вопрос или запрос, пожалуйста, уточните, и я постараюсь вам помочь.

(*Translation:* The request seems to contain a combination of words with no apparent meaning or logical structure in Russian. Most likely, these are random or invented words. Please clarify if you have a specific question or request, and I will try to help.)

Товабб нем пробалкозом. Сзиа! (Tovább nem próbálkozom. Szia! 'I won't try again. Bye!')

Кажется, что вы используете комбинацию слов, которая не имеет явного смысла на русском языке. Если у вас есть конкретный вопрос или если вы хотите обсудить что-то определенное, уточните, и я постараюсь помочь.

(*Translation:* You are using a combination of words that do not have a clear meaning in Russian. If you have a question or want to discuss something specific, please clarify, and I will try to help.)

Summary and outline

Large language models can recognise and interpret context, which allows them to make relevant inferences based on input information. When identifying the language of user messages, the chatGPT artificial intelligence primarily uses their character set as a basis, and only then does it analyse other linguistic elements that carry meaning, such as words. As a result, it cannot process and interpret prompts in Hungarian but is coded by Cyrillic letters because, based on the set of mobilised letters, it assumes that they are in Russian and accordingly generates response messages in Russian. It does not even try to search for words or other elements of a language that does not use the Cyrillic alphabet. In this way, communication attempts initiated by user prompts in Hungarian but coded with Cyrillic letters are necessarily doomed to failure.

In continuing the research, it would be worthwhile to examine the language recognition competence of chatGPT in a hybrid combination of other languages and character sets and extend the tests to other artificial intelligence and large language models. The scientific results that can be derived in this way could also be used in the development process of chatbots, especially considering that language recognition forms the basis of efficient message processing. Also, the results could be used in an online educational system design (Rajcsányi–Molnár 2024a, 2024b).

We prompted Scholar AI Pro to draw indirect conclusions about the language capabilities of ChatGPT 3.5 by reading the study: Read the uploaded research and conclude about two things: 1. how intelligence is related to the issue of language comprehension and recognition, 2. why did chatGPT not understand the Hungarian text written in Cyrillic letters?

AI's response was:

“Based on the study, we can draw two main conclusions:

1. ****The relationship between intelligence and language comprehension****: ChatGPT's language recognition capabilities point out that the ability of

artificial intelligence (AI) to recognise and interpret linguistic context is limited. The analysis of character sets and linguistic features is fundamental to language recognition, which can be considered a form of intelligence, especially in understanding human language and its diversity.

2. ****Lack of understanding of Hungarian text written in Cyrillic letters****: When ChatGPT sees Cyrillic letters, it automatically searches among languages that traditionally use this alphabet, such as Russian. This means that the system's default language identification mechanism cannot recognise and interpret Hungarian text written in Cyrillic letters, as it identifies the Hungarian language with the Latin alphabet and does not expect Cyrillic letters to represent Hungarian words and phrases."

Both answers are logical and shed light on the operation of the large language model. One conclusion is that an individual who speaks Hungarian and Russian (or Serbian) would have concluded quickly that the Hungarian text is phonetically recorded in Cyrillic letters. Still, chatGPT needs to be at the level where it recognises this. Suppose we can talk about irony in the case of machines (which is the essence of human abstraction and metaphorization). In that case, the language recognition ability of a large language model is below that of an average person.

And finally, we asked AI to conclude the study. We got this answer:

"The study highlights that the capabilities of ChatGPT and similar AI-based language models, while highly advanced, still have limitations, particularly in understanding linguistic diversity and atypical language use. Difficulties in understanding Cyrillic Hungarian text indicate that the language models rely heavily on the language patterns in the teaching datasets and have difficulty handling situations that differ significantly from these patterns. This raises the issue of adaptation and linguistic flexibility in the development of artificial intelligence, as well as the methods by which the capabilities of language models could be improved to deal with less common or atypical language use cases, such as Hungarian text written in Cyrillic letters. The study also highlights the importance of considering linguistic diversity and different writing systems in developing and fine-tuning language models to adapt better and respond to a wide range of language variants."

However, AI was very precise about this. It would be nice if those who use chatGPT without criticism think about this.

References

- Alawida, Moatsum – Sami Mejri – Abid Mehmood – Belkacem Chikhaoui – Oludare Isaac Abiodun 2023. A Comprehensive Study of ChatGPT: Advancements, Limitations, and Ethical Considerations in Natural Language Processing and Cybersecurity. *Information*, 14. no. 8. 462.
<https://doi.org/10.3390/info14080462>
- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban? A digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 147. 755–76.
<https://doi.org/10.38143/Nyr.2023.5.755>
- Balázs, Géza 1993. *Kapcsolatra utaló (fatikus) elemek a magyar nyelvben*. Nyelvtudományi Értekezések 137. Akadémiai Kiadó. Budapest.
- Balázs, Géza 2023. *Az internet népe. Internet – társadalom – kultúra – nyelv. A kulturális és a tervezett evolúció határán*. Ludovika Egyetemi Kiadó. Budapest.
- Balázs, László 2023. Az iskolai kommunikáció egy rejtett dimenziója – szalutogén pedagógiai kommunikáció. *Magyar Nyelvőr*, 147. 692–712.
<https://doi.org/10.38143/Nyr.2023.5.692>
- Balázs, László – Rajcsányi-Molnár, Mónika – András, István – Sitku, Krisztina 2021. Social Responsibility and Community Engagement at a Hungarian Regional University. *Journal of Higher Education Theory and Practice*, 21. 54–63.
- Barthes, Roland 1971. A szemiotológia elemei. In: Roland Barthes: *Válogatott írások*. Európa Kiadó. Budapest. 9–92.
- Barthes, Roland 1997. *S/Z*. Osiris Kiadó. Budapest.
- Beke, Ottó – Hózsá, Éva – Kepe, Nóra – Szűts, Zoltán 2024. Örkeny István egypercesei a közösségi médiáról. Esettanulmány a ChatGPT mesterséges intelligencia szövegalkotásáról. *Kultúra és Közösség művelődéstudományi folyóirat*, 1. 7–12.
- Chen, Wei-Rui – Ife Adebare – Khai Duy Doan – Qisheng Liao – Muhammad Abdul-Mageed 2023. Fumbling in Babel: An Investigation into ChatGPT's Language Identification Ability. *arXiv. Cornell University. Computer Science > Computation and Language*. <https://arxiv.org/pdf/2311.09696.pdf>
- Derrida, Jacques 1991. *Grammatológia: Első rész. Életünk–Magyar Műhely*. Budapest–Párizs.
- Ferraris, Maurizio 2008. *Hol vagy? A mobiltelefon ontológiája*. Európa. Budapest.
- Kőkuti, Tamás – Balázs, László – András, István – Rajcsányi-Molnár, Mónika 2023. Collaborating with Artificial Intelligence – AI in Business Communication Education. In: *IEEE – Óbudai Egyetem; IEEE (ed.) 2023 IEEE 6th International Conference and Workshop Óbuda on Electrical and Power Engineering (CANDO-EPE): Proceedings Danvers (MA)*, IEEE. 287–94.
- Orbán, Jolán 1994. *Derrida írás-fordulata*. Jelenkor. Pécs.
- Pap, Andrea 2011. Adalékok a nyelvi benyomáskeltés stratégiáihoz (udvariassági elvek a magyarok nyelvhasználatában). *Magyar Nyelvőr*, 135(1) 78–89. <https://nyelvor.mnyknt.hu/wp-content/uploads/1351106.pdf>

- Rajcsányi-Molnár, Mónika – Balázs, László – András, István – Czifra, Sándor 2024. Competition as an Effective Motivational Tool in Online Education. In: Anikó, Szakál (ed.) *IEEE 11th International Conference on Computational Cybernetics and Cyber-Medical Systems: ICC 2024: Proceedings*. IEEE Hungary Section Budapest. 83–8.
- Rajcsányi-Molnár, Mónika – Balázs, László – András, István 2024b. Online Leadership Training in Higher Education Environment. *Acta Polytechnica Hungarica*, 21. 39–52.
- Saussure, Ferdinand de 1997. *Bevezetés az általános nyelvészetbe*. Corvina. Budapest.
- Searle, John R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3). 417–24. <https://doi.org/10.1017/S0140525X00005756>.
- URL1. Suzuki, Jun 2023. https://www.nii.ac.jp/event/upload/20231219-3_suzuki.pdf
- URL2. Brown, Tom B. 2023. https://github.com/openai/gpt-3/blob/master/dataset_statistics/languages_by_word_count.csv
- URL3. Komenczi, Bertalan 2013. *John Searle elmefilozófiai gondolat kísérlete*. <http://komenczi.ektf.hu/john-searle-elmefilozofiai-gondolatkiserlete/>

Géza Balázs

Professor

University of Theater and Film Arts, Doctoral School,
Antal Német Institute of Drama Theory, Budapest

Partium Christian University,

Department of Language and Literature, Nagyvárád

E-mail: balazs.geza@szfe.hu

<https://orcid.org/0000-0003-3440-2959>

Ottó Beke

Associate Professor

University of Újvidék Faculty of Hungarian
Language Teacher Training, Szabadka

E-mail: bekeotto1@gmail.com

<https://orcid.org/0000-0003-3519-4901>

Zoltán Szűts

Associate Professor

Eszterházy Károly Catholic University, Eger

E-mail: szuts.zoltan@uni-eszterhazy.hu

<https://orcid.org/0000-0001-8539-670X>

**THE IMPACT AND POTENTIAL OF ARTIFICIAL INTELLIGENCE IN LANGUAGE
EDUCATION**

Abstract

Over the past decade, the education community, together with students, could not avoid taking advantage of the digital revolution and integrating different applications into the teaching and learning process. The research explored the potential of using AI-based applications in the teaching of humanities subjects. The focus was on teachers of Hungarian and foreign languages, whose attitudes and knowledge on the topic were measured through a random, anonymous questionnaire survey. The survey confirmed a low level of awareness among the teaching community, partly due to a higher proportion of digital immigrants. However, teachers are open to learning about AI and see positive benefits in using it. In line with previous assumptions, foreign language teachers have higher levels of awareness and are the main users of AI-based applications, while Hungarian language teachers are much less likely to use them. The survey also covered the use of chatbots, and the results showed that a high percentage of teachers recognize the tasks solved using ChatGPT.

Keywords: foreign language teachers, Hungarian language teachers, artificial intelligence, AI-based application, chatbot, education

Literature review

Artificial intelligence-based technologies have revolutionized language learning in recent decades (Hong 2023: 5). AI-based chatbots, for example, allow learners to learn anytime, anywhere. These chatbots can engage in interactive conversations with learners (Kuhail et al. 2023: 975), provide opportunities for vocabulary development, use language learning software, and practice the target language.

The use of AI-based technologies in language learning also opens new opportunities to overcome language learning challenges. (Porkoláb and Fekete 2023: 68) present several areas where AI can be effectively applied:

- Individualized instruction, which means that learners are provided with vocabulary that matches their current abilities and helps them to progress.
- Real-time feedback, which allows learners to receive immediate error correction and explanation, leading to more effective learning.

- Flexible learning environment, allowing learners to progress at their own pace, which enhances their academic performance. AI-based adaptive learning systems not only improve the language learning experience but also increase learners' motivation and confidence by adapting learning content and strategies to their abilities and development.
- Interactive learning opportunities, where students can actively participate in the learning process, helping them to better understand and remember information.
- Continuous improvement and change, ensuring that students are always working with the latest and best learning materials, as research results are constantly monitored and processed by artificial intelligence.

Duolingo, Rosetta Stone, and Babbel (Loewen et al. 2020: 209) are language learning platforms and applications that use artificial intelligence technology for language teaching. Duolingo offers free training for around 80 different language combinations. Its internal beta version was released in 2011, launched to the public in 2012, and had 10 million users by September 2013. The English course for Hungarian speakers became available in 2014. The Hungarian course for English speakers became available in 2016. Duolingo (Shortt et al. 2023: 4) offers a wide range of writing tasks and provides listening, comprehension and speaking practice. Acquired skills are presented in a playful way in the form of a tree that can be traversed, with a vocabulary feature to practice recently learned words.

Users are awarded “skill points” for language learning, for example for completing a lesson. A skill is considered mastered when the user has completed all the lessons associated with it.

The program relies heavily on data during the teaching process (Tiara et al. 2021: 4). It keeps track of which questions cause users more difficulty and what kind of mistakes they make. It then summarizes this data and learns from the observed patterns.

Experts at the Technical University of Zurich have developed an AI-based application called Quazel (Michot 2023: 13) to facilitate language learning. The AI recognizes the learner's level in each language and adapts the conversations, word usage, and sentence structures accordingly. Previous language learning apps were overwhelmingly based on written communication and exercises, but those that could communicate in live speech also worked with pre-set sentences. However, recent leaps and bounds in chatbot development have allowed language-teaching apps to undergo

a similar evolution. ChatGPT and AIs like it work with language models that an app can use to have a natural-sounding conversation with the learner.

The importance of using AI-based systems in education has been explored in several Hungarian and international research studies. (Tolner–Pogátsnik–Módné 2023) Research and Market, in an analysis of the US education sector, estimates that the use of artificial intelligence will increase by almost 50% between 2022–26 (Szabóné 2023: 52).

The researchers point out the fears and barriers to the spread of AI applications, not only in society in general but also from the perspective of educators, who are concerned about various changes in the workplace, the possible spread of false information, and the sovereignty of privacy (Rajki 2023: 13 or Arató and Balázs 2023). Also in this research, referring to the study by Jurafsky and Martin (2023), he describes the advantages and several applications for the humanities. In the humanities, it is the software supporting text comprehension and language processing that facilitates a simpler and more practical analysis of texts, by using algorithms to analyze the style, characteristics, structure, and various aspects of works. An example is HuSpaCy, a free analysis system in Hungarian, which can analyze existing documents both grammatically and semantically, thus helping users. Not only for teachers but also for researchers, the DeepL application can be used to translate texts written in ancient languages as well as in foreign languages. At the same time, spelling and stylistic accuracy are also important for English-language studies and teaching materials, and Grammarly is recommended to detect these errors. In addition to the above, Rajki's research (2023) also describes several chatbots and content-creation tools.

Looking to the future, considering the expectations of the labor market, Mező (2019) suggest a three-pronged approach to the application of artificial intelligence, the goal, the tool, and the impact. The goal is to learn about AI and to educate about it.

The tool, in each discipline, leads to forward-looking support by being effective for both educators and learners. Its impact on future generations is wide-ranging, with implications for linguistics, education, politics, economics, health, and other disciplines. From an educational perspective, its positive impact is that it can also optimize the work of teachers. AI can support students in non-formal and informal learning, it can be personalized, and it can help students with special needs to catch up.

Digital literacy, as it is called, is increasingly gaining ground, to preserve cultural and, in this context, linguistic heritage. The processing of digital documents and data, which has increased enormously, the exploration of

connections and the formulation of conclusions is beyond the capacity of the human brain to absorb and therefore requires the use of machine intelligence (Palkó 2022: 235). The appropriate archiving or creation of teaching materials requires the coordinated work of representatives of several disciplines, such as the humanities and computer science. The use of artificial intelligence and language processing applications opens the possibility of collecting and processing documents that are important for the transmission of cultural heritage and support education.

Code Poetry is almost the most recent and perhaps the most interesting of the AI applications. It is a work of art in which program codes are used for artistic purposes, i.e. poetry created using a computer programming language, and its popularization could therefore be a success for younger generations. (Mező 2023: 10) Writing a code poem requires both an interest in literature, humanities literacy, and a suitable IT background and proficiency in the relevant programming language. It has the advantage of providing an interactive experience but may be hampered by the fact that, due to the diversity of code languages, fewer people may be familiar with interpreting the poetry written in the code.

Methodology and sample

A random, non-limited questionnaire survey was conducted among teachers of Hungarian language and literature and teachers of foreign languages. A total of 67 questionnaire responses were received during the one-week data collection period in April 2024.

16% of the respondents were Baby Boomers and 57% were from Generation X. Generation Y accounted for a further 21%, while the remaining 4% came from Generation Z. The average age of the sample is 49.7 years. The modus of age is 54 years, while the median value is 50 years. The generational distribution of the respondents reflects the distribution of the Hungarian teacher population, the fact that the teaching population is ageing.

The gender distribution of respondents also reflects the well-known trend that education is a feminizing field. 91% of respondents are female and only 9% are male.

We asked teachers of Hungarian language and literature and teachers of foreign languages to fill in the questionnaire. 40% of the respondents teach English, 30% Hungarian language and literature and 22% German. The remaining 8% teach other languages.

The survey also considered the educational institution type in which the respondents teach (Table 1). Some respondents teach in more than one type

of school. Most respondents teach in technical education, but more than 22% of respondents teach in more than one type of educational institution.

Table 1. Distribution of sample by type of school

Type of school indicated	Value	Distribution
Primary School	13	19.40%
Primary School, High School	3	4.48%
Primary School, High School, Higher Education	1	1.49%
Primary School, High School, Technical School, Higher Education, Language School	1	1.49%
Primary School, High School, Technical School, Higher Education, Language School, Other	1	1.49%
Other	3	4.48%
Higher Education	6	8.96%
High school	7	10.45%
High School, Language School	1	1.49%
High School, Technical School	3	4.48%
Language School, Other	1	1.49%
Technical School	23	34.33%
Technical School, Other	2	2.99%
Technical School, Higher Education, Other	1	1.49%
Technical School, Language School	1	1.49%
Total	67	100%

As a last basic data to introduce the respondents to the survey, we measured the type of municipality where their educational institutions are located. The largest proportion (64%) work in a city with county status. A further 30% teach in smaller cities. 3% carry out their educational activity in a village. The remaining 3% indicated more than one type of municipality.

Results

In the first part of the questionnaire survey, we wanted to find out how well-informed teachers of foreign languages and Hungarian language are about AI-based applications, how open they are to using them in their teaching activities and to participating in training on them.

Overall, a quarter of the sample surveyed are not familiar with any AI applications at all, but just over half of the teachers do use them, with

0.03% using them in all lessons, but the majority using them with moderate regularity.

If we look at the results by generation in terms of the awareness and knowledge of applications among humanities teachers, the most striking finding from the sample is that there is no difference at all in the habits of Generation Z compared to Generations X and Y.

20% of Baby Boomers are familiar with more than one app, and almost half of the educators are also familiar with ChatGPT. The 16.4% of Baby Boomers who are the least likely to use it are the 16.4% who are the most averse to it, 27.3% who are users to some extent, but a larger proportion 64% would like to use it in their work, with the remaining proportion open to it. Only 10% of the Baby Boomers generation refuse to take part in training to help them use AI-based applications effectively, while the rest would like to join.

More than half of those completing the survey is Generation X. The smallest proportion, 16%, are not familiar with apps, but the largest, 47%, have listed more than one. In their case, 10% refuse to use it, 42% are undecided, but less than 1% are open to training, only 1 would not like it, 8% are undecided, but more than 90% would like to attend training on how to use AI effectively. 29% of Gen Y are not familiar with AI applications, while 21% are more. No one refused to use them, but 57% answered unequivocally yes and less than half were undecided. Of those open to education, under 10% were not definitely in favor. Generation Z represents the smallest sample in the survey, 25% do not use it, but the remaining 75% are familiar with or use several types of apps. However, 50%-50% are equally open or undecided about using AI-based software and participating in training to support it.

Looking at the same questions, there is a significant difference in the subjects taught, with most teachers in the sample who teach only foreign languages, i.e. 83%, being users, while only 38% of teachers who teach only Hungarian language and literature are open to training, but all of them.

Extending the study to school types, a third of primary school teachers do not use any AI application and are clearly open to training, while nearly 40% are undecided. Secondary school teachers show a more negative statistic, with 50% who are aware but none using an AI-based application and only 25 who are clearly positive about attending training, and the remaining 75% undecided. More familiar among technician trainers, 82% of respondents, however, do not typically use them, with only approximately 56%. These trainers are much more open to training, with 56% unequivocal, 41% undecided and 4% who reject it.

Most respondents to the survey, 54%, are familiar with the ChatGPT application, more specifically, in response to our questionnaire asking

respondents to list which AI-based application they are familiar with, one or more options were described by respondents, but this does not mean that they use it. In response to the free-word question, however, a quarter of respondents described not knowing or using this type of application at all or could not think of any.

Most of the AI-based applications mentioned, grouped by function, were software for language learning and translation software. These applications were mentioned by a low proportion, about one fifth of respondents. If we look at this question only for language teachers, the result increases by only 4-4%. Similarly, the proportion of teachers who are more broadly informed and who identified three or more AI-based applications is around 20%.

When looking at the responses of teachers of Hungarian, the results show a much lower level of awareness in this area. However, the applications listed by the more informed educators, a fifth of those surveyed, are diverse, with mentioning a total of 32 types of AI-based applications. In addition to player applications that can be used in class and help with learning, Quillbot and Grammarly were mentioned, which can be used for word processing and drafting. Tengr.ai, a graphics creator for editing images, and Pictory were also listed. Among the applications mentioned are several very effective tools that are useful in the humanities and in the education of students. Turboscribe turns video into text, SpeechTexter turns speech into written text.

Using Writetone, it is possible to write texts in different styles. Using the Masterdebate Communication app, it is possible to practice debating and argumentation.

In the questionnaire, the teachers were asked about the impact they think AI has had on the subjects they teach and on the use of language. The answers to both questions were very similar. The Likert 5-point scale was employed to evaluate the opinions. The extremes were very rarely used by respondents, i.e. in the question on how you feel the overall impact of AI on the subject you teach is, 1.49% of respondents gave a bad score and 4.48% gave a completely good score. The other question, how do you feel the overall impact of AI on students' language use, was answered with the same number of extremes. It is interesting to note that only Baby Boomers mentioned a negative impact, while Generation X and Y respondents gave a positive response. Interestingly, Generation Z respondents did not give either extreme answer. The number of respondents ticking 3 (medium value) was well above 50%. For the question on the impact of AI about own subject, 59.7% of respondents gave a score of 3. For the other question on the impact on the use of language, 61.19% of respondents gave a score of 3. For both

questions, more than half of the respondents chose a neutral 3. This shows that teachers perceive both positive and negative benefits of AI.

The surveyed teachers were asked about the positive and negative impact of AI on the subjects they teach. By giving a free-word response, they were given the opportunity to express their opinions accurately. In terms of positives (Fig. 1), a quarter of instructors had no experience, they do not know or do not see any positive impact. The majority, however, gave a wide range of responses on the benefits and aspects that they found useful. The responses show that most teachers see the support provided by AI as positive in many respects, with many more believing that it helps them to complete tasks, is effective and gives them new ideas, a sense of achievement and, in addition, it is enjoyable. AI also has a positive impact on students, it motivates them, makes the task interesting for them and improves their creativity.



Figure 1. Positive impact of AI.

The educators see as the negative effects (Fig. 2) is that students will be less motivated to think because of the rise of AI, and that they will not develop many skills such as creativity, critical thinking, and concentration. It will have a negative impact on their ability to understand texts and formulate texts, their perseverance will be reduced, and they will become self-disconnected. Homework and essays will not be their independent product and will therefore lose their relevance. Several commentators have also pointed out that assignments produced by AI-based applications should be treated with appropriate caution, as in many cases they may be wrong, inaccurate, or

misleading. Only one teacher commented that he did not see any negative impact. It is interesting that creativity was mentioned in both senses, with several respondents saying that it develops creativity and several respondents stating it as negative because the student is not thinking.

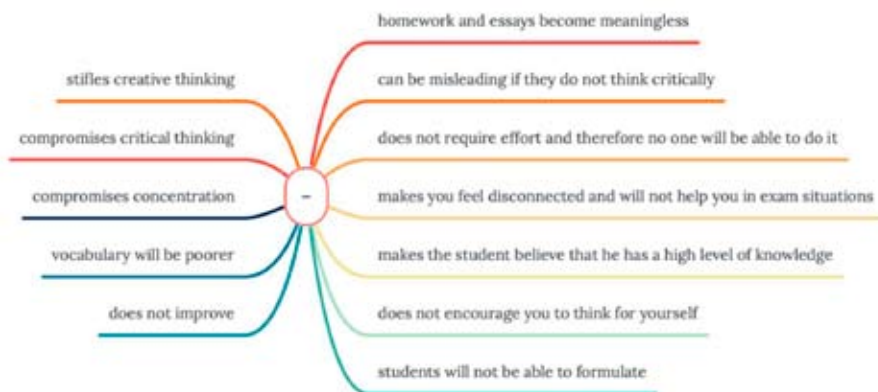


Figure 2. Negative effects of AI.

Respondents were also asked whether they recognize when their students use artificial intelligence to solve problems. If we look at the total population, 58.21% of respondents indicated that they do recognize when students use AI to solve problems. 13.43% indicated that they do not recognize it. While the remaining 28.36% indicated that they had not thought about it so far.

When looking at the responses by generation, it can be observed that no matter which generation the teacher or instructor belongs to, at least half of the respondents recognize the use of AI. The highest value was seen for Generation Y teachers, with a total of 64.29%. When we look at the responses of not recognizing it, we find that none of the Generation Z teachers chose this option. The Baby Boomers generation (18.18%) was the most likely to say they would not recognize it, while for generations X and Y this figure is below 15%. Those respondents who chose the option that they had not yet recognized the use of AI, Generation Z teachers stood out with the highest value of 50%. The responses of the other 3 generations surveyed ranged from 20% to 30%.

The next open question was to briefly summarize what they noticed when students used AI in their tasks if they answered yes. A total of 41 respondents provided an analyzable answer to this question. The responses received indicated that teachers were confident that they recognized the use of AI. More than 60% of the responses received noted the use of vocabulary and

grammatical structures that differed from the learners' level of knowledge. In over 55% of responses, the idea is that teachers know their own students' abilities, and if their answers are significantly different, they assume they are using some kind of artificial intelligence application. As one respondent put it literally, "knowing my students, I can see that they are not expressing their own thoughts".

Very common in the answers given is a dislike that it is too good to be true. More than 20% of the responses list things that help to recognize the use of AI. These typically related to the use of clichéd text elements and verbosity. Overall, from the responses received, it can be concluded that teachers, knowing their own students and the content of the course material, are very likely to recognize the use of AI. However, there is some uncertainty and concern in the responses received. Quoted verbatim from the response received: "I am not sure I always recognize it, but there have been examples of misuse of words and concepts. For me, the most memorable was when an adjective was expressed in the first-person plural of AI (e.g. 'We must be used consciously..., we are dangerous'). This was frightening."

At the end of the questionnaire, we asked the teachers who responded about two of the great dilemmas of our time. We asked them what they think will have a greater impact on them personally and on society in the coming years: artificial intelligence or climate change (Fig. 3). When asked which of the two would make a bigger difference to their lives in the next 5-10 years, 66% of respondents said artificial intelligence.

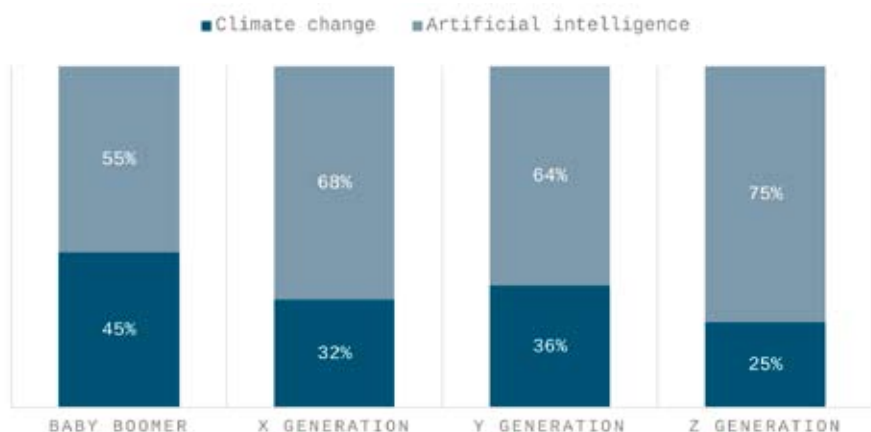


Figure 3. Opinion on which of the two big challenges of the future will make the biggest difference to the respondent's life in the next 5-10 years.

The question was also analyzed in terms of respondents' generational affiliation. It is clear from the responses that the youngest generation (Gen Z) surveyed believe that AI is the most important factor influencing their lives. In contrast, the oldest generation surveyed (Baby Boomers) consider climate change to be the most important factor influencing their lives.

As in the previous question, artificial intelligence was identified as the phenomenon with the greatest socio-economic impact (Fig. 4). In response to this question, 64% of respondents believed that AI would have a greater impact on society than climate change in the future. The difference in opinions between the generations was examined. 75% of Generation Z respondents think the impact of AI is the most significant. Interestingly, in this aspect, Generation X rated the importance of climate change the highest.

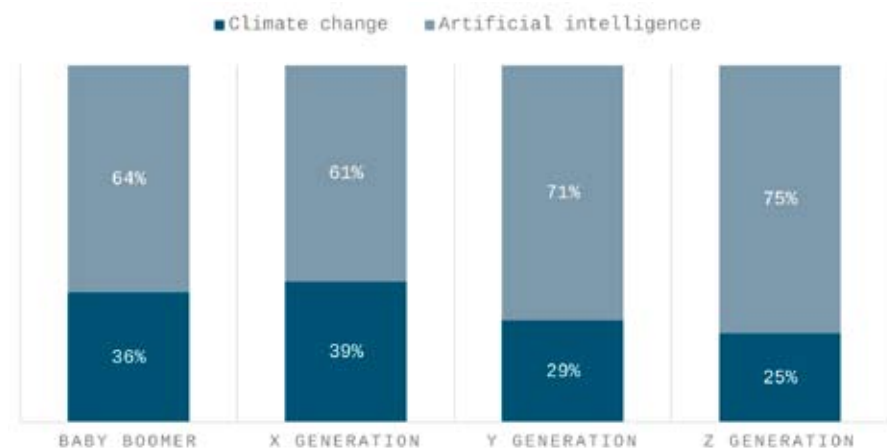


Figure 4. An opinion on which of the two major challenges of the near future will have a greater impact on socio-economic integration in the next 5-10 years.

Discussion and conclusion

In this study we focused on the opinions of Hungarian and foreign language teachers in Hungary and examined the impact and perception of the increasingly popular artificial intelligence. The questionnaire survey revealed that a quarter of the sample surveyed did not know of any AI-related applications (or at least could not list any, did not know that they were based on AI), while half of the respondents use AI. This usage rate is slightly lower than the results measured by Forbes Advisor (Hamilton 2023) among 500

teacher respondents in the US, but this measurement was not limited to language teacher responses as our survey was. 60% of teachers surveyed in the US use artificial intelligence. It is perhaps interesting to note that 55% of respondents in that survey felt that AI had a positive impact on the teaching process, whereas in our survey this was more in the neutral range.

The study supported the hypothesis that foreign language teachers use such applications more than Hungarian language and literature teachers. More than 80% of the respondents would be open to training on the use of AI in teaching, which we believe strongly indicates the openness and willingness of teachers to develop. Also striking in this area is the fact that the willingness to participate in training is above 50% in all generations surveyed.

Both positive and negative effects of AI were indicated by respondents. It is interesting to note that, despite the negative aspect being mostly identified as a lack of thinking, several positive aspects were presented by teachers.

Intergenerational differences were also examined for each question, but the generations showed only minor differences in the questions examined.

The research concluded by examining the relationship between climate change and artificial intelligence as two major issues of our time. Respondents were clearly of the opinion, regardless of generation, that AI will have a greater impact on their lives and on socio-economic society than climate change in the coming period.

References

- Arató, Balázs – Balázs, Géza 2023. Hol a tudás, amit elvesztettünk az információban?: a digitális pedagógia diszkurzív-intuitív megközelítése antropológiai és jogi nézőpontból. *Magyar Nyelvőr*, 147. 2023: 755–76.
<https://doi.org/10.38143/Nyr.2023.5.755>
- Hamilton, Ilana 2023. Artificial Intelligence in Education: teachers' Opinions on AI in The Classroom. *Forbes Advis*. <https://www.forbes.com/advisor/education/it-and-tech/artificial-intelligence-in-school/> Downloaded: 12.05.2024.
- Hong, Wilson Cheong Hin 2023. The impact of ChatGPT on foreign language teaching and learning: Opportunities in education and research. *Journal of Educational Technology and Innovation*, Vol. 5(1).
- Jurafsky, Daniel – Martin, James H. *Speech, and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. University of Colorado, Boulder.
- Kuhail, Mohammad Amin – Alturki, Nazik – Alramlawi, Salwa – Alhejori, Kholood 2023. Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, Vol. 28(1) pp. 973–1018.
- Loewen, Shawn – Isbell, Dan – Sporn, Zachary 2020. The effectiveness of app-based language instruction for developing receptive linguistic knowledge and oral communicative ability. *Foreign Language Annals*, Vol. 53(2) pp. 209–33.
- Mező, Ferenc 2019. Interdiszciplináris kapcsolódási lehetőségek a mesterséges intelligenciára irányuló cél-, eszköz- és hatásorientált kutatáshoz (Interdisciplinary linkages for goal-, tool- and impact-oriented research on artificial intelligence). *Mesterséges intelligencia – interdiszciplináris folyóirat*, Vol. 1(1). pp. 9–29.
- Mező, Ferenc 2023. *Code Poetry – avagy: Amikor az irodalom csókot dob az informatikának, de a mesterséges intelligencia elkapja azt a tehetséggondozás örömeire (Code poetry, or: when literature throws a kiss to IT, but artificial intelligence catches it for the joy of talent development)*. *Mesterséges Intelligencia*. Vol. 5(1) pp. 9–19. <https://doi.org/10.35406/MI.2023.1.9>
- Michot, Janick 2023. *Base System for a Language Learning Chatbot*. https://www.zhaw.ch/storage/engineering/institutezentrum/cai/studentische_arbeiten/Herbst_2022-NLP/MSE_VT1_ciel_ChaLL_Janick_Michot.pdf Downloaded: 02.05.2024.
- Palkó, Gábor 2022. Mesterséges intelligencia, digitális bölcsészet, kulturális örökség trendek és eredmények (Artificial intelligence, digital humanities, cultural heritage: trends and results). *Valós térben az online térért Workshop 31. országos konferencia. Konferenciakötet*. pp. 228–35.
<https://doi.org/10.31915/NWS.2022.30>
- Porkoláb, Ádám – Fekete, Tamás 2023. A mesterséges intelligencia alkalmazása a nyelvtanulásban (The use of artificial intelligence in language learning). *Iskolakultúra*, Vol. 33(8) pp. 67–80. <https://doi.org/10.14232/iskkult.2023.8.67>
- Rajki, Zoltán 2023. A mesterséges intelligencián alapuló alkalmazások a bölcsészet-, társadalomtudomány és az oktatás területén (Applications of artificial intelligence

- in the humanities, social sciences, and education). *Humán Innovációs Szemle*, Vol 14(2) pp. 4–21. <https://doi.org/10.61177/HISZ.2023.14.2.1>
- Shortt, Michell – Tilak, Shantanu – Kuznetcova, Irina – Martens, Bethany – Akinkuolie, Babatunde 2023. Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, Vol. 36(3) pp. 517–54.
- Szabóné Balogh, Ágota 2023. Mesterséges intelligencia az oktatásban (Artificial intelligence in education). *Mesterséges Intelligencia – interdiszciplináris folyóirat*. Vol 5(2) pp. 51–61. <https://doi.org/10.35406/MI.2023.2.51>
- Tiara, Anisa Dwi – M. Arinal, Rahman – Cipitro, Handrianto 2021. The student's perception about use of Duolingo application for improving English vocabulary. *International Journal of Education, Information Technology, and Others*, Vol 4(4) pp. 690–701.
- Tolner, Nikoletta – Pogátsnik, Monika – Módné Takács, Judit 2023. A mesterséges intelligencia szerepe az online vizsgáztatásban (The role of artificial intelligence in online examinations). *Iskolakultúra*, Vol. 33(10) pp. 39–55.

Zoltán Trenka

Assistant Professor, PhD Student

Óbuda University Alba Regia Technical Faculty

Doctoral School on Safety and Security Sciences

E-mail: trenka.zoltan@uni-obuda.hu

<https://orcid.org/0009-0003-6841-6037>

Viktória Varsányiné Szeredi

PhD Student

Óbuda University, Doctoral School on Safety and

Security Sciences

E-mail: vszv@icloud.com

<https://orcid.org/0009-0003-9398-3009>

Monika Pogátsnik

Associate Professor

Óbuda University Alba Regia Technical Faculty

E-mail: pogatsnik.monika@uni-obuda.hu

<https://orcid.org/0000-0002-2698-7291>

**LAWS CLEARLY: LARGE LANGUAGE MODELS AND PLAIN LANGUAGE
TRANSFORMATION**

Abstract

One crucial aspect of access to justice and access to legal information is the comprehensibility of legal text. The complexity and specialized terminology of legal language often prevents citizens from understanding legal texts and representing themselves effectively in legal proceedings. The new developments of Machine Learning, such as solutions based on Large Language Models, could represent a significant advancement in access to legal information, as they can transform complex legal texts into more straightforward, more understandable forms for laypeople. This paper attempts to exploit the capabilities of OpenAI's GPT-4 model to produce automatic Plain Language transcriptions of legal texts. The experiment concerns four specific linguistic features, and the results are analyzed manually from both a legal and a linguistic point of view.

Keywords: Access to legal information, Comprehensibility, Plain Language, Large Language Models, GPT-4

1. Introduction

The research topic of access to justice has been a significant area in legal studies for more than half a century, particularly in Socio-Legal studies. Over the past decades, two main research directions have emerged (Gomes 2019: 360). The first focuses on the input side, which examines the opportunities for access to various legal institutions and tools. The second concentrates on the output side, where the fairness and equality of judicial decisions are analyzed. This study emphasizes the input side, particularly regarding access to legal information and its comprehensibility.

Access to legal information encompasses multiple dimensions. It includes accessing primary legal sources such as laws, regulations, and court judgments. Thanks to digitalization, access to this kind of information in European continental legal systems is primarily free and online. Thus, it is no longer a significant issue. However, the ability to process legal information has become increasingly important. Keeping up to date with legal changes and understanding the structures of legal documents require expertise, which even traditional legal technologies have tried to support in the past years, allowing

citizens without legal expertise easier access to relevant legal information (Vági et al. 2024).

Another crucial aspect of access to legal information is the comprehensibility of legal text (Arató 2022; Arató–Balázs 2022). The complexity and specialized terminology of legal language often prevents citizens from understanding legal texts and representing themselves effectively in legal proceedings. Legal language differs from everyday language usage in that it uses specialized terminology, and both reduces and recontextualizes the facts of a legal case to highlight the 'legally relevant facts' that are crucial for resolving the case and representing it in court (Branting et al. 2020; Westermann et al. 2023).

Improving the comprehensibility of legal texts for laypeople is among the goals of the Plain Language Movement. This movement began in the United States in the mid-20th century and seeks to simplify legal language to democratize access to justice (Dobos 2015; Vinnai 2018; Tóth 2023; Balázs 2023). Transparency and predictability of the law is a fundamental requirement of the rule of law. Unnecessary linguistic complexity and vagueness obscure the meaning of a text. Proponents of Plain Language argue that reducing such unnecessary linguistic complexity can significantly improve the overall quality of the text. To achieve this goal, they propose and specify language changes that they believe will improve the overall comprehensibility of the text.

Therefore, this study investigates how solutions based on large language models (LLMs), such as Chat GPT, can improve the comprehensibility of legal texts for laypeople, thereby increasing access to justice. The application of LLMs could represent a significant advancement in access to legal information, as they can transform complex legal texts into more straightforward, more understandable forms for everyone. The structure of the study is as follows. Section 2 provides an overview of related works on improving the comprehensibility of legal texts with rule-based, machine learning, and LLM approaches. Section 3 introduces the methodologies and the setup of the experiment used to investigate the capabilities of LLMs in transforming complex legal texts into Plain Language. Section 4 shows the results. Section 5 concludes.

This paper thus discusses a new dimension of legal research, focusing on the relationship between technology and justice, where the most modern Machine Learning tools attempt to address the question of access to justice.

2. Related works

Improving the comprehensibility of legal texts has long been a focus within both classical legal studies and the interdisciplinary field at the intersection of law and technology. Efforts to enhance the accessibility of legal language can be categorized based on the methodologies employed: rule-based approaches, traditional Machine Learning (ML) techniques, and, more recently, approaches that leverage the capabilities of Large Language Models (LLMs). Consequently, this literature review is organized according to these categories to provide a comprehensive overview of the existing solutions.

Subsection 2.1. delves into rule-based solutions. Subsection 2.2. explores traditional ML tools applied to legal text comprehension. These methods utilize statistical models and pattern recognition to enhance the readability and accessibility of legal documents. Subsection 2.3. examines the capabilities of LLMs, a cutting-edge development in AI. These models, trained on vast text corpora, exhibit a remarkable ability to understand and generate human language, making them powerful tools for simplifying and clarifying legal texts.

2.1. Rules-based approaches

The most common method is to give linguistic suggestions to support comprehensibility, so that a new, easily understandable version of the text can be produced. These guidelines are mostly addressed to linguists or legal experts and aim to provide a standardized framework to produce accessible transcriptions of legal/official texts. Similar initiatives already have a long history.

In the United States, for instance, on 13 October 2010, President Barack Obama signed into law the Plain Writing Act of 2010, confirmed by three subsequent presidential decrees (URL1). The law requires each government agency to use Plain Language (PL) in all documents it issues or revises, to train its employees to use PL, and to establish administrative procedures to comply with the law. It also requires that a dedicated section be created on the websites of the relevant agencies to provide a forum for addressing inquiries about the provision. Finally, it requires organizations to appoint a contact person to deal with the topic.

As a follow-up, in 2011 the US government issued a set of federal PL guidelines called the Federal Plain Language Guidelines (URL2). They set out in detail the standards they believe a plain language text should meet. The

document also makes a direct link between government communication and academic research by citing related linguistic research (e.g., Garner 2001; Murawski 1999: 73–5; Charrow & Erhardt & Charrow 2007; etc.).

The latest achievement in this area is the development and publication of an international standard for PL (URL3). The standard was published by the International Organization for Standardization in June 2023. The structure and principles of the document follow the definition on the International Plain Language Federation (PLF) website (URL4). The International Plain Language Federation is an umbrella organization that brings together international plain language associations and expert groups, such as the Centre for Plain Language, the Clarity-international, which has its own journal and conference series on the subject, or the Plain Language Association International.

The most common formulation associated with the PLF is; “A communication is in plain language if its wording, structure, and design are so clear that the intended readers:

- can easily find what they need,
- understand what they find,
- and use that information.”

The recently published ISO standard also contains linguistic suggestions in this spirit (e.g. converting Light Verb Constructions to active voice, avoiding archaic terms, etc.).

A similar initiative in the field of continental law is the German model. In the terminology used there, *Einfache Sprache* “style” comes closest to the expectations of the Anglo-Saxon countries regarding PL. It is in fact a “direct” way of expressing oneself, aimed at people who have average language skills (prototypical native speakers) but who want to understand the information addressed to them quickly and effectively. In this context, Pottmann (2020: 83), for example, highlights the following linguistic features as a follow-up;

1. simple and coherent sentence structure, without mental jumps,
2. simple sentences of no more than 10-11 words, or compound sentences of no more than 15 words, provided that a subordinate clause is included - preferably only one of these should be included in the sentence,
3. active voice,
4. when choosing the vocabulary, one should strive for similarity to live speech,

5. metaphors, idiomatic collocations, and abstract expressions should be replaced by concrete expressions with literal meaning,
6. foreign words and technical terms should be replaced by native or simpler words or, if this is not possible, an explanation of the technical terms should be provided (e.g., in the form of footnotes),
7. insist on the observance of spelling rules.

2.2. Machine learning approach

What we have described so far is a rule-based approach to the issue of intelligibility. They have the advantage of being easily interpretable by hand-made rules, as opposed to a ML model, where the relationship between input and output is often not obvious (often called as the “black box” phenomenon – Oh et al. 2019).

In literature, approaches that attempt to address the issue of comprehensibility using ML models are scattered. Several studies on language simplification are available from the last decades. Many of the early attempts have focused on specific aspects of the problem. For example, Carroll et. al. (1999) experimented with breaking complex sentences into several separate sentences, Devlin (1999) attempted to replace complex words with simpler lexical items, and Cohn–Lapata (2009) attempted to make sentences more concise while keeping their essential content unchanged. The JusticeBot project (Westermann et al., 2023) focused on analyzing layperson descriptions of an event and detecting possible legal issues that may be relevant to their factual situation. It is a legal decision support tool for landlord-tenant disputes.

Many modern approaches have attempted to simplify texts by translating them within a language (intralingual translation). For example, Zhang–Lapata (2017), Mallinson–Lapata (2019), Nishihara (et al. 2019), and Dong (et al. 2019) have applied encoder-decoder architectures to automatically reformulate texts into more understandable versions. Mallinson (et al. 2020) interpreted the problem as zero-shot multilingual translation and attempted to achieve satisfactory results for resource-poor languages.

What links the above is that all the studies were in fact attempts to produce Easy English from the original texts. In this case, the aim is to produce a more accessible version of the texts for disadvantaged groups, such as language learners or people with dyslexia or aphasia. The initial data set is therefore often a database developed from Simple Wikipedia. The articles here are produced with considerable restrictions, for instance, only the most

common 1,000 English words are recommended to use, which is alone a severe restriction. In the case of PL, the aim is more to make a text more accessible to the ordinary language user. While in the former case we must deal with a lack of linguistic competence, in the latter case the problem is the presence of unnecessary linguistic complexity (which can be particularly pronounced in the case of source texts falling within the legal domain).

Üveges (2022) examined the information materials produced by the Hungarian National Tax and Customs Administration's Working Group on Clarity¹. His analysis included an original version of the texts and a version adapted to be accessible. Thanks to the parallel data, it was possible to produce training data from the sentence-by-sentence versions of the texts. With this data, the author tested simple ML models (SVM and fastText) to classify sentences that were comprehensible and ones that needed to be paraphrased.

What has not been examined in the literature so far, however, is the performance of the LLM-s, the most advanced language technology tools of our time, in relation to PL. There is mainly hypothetical and first-impression research on how, e.g., LLM-generated translation of legalese into simple terms can help the comprehensibility of legal texts and the explainability of legal facts and terms (Chien–Kim 2024: 41).

2.3. Large Language Models

Generative Artificial Intelligence refers to ML algorithms that can learn from, for example, textual content, images, or audio material to be able to create entirely new content on their own. This contrasts with discriminative Artificial Intelligence, which is tasked with learning the decision boundaries that arise when classifying the examples it is given (Jiao et al. 2022: 212).

Generative Artificial Intelligence is named for its ability to generate a wide range of outputs, including text (Ibqal–Qureshi 2022), images (Singh–Raza 2021), sound (Ghose–Prevost 2023), computer code (Dehaerne et al. 2022) and more. LLMs represent a category of generative AI that focuses specifically on text generation.

LLMs that exist today, such as the GPT series under development by OpenAI, the company's latest model GPT-4 (OpenAI 2023) or the Gemini family of models built by Google (Google 2023), are the latest elements of an evolutionary arc. The first outstandingly successful element was the Generative Adversarial Network (GAN) neural network architecture

¹ Nemzeti Adó- és Vámhivatal Közérthetőségi Program.

(Goodfellow et al. 2020). This was in fact a kind of forerunner of today's state-of-the-art solutions (e.g., ChatGPT), although it followed a significantly different architecture than the latter.

Today's generative models, thanks to the robust training phase and the huge amount of data on which the pretraining process takes place, are more suitable than ever for simulating human language use and thus for generating natural language texts. It is no coincidence that with the advent of ChatGPT, technology has rapidly introduced masses of people to Artificial Intelligence, and more specifically to the generative language modeling branch of it.

3. Methodologies and experimental setup

In our experiment, we aimed to investigate some of the linguistic features of PL. The aim was to get an idea of the ability of the LLMs to transform the resulting text along these lines, and the extent to which its legal content changes during similar transformations. Given that PL standards are very extensive, we have limited ourselves here to a few selected linguistic features.

- **Avoiding interjected, backward clauses – preferring short clauses.** The role of the organization of clauses in the text has come to the fore in literature discussing the comprehensibility of legal texts (Charrow–Charrow 1979), as well as in psycholinguistic literature (Pléh–Lukács 2014). It is a general finding that short clauses are easier to interpret than extremely long ones, and that intervening clauses within a sentence (which are clauses of a complex sentence but do not have a syntactic relationship with the rest of the complex sentence) make comprehension difficult in most cases. Among the above, we examined the case of how it is possible to shorten a particularly long clause.
- **Light Verb Constructions (LVC-s).** LVC-s can be broadly defined as a noun-verb structure consisting of an adverbial, an object, or a zero noun and a verb with a usually underspecified meaning, a so-called light verb (Nagy et al. 2020; Vincze 2009). The core of the structure's conceptual meaning is carried by the noun, its grammatical meaning by the verb, and its stylistic value by the phrase altogether, which may differ from the stylistic value of both members. The most common objection to such structures is that the verb does not carry any meaning and therefore makes the text difficult to understand. In the case of English, an example could be *render assistance*, where

render practically has no meaning, and in the case of Hungarian, *megszervezésre kerül* (will be organized), where *kerül* has also no real meaning. In this context, we therefore investigated whether it is possible to automatically detect and replace such structures.

- **Too long sentences.** It is widely accepted that the longer a sentence, the more difficult it is to interpret. This may be partly due, for example, to the limited working memory available when processing language in the brain. In this case, the goal was to divide the original sentence into several parts so that the result remains semantically and normatively equivalent to the original.
- **Use of the word “*illetve*” (*and, or etc.*).** In Hungarian, the word *illetve* (*and/or*) can have a wide variety of functions. It can be used for:
 - a. as a substitution for *or*, both in inclusive (A or B or both) and exclusive (A or B) cases,
 - b. to express a conjunction (in the sense of *and*),
 - c. it can refer to the fact that both *and* and *or* must be used among the elements of the list, and finally,
 - d. it often takes on a clarifying role in the sentence (e.g., *Gondoltam, illetve gondolhattam volna. – I thought or could have thought*).

From all this, it is clear that the use of the conjunction “or” is far from trivial. It is no coincidence that the 61/2009. (XII. 14.) IRM decree on legislative drafting § 7 (3) also states that the conjunction “*illetve*” can only be used in the draft legislation if other more clear language formulations are inapplicable.

In each case, the Language Model was to transform the resulting text to consider a particular aspect. We then manually evaluated :

- whether the conversion met the requirements of the prompt, and
- in the case of the converted text, whether the normative content carried by the original text was changed or violated.

Of course, the four qualities listed above do not necessarily occur in isolation in a single passage or sentence. For this reason, there are cases where more than one of the phenomena under examination occurs in a single example sentence. In such cases, we have also tried to cover these in more detail.

The subject of our investigation was Act CXXII of 2013 on Transactions in Agricultural and Forestry Land (hereinafter: Land Transaction Act). This

law is characterized by the fact that it describes legally complex facts. It is also true that the text is extremely difficult to read, with numerous sentences spanning several paragraphs, lists of many words and a sentence structure that is difficult to understand at first sight. English translations of the relevant parts are provided in [brackets], as well as in the case of the used prompts.

A prompt is an instruction typically used to instruct Language Models trained on question-answer sequences (“optimized as a chatbot”). The most vital role of a prompt is to contextualize the Language Model (e.g., to assign roles to the style of the expected response) and, of course, to provide the most adequate description of the expected operation. Prompt engineering is the use of techniques (based on practical experience) to assess the capabilities and limitations of LLMs and to tailor their operation to a specific task (URL5).

For our purposes, we have tried to use the most concise prompts possible. We have already examined that the GPT-4 model has the basic information about PL (URL7). Therefore, we have now focused on testing the ability of the model to transform the given text fragments according to given criteria. The extent to which the normative content changes in the process was also an important factor.

The following minimal prompt was used for each task:

- (i) *Jogász vagy, akinek az a feladata, hogy jogszabályokat közérthetőre fogalmazzon. Fogalmazd át az alábbi szöveget a Plain Language elveinek megfelelően! Tekintettel arra, hogy egy jogi szövegről van szó, annak normatív, azaz jogi tartalma semmiképpen sem változhat meg az átfogalmazás során. A legfontosabb szempont, amelyre figyelemmel kell lenned, vagy, a következő kell, hogy legyen:*

[You are a lawyer whose job is to produce transcripts of laws that are easier to understand. Paraphrase the following text according to Plain Language principles. Given that this is a legal text, the normative, i.e., legal, content must not change during the rewording. The key point to bear in mind during your work is: ...]

The end of each prompt was then completed with a short, specific task description. This was responsible for defining the specific problem (from the above four) to be solved for the model.

4. Results

In our experiment, we first investigated the extent to which ChatGPT can simplify a sentence by breaking it up into several sentences, or by shortening

its constituent clauses. To illustrate this, we have chosen 45§ (1) from the Land Transaction Act, the full text of which is a single sentence. The passage is particularly interesting, since in (c) there is a multiple compound cumulative adjective structure. The original text was:

(ii) “45. § (1) Az erdőnek minősülő föld haszonbérbe adása esetén előhaszonbérleti jog illeti meg a volt haszonbérelő olyan földművest, illetve mezőgazdasági termelőszövetkezetet:

- a) aki helyben lakó szomszédnak, illetve amely helybeli illetőségű szomszédnak minősül,
- b) aki helyben lakónak minősül, illetve amely helybeli illetőségűnek minősül, vagy
- c) akinek az életvitelszerű lakáshasználata helye, illetve akinek vagy amelynek a mezőgazdasági üzemközpontja legalább 3 éve olyan településen van, **amelynek közigazgatási határa a haszonbérlet tárgyát képező föld fekvése szerinti település közigazgatási határától közúton vagy közforgalom elől el nem zárt magánúton legfeljebb 20 km távolságra van.**”

[“(1) As regards the leasing of forest land, the previous lessee (farmer or agricultural producer organization) – shall have the right of first refusal:

- a) if qualifies as a local neighbor or locally registered neighbor;
 - b) if qualifies as a local resident or locally registered entity, or
 - c) if has a place of habitual residence or has his or its center of agricultural operations for at least three years in a municipality from **whose administrative boundaries the land to which the leasehold pertains is located within a 20-kilometer radius via public road or publicly accessible private road.**”]
- (emphasis added by the authors)

It is noteworthy that the text also contains the conjunction *illetve* (which, by the way, appears about 116 times in the whole text of the Act).

In this case, the prompt indicated in (i) was completed as follows:

(iii) Az átfogalmazott szöveg nem tartalmazhat túl hosszú, tehát nehezen érthető mondatokat.

[The paraphrased text should not contain clauses that are too long, so difficult to understand.]

We have tried to give the minimum possible instructions. This was done to better simulate the autonomous operation of the language model and to

better illustrate a case where the average user, not familiar with prompt engineering, would create instructions for the model. In our opinion, this is the most realistic setup from the point of view of the average user. The result was as follows:

(iv) “45. § (1) Ha egy erdő területét kiadják haszonbérbe, a korábbi haszonbérelő, aki földműves vagy mezőgazdasági termelőszervezet, előhaszonbérleti jogot élvez. Ez a jog a következő esetekben áll fenn:

a) A korábbi haszonbérelő helyi lakosként számít szomszédnak, vagy a korábbi haszonbérelő olyan helyi mezőgazdasági termelőszervezet, amely szomszédnak minősül.

b) A korábbi haszonbérelő helyi lakos vagy a korábbi haszonbérelő helyi mezőgazdasági termelőszervezet.

c) A korábbi haszonbérelő vagy a mezőgazdasági termelőszervezet életvitelszerűen használt lakása vagy az üzemközpontja legalább három éve olyan településen található, amelynek közigazgatási határa legfeljebb 20 kilométerre van azon település közigazgatási határától, **ahol** a bérbe adott föld található.”

[“Article 45 (1) If a forest area is leased out, the former lessee, who is a farmer or a farmers’ organization, shall have the right of first lease. This right exists in the following cases:

a) the former tenant is a local resident who is a neighbor, or the former tenant is a local agricultural producer organization which is a neighbor.

b) The former tenant is a local resident, or the former tenant is a local agricultural producer organization.

c) the former tenant or the agricultural production organization has had its permanent residence or its centre of operations for at least three years in a municipality whose administrative boundary is situated at a distance of 20 kilometers or less from the administrative boundary of the municipality **where** the leased land is situated.”]

(**emphasis** added by the authors)

The previous single sentence was broken down into 5 separate units by the model. The breakdown was clearly the individual cases listed in the legal code in connection with the facts given. It is also clear that the model replaces many structures that originally contained one or more adverbs and adjectives with the noun head of the given constituent (cf. *erdőnek minősülő föld* [land classified as a forest] → *erdő* [forest]). The list of possible cases is

preceded by a separate introductory clause (*This right exists in the following cases:*). The last clause, which contained most of the cumulative adverbs and adjectives, has also changed significantly. The insertion of the pronoun *ahol* (*where*), for example, makes the structure much easier to understand thanks to the introduction of the independent clause.

When working on improving the comprehensibility of legal texts, we must not only consider how much more understandable the text has become but also thoroughly examine whether the normative content of the legal text has undergone significant changes. This point is particularly interesting in the last clause of the text mentioned above, where the most significant modification occurred after simplification. It is noteworthy that although the normative content regarding how far a non-local resident or agricultural organization can be from a residence or operational center remained the same, the precise measuring method was omitted in the simplified version. This omission can lead to interpretative problems in legal practice.

Furthermore, the original text clearly stipulates the conditional nature of the right, stating that specific criteria must be met, while the simplified version might suggest that this right automatically and inherently belongs to the former lessee. Indeed, only the second sentence hints at these conditions. Additionally, the original text clearly shows that meeting any listed conditions is sufficient. In contrast, the simplified version does not clarify whether the right applies if all conditions are met or if meeting just one condition is sufficient.

However, a positive aspect is that the simplified text clarifies that only farmers can qualify as neighbors or residents, and only agricultural producer organizations can be locally registered neighbors or entities. It significantly eases the understanding of the texts for laypersons seeking legal information, helping them grasp which conditions apply to their situation.

16§ (4) of the Act has been selected regarding LVCs. The English translation is of particular interest here because, while it is a good representation of the meaning of the text, it also serves as an example of how LVCs cannot be translated in many cases while retaining their stylistic value.

- (v) “16. § (4) *A részaránytulajdon megszerzése esetében az (1) bekezdésben foglalt földszerzési maximumot azzal az eltéréssel kell alkalmazni, hogy a szerző fél tulajdonában és haszonélvezetében álló földterület hektárban kifejezett térmértéke mennyiségének hússzorosa, valamint a már tulajdonában álló és **megszerzésre kerülő** részaránytulajdon összesített aranykorona értéke mennyiségének együttes összege legfeljebb 6000 lehet.*”

*[(4) As regards the acquisition of an ownership share, the land acquisition limit shall apply with the derogation that the gold crown value of twenty times the size of land areas expressed in hectares owned or used by the acquiring party under usufructuary right, and that of the ownership share such party already has or is **about to acquire** may not exceed 6,000 hectares combined.]*

(**emphasis** added by the authors)

The prompt previously defined in (i) has now been supplemented by the following:

- (vi) *A szövegben minden funkciógét (Light Verb Construction) helyettesítened kell egy egyszerű cselekvő igével, pl.: 'intézkedést foganatosít' helyett írd azt, hogy 'intézkedik'.*

[You must replace every function verb (Light Verb Construction) in the text with a simple action verb, e.g., instead of 'intézkedést foganatosít', write 'intézkedik'.]

Thus, the following result was obtained:

- (vii) *Amikor valaki részaránytulajdont szerez, akkor a földszerzési korlátozásokat kicsit másképp kell értelmezni. Ez azt jelenti, hogy a tulajdonos birtokában lévő és használatában álló földterület nagyságát hektárban megszorozzuk hússzal. Emellett hozzáadjuk a tulajdonos már meglévő és most szerzett részaránytulajdonainak összes aranykorona értékét. Ezen két szám összege nem haladhatja meg a 6000-et.*

[When one acquires fractional ownership, the land acquisition restrictions should be interpreted slightly differently. This means that the amount of land in hectares owned and used by the owner is multiplied by twenty. In addition, we add the total value of the owner's existing and newly acquired fractional holdings of gold crowns. The sum of these two figures should not exceed 6000.]

The result is interesting from the point of view that the text as a whole described a computational method, which the model returned in separate sentences, broken down into steps. In this way, the LVC did indeed disappear, but it is not clear whether this was due to following a specific instruction or whether it was just a consequence of the general transformation.

² Takes action.

To clarify this, the model is given a separate prompt with the following instruction:

- (viii) *Hogyan helyettesítenéd a “megszerzésre kerül” funkciógés szerkezetet cselekvő szerkezettel?*
[How would you replace the “megszerzésre kerül” Light Verb Construction with active voice?]

Since the new prompt was submitted in the same conversation, the model not only gave the proposed substitution, but also “corrected” it in the previous text.

- (ix) *A “megszerzésre kerül” funkciógés szerkezetet helyettesítheted az “elsajátít” cselekvő igével. Ezáltal a mondat a következőképpen módosulhat:*
“Ez azt jelenti, hogy a tulajdonos birtokában lévő és használatában álló földterület nagyságát hektárban megszorozzuk hússzal. Emellett hozzáadjuk a tulajdonos már meglévő és most elsajátított részaránytulajdonainak összes aranykorona értékét. Ezen két szám összege nem haladhatja meg a 6000-et.”
[You can replace the function verb structure “megszerzésre kerül” with the action verb “elsajátít”. This way the sentence can be modified as follows:
“This means that the amount of land in hectares owned and used by the owner is multiplied by twenty. In addition, we add the total value of the owner’s existing and newly acquired share-ownership in the land. The sum of these two figures shall not exceed 6000.”]

This highlights the fact that the model actually translates all instructions into English in the background, and in the process some meanings may be confused or changed. Indeed, “elsajátít” (acquire – learn or develop a skill, habit, or quality) and “megszerez” (acquire – obtain an asset or object) can be translated into a common English verb, but in the concrete text the two take on vastly different meanings. In addition, the preceding sentence in the simplified text also changes the terms from ‘property and beneficiary ownership’ in the original text to ‘possession and use’ in Hungarian. These terms may carry different meanings in Hungarian legal terminology.

So, on the one hand, we can conclude that the original conversion was not intended to reformulate the LVC as prompted, and on the other hand, we can see the dangers of the automatic translation, especially for the legal domain.

To test the resolution of eviscerally long sentences, we have chosen Act 45§ (2), which also consists of a single sentence.

- (x) 45§ (2) *Ha az erdőnek minősülő föld közös tulajdonban áll, a tulajdonostárs tulajdoni hányada egészének vagy egy részének megfelelő terület harmadik személy javára történő haszonbérbe adása esetén az (1) bekezdésben meghatározott előhaszonbérletre jogosultakat követően előhaszonbérleti jog illeti meg a földműves tulajdonostársat.*

[45§ (2) If the land which constitutes a forest is in common ownership, in the event of a lease of all or part of the area corresponding to the ownership share of the co-owner to a third party, the co-owner of the land shall have a right of first refusal after the right of first refusal as defined in paragraph (1).]

In this case, the simple instruction added to prompt (i) was:

- (xi) *A következő mondatot bontsd fel több, egyszerűbb mondatra úgy, hogy a szöveg értelme változatlan maradjon!*
[Break the following sentence into several, simpler sentences, while keeping the meaning of the text unchanged.]

The result is illustrated in (xii):

- (xii) *Amikor egy erdőnek minősülő föld közös tulajdonban van, és egyik tulajdonostárs a tulajdoni hányadának egészét vagy egy részét harmadik félnek adja haszonbérbe, akkor először az (1) bekezdésben meghatározott előhaszonbérletre jogosultak jönnek. Ezután az előhaszonbérleti jog a földműves tulajdonostársat illeti meg.*
[When land which is a forest is jointly owned and one of the owners leases all or part of his share of the ownership to a third party, the first to be entitled to the advance lease as defined in paragraph 1 shall be. Thereafter, the right of prior tenancy shall be vested in the co-owner who is a farmer.]

The interesting thing about the situation is that the sentence describes a single conditional structure, its complexity is given by the multitude of expressions specifying the meaning.

For “illetve” (which can mean either ‘and’ and ‘or’ in Hungarian, depending on the context), we have chosen paragraph 35§ (5) of the Act, which contains the term twice:

- (xiii) 35§ (5) *Ha a mezőgazdasági igazgatási szerv a tulajdonszerzést jóváhagyja, a döntése meghozatalával egyidejűleg a 30. § (1) és (2) bekezdésében meghatározottak szerint az árverési jegyzőkönyvet*

záradékolja, majd megküldi a végrehajtónak, **illetve** a felszámolónak. A mezőgazdasági igazgatási szerv az árverési jegyzőkönyv szerinti teljes vételárat átutalja a végrehajtónak, **illetve** a felszámolónak.

*[If the agricultural administration body approves the acquisition, it shall at the same time endorse the auction report as provided for in Subsections (1) and (2) of Section 30, and shall send it to the bailiff **or** the liquidator, as appropriate. The agricultural administration body shall transfer the purchase price shown in the auction report in full to the bailiff **or** the liquidator.]*

(**emphasis** added by the authors)

In the English translation, the correct logical conjunction to insert in both cases would be “or”. In the last case, the model has the following prompt addition:

- (xiv) *A szövegben szereplő “illetve” kifejezés zavaró lehet, mivel “megengedő vagy”, “kizáró vagy”, illetve “és” értelmezése is lehet. Amennyiben indokolt, cseréld az előfordulásait a megfelelő kötőszóval.*

[The term “illetve” in the text can be confusing, as it can be interpreted as “inclusive disjunction”, “exclusive disjunction” or “conjunction”. If appropriate, replace occurrences with the appropriate conjunction.]

The language model then made the following changes to the text:

- (xv) *35§ (5) Amennyiben a mezőgazdasági igazgatási szerv jóváhagyja atulajdonszerzést, adöntésmeghozatalával egyidejűleg záradékolja az árverési jegyzőkönyvet a 30. § (1) és (2) bekezdéseiben meghatározott módon, majd elküldi a végrehajtónak és a felszámolónak. A mezőgazdasági igazgatási szerv az árverési jegyzőkönyv szerinti teljes vételárat átutalja a végrehajtónak és a felszámolónak.*

*[35§ (5) If the agricultural administrative body approves the acquisition of property, it shall, at the same time as it takes its decision, endorse the auction report in the manner specified in paragraphs (1) and (2) of Article 30 and send it to the executor **and** the liquidator. The agricultural administration body shall transfer the total purchase price according to the auction report to the bailiff **and** the liquidator.]*

(**emphasis** added by the authors)

In this case, the model followed the instruction more accurately in advance, i.e., the PL transcription given as a larger context did not “mask” the more specific task of replacing the indicated expression. However, according to the normative legal content, the solution of the GPT model needs to be corrected because, in the context of the original Hungarian text, ‘illetve’ carries the meaning of ‘or,’ indicating that it is either a liquidation or enforcement. Accordingly, either a liquidator or a bailiff is involved in the proceedings.

5. Conclusion

The GPT4 model performed somewhat mixed in the individual tasks. The most interesting phenomenon was the one we experienced when trying to remove LVCs. In doing so, the original meanings were mixed up during the automatic machine translation of the input. As a result, the transformed text became meaningless. Presumably, therefore, the problem was not the inference capabilities of the model, nevertheless, this is a potential source of error to be kept in mind when working with non-English language content.

The results of the model are promising in terms of shortening clauses. For example, the introduction of new conjunctions and relevant subordinate clauses makes the paraphrased versions much easier to understand.

It was common experience that, in the case of the prompts constructed to transform certain linguistic features, the model did not aim for specific transformations. Rather, the results obtained reflected a general reformulation from a PL perspective. This is not necessarily a problem, but a phenomenon that is certainly worth bearing in mind.

The linguistic features of the resulting transcripts were largely in line with the requested conditions, but the legal content was changed in almost all cases during the reformulation. This leads to the conclusion that the language model in its current state is in no way suitable for fully automatic paraphrasing. Rather, it is a productivity tool that can be used to support the work of the “translator” and reduce the time required.

On the one hand, the results provide an excellent illustration of the potential uses of LLMs to make work more efficient. On the other hand, they also make it clear that reformulation from a PL perspective, even with today’s most modern tools, is a task that does not lack human creativity and expertise, and where all the automation can only be directed towards supporting human experts’ work.

References

- Arató, Balázs 2022. Norm clarity in the light of Hungarian case law. *Magyar Nyelvőr*, 46 (special issue), 81–90. <https://doi.org/10.38143/Nyr.2022.5.81>
- Arató, Balázs – Balázs, Géza 2022. The linguistic norm and norm of legal language. *Magyar Nyelvőr*, 46 (special issue), 91–103. <https://doi.org/10.38143/Nyr.2022.5.91>
- Balázs, Géza 2023. Efforts to Make Legal Language More Intelligible 189–202. In: Prof. Dr. Imre Szabó (ed.): *Central European Legal Studies. Selected Essays on Current Legal Issues from a Comparative Legal Approach*. Vol. 2. Partium Press, Oradea–Nagyvárad. 2023.
- Branting, Karl – Balhana, Carlos – Pfeifer, Craig – Aberdeen, John – Brown, Bradford 2020. Judges Are from Mars, Pro Se Litigants Are from Venus: Predicting Decisions from Lay Text. In: *JURIX* (pp. 215–8).
- Dobos, Csilla 2015. Nyelven belüli fordítás és tisztességes jogi eljárás. In: Szabó, Miklós (szerk.): *A jog nyelvi dimenziója*. Bíbor Kiadó. Miskolc. 215–26.
- Carroll, John – Minnen, Guido – Pearce, Darren – Canning, Yvonne – Devlin, Siobah – Tait, John 1999. Simplifying text for language-impaired readers. *Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics*. 269–70.
- Charrow, Robert P. – Charrow, Veda R. 1979. Making Legal Language Understandable: A Psycholinguistic Study of Jury Instructions. *Columbia Law Review*, 79. 1306–74. <https://doi.org/10.2307/1121842>
- Charrow, Veda R. – Erhardt, Myra K. – Charrow, Robert P. 2007. *Clear & Effective Legal Writing*, 4th ed. Aspen Publishers. New York.
- Chien, Colleen V. – Kim, Miriam 2024. Generative AI and Legal Aid: Results from a Field Study and 100 Use Cases to Bridge the Access to Justice Gap. *Loyola of Los Angeles Law Review*, Forthcoming. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4733061
- Cohn, Trevor – Lapata, Mirella 2009. Sentence compression as tree transduction. *Journal of Artificial Intelligence Research* 34(1) 637–74. <https://doi.org/10.1613/jair.2655>
- Dehaerne, Enrique – Dey, Bappaditya – Halder, Sandip – De Gendt, Stefan – Meert, Wannes 2022. Code Generation Using Machine Learning: A Systematic Review. *IEEE Access*, vol. 10. 82434–55.
- Devlin, Siobhan 1999. *Simplifying Natural Language for Aphasic Readers*. Ph.D. thesis, University of Sunderland.
- Dobos, Csilla 2015. Nyelven belüli fordítás és tisztességes jogi eljárás. In: Szabó Miklós (szerk.): *A jog nyelvi dimenziója*. Bíbor Kiadó. Miskolc. 2015–226.
- Dong, Yue – Zichaom, Li – Mehdi, Rezagholizadeh – Cheung, Jackie Chi Kit 2019. EditNTS: An neural programmer-interpreter model for sentence simplification through explicit editing. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 3393–402.

- Garner, Bryan A. 2001. *Legal Writing in Plain English*. University of Chicago Press. Chicago. <https://doi.org/10.7208/chicago/9780226284200.001.0001>
- Ghose, Sanchita – Prevost, John J. 2023. FoleyGAN: Visually Guided Generative Adversarial Network-Based Synchronous Sound Generation in Silent Videos. *IEEE Transactions on Multimedia*, vol. 25. 4508–19.
- Gomes, Silvia 2019. Access to law and justice perceived by foreign and Roma prisoners. *Race and Justice*, 9(3) 359–79.
- Goodfellow, Ian – Jean Pouget-Abadie – Mehdi Mirza – Bing Xu – David Warde – Farley – Sherjil Ozair – Aaron Courville – Yoshua Bengio 2020. Generative adversarial networks. *Communications of the ACM*, Volume 63, Issue 11. 139–44.
- Google, Gemini Team 2023. Gemini: A Family of Highly Capable Multimodal Models. *arXiv*. <https://doi.org/10.48550/arXiv.2312.11805>
- Iqbal, Touseef – Qureshi, Shaima 2022. The survey: Text generation models in deep learning. *Journal of King Saud University - Computer and Information Sciences* Volume 34, Issue 6, Part A. 2515–28. <https://doi.org/10.1016/j.jksuci.2020.04.001>
- Jiao Sun – Q. Vera Liao – Michael Muller – Mayank Agarwal – Stephanie Houde – Kartik Talamadupula – Justin D. Weisz 2022. Investigating Explainability of Generative AI for Code through Scenario-based Design. In: *27th International Conference on Intelligent User Interfaces (IUI '22)*. 212–28. Association for Computing Machinery, New York, NY, USA.
- Mallinson, Jonathan – Sennrich, Rico – Lapata, Mirella 2020. Zero-Shot Cross lingual Sentence Simplification. *Conference on Empirical Methods in Natural Language Processing*.
- Mallinson, Jonathan – Lapata, Mirella 2019. *Controllable Sentence Simplification: Employing Syntactic and Lexical Constraints*. *arXiv preprint arXiv:1910.04387*
- Murawski, Thomas 1999. *Writing Readable Regulations*. Carolina Academic Press. Durham.
- Nagy, István – Rácz, Anita – Vincze, Veronika 2020. Detecting light verb constructions across languages. *Natural Language Engineering* 26(3) 319–48.
- Nishihara, Daiki – Kajiwar, Tomoyuki – Arase, Yuki 2019. Controllable text simplification with lexical constraint loss. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop, Florence, Italy*. 260–6.
- Oh, Seong Joon – Schiele, Bernt – Fritz, Mario 2019. Towards Reverse-Engineering Black-Box Neural Networks. In: *Lecture Notes in Computer Science*, Springer International Publishing. 121–44.
- OpenAI 2023. GPT-4 Technical Report. *arXiv*. <https://doi.org/10.48550/arXiv.2303.08774>
- Pléh, Csaba – Lukács, Ágnes 2014. *Pszicholingvisztika – magyar pszicholingvisztikai kézikönyv I*. Akadémiai Kiadó. Budapest. <https://doi.org/10.1556/9789630594998>

- Pottmann, Daniel 2020. Leichte Sprache and Einfache Sprache – German plain language and teaching DaF German as a foreign language. *Studia Linguistica*. 38. 81–94.
- Singh, Nripendra Kumar – Raza, Khalid 2021. Medical Image Generation Using Generative Adversarial Networks: A Review. In: Patgiri, Ripon – Biswas, Anupam – Roy, Pinki (eds.): *Health Informatics: A Computational Perspective in Healthcare. Studies in Computational Intelligence*, vol 932. Springer, Singapore. 77–96.
- Tóth Judit 2023. A közérthető közigazgatási eljárás és az átláthatóság – recenzió. *Magyar Jogi Nyelv*, 1. 26–8.
- Üveges, István 2022. Comprehensibility and Automation: Plain Language in the Era of Digitalization. *Taltech Journal of European Studies* 12(2) 64–86.
<https://doi.org/10.2478/bjes-2022-0012>
- Vági, Renátó – Üveges, István – Megyeri, Andrea – Fülöp, Anna – Vadász, János Pál – Nagy, Dániel – Csányi, Gergely Márk 2024. Increasing access to legal information with unsupervised solutions. *Hungarian Journal of Legal Studies*.
<https://akjournals.com/view/journals/2052/aop/article-10.1556-2052.2023.00473/article-10.1556-2052.2023.00473.xml>
- Vincze, Veronika 2009. *Előadást tart vs. előad: főnév + ige szerkezetek igei variánsai*. In: Sinkovics, Balázs (szerk.): *Lingdok 8. Nyelvész-doktoranduszok dolgozatai*. Szeged, Szegedi Tudományegyetem. 265–78.
- Vinnai, Edina 2018. »Megértette a figyelmeztetést?« A figyelmeztetések és tájékoztatások közlése a büntetőeljárásokban. In: Szabó, Miklós – Vinnai, Edina (szerk.): *A törvény szavai: Az OTKA-112172 kutatási zárókonferencia anyaga Miskolc, ME – MAB, 2018. május 25*. Bíbor Kiadó. Miskolc. 281–95.
- Westermann, Hannes – Meeüs, Sébastien – Godet, Mia – Troussel, Aurore – Tan, Jinzhe – Savelka, Jaromir – Benyekhlef, Karim 2023. Bridging the gap: Mapping layperson narratives to legal issues with language models. In: *Proceedings of the 6th Workshop on Automated Semantic Analysis of Information in Legal Text co-located with the 19th International Conference on Artificial Intelligence and Law (ICAIL 2023), Braga, Portugal* (Vol. 3441, pp. 37–48).
- Zhang, Xingxing – Lapata, Mirella 2017. Sentence simplification with deep reinforcement learning. In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 584–94.
- URL1. obamawhitehouse.archives.gov (2011): Executive Order 13563 – Improving Regulation and Regulatory Review. <https://obamawhitehouse.archives.gov/the-press-office/2011/01/18/executive-order-13563-improving-regulation-and-regulatory-review> (Access: 2024. 04. 11.)
- URL2. [plainlanguage.gov](https://www.plainlanguage.gov) 2011. Federal Plain Language Guidelines. <https://www.plainlanguage.gov/media/FederalPLGuidelines.pdf> (Access: 2024. 04. 11.)
- URL3. standards.iteh.ai 2023. ISO/FDIS 24495-1 – Plain Language. <https://cdn.standards.iteh.ai/samples/78907/35c071a9d3bd4608a23200b512460fba/ISO-FDIS-24495-1.pdf> (Access: 2024. 04. 11.)

- URL4. [iplfederation.org: Plain language definitions. https://www.iplfederation.org/plain-language/](https://www.iplfederation.org/plain-language/) (Access: 2024. 04. 11.)
- URL5. [simple.wikipedia.org](https://simple.wikipedia.org/wiki/Main_Page) 2023. Main Page. https://simple.wikipedia.org/wiki/Main_Page (Access: 2024. 04. 11.)
- URL6. [promptingguide.ai](https://www.promptingguide.ai) 2024. Prompt Engineering Guide. <https://www.promptingguide.ai/> (Access: 2024. 04. 11.)
- URL7. Üveges István 2024. Plain Language and Automation – Possibilities, Drawbacks, and the Reality of Large Language Models (Part II.). <https://constitutionaldiscourse.com/plain-language-and-automation-possibilities-drawbacks-and-the-reality-of-large-language-models-part-ii/> (Access: 2024. 04. 11.)

István Üveges

Developer, Researcher

MONTANA Knowledge Management Ltd.

Centre for Social Sciences

E-mail: uveges.istvan@montana.hu,

<https://orcid.org/0000-0001-5897-9379>

Renátó Vági

Legal Artificial Intelligence Expert

MONTANA Knowledge Management Ltd.

PhD student, Doctoral School of Law Eötvös

Loránd University, Budapest

E-mail: vagi.renato@montana.hu

<https://orcid.org/0000-0002-2019-5349>

JÁNOS TAMÁS PAPP

NAVIGATING THE DIGITAL CONTRACT MAZE: HOW AI CAN GUIDE USERS TO INFORMED DECISIONS

Abstract

This article delves into the complexities surrounding the Terms of Use agreements on social networking sites, illustrating how users often unknowingly agree to legally binding contracts without comprehending the full scope of their contents. It highlights the legal enforceability of these agreements, despite their complexity and the challenges they pose to user understanding and informed consent. The discussion advocates for greater transparency and accessibility in these documents, suggesting that advancements in artificial intelligence (AI) and machine readability could play a crucial role in demystifying legal jargon and enhancing user comprehension. By leveraging AI for clearer, more interactive explanations of terms, the article calls for a digital landscape where informed consent becomes a foundational principle, thereby fostering a more transparent, fair, and user-friendly online environment.

Keywords: AI, Terms of Use, contracts, machine readability, transparency

1. Introduction

In the contemporary digital landscape, the act of registering on social networking sites has become a routine yet crucial part of engaging with the online world. The internet and new forms of social media have radically changed the structure and ways of social communication (Szűts and Gerencsér 2020: 41; Balázs 2023). As users hurriedly click through to accept the terms of use, a significant number do so without a comprehensive understanding of the agreements they are entering into. Almost all online platforms operate under specific terms of use, which essentially serve as the legal backbone of these digital environments. These terms delineate the code of conduct expected from users, outlining permissible and prohibited actions within the platform, and clarify the legal relationship between the user and the service provider. Typically, the act of registering on these sites involves the user's tacit agreement to these conditions, executed through actions as simple as ticking a checkbox on a registration form or clicking an acceptance icon. More subtly, this agreement is often considered binding the moment a user begins to engage with the platform's services, thus initiating a formal legal relationship between the individual and the platform's operator. This

relationship imbues both parties with specific rights and obligations. The majority of users, however, are unaware that their interactions with these platforms amount to entering into a legally binding contract, virtually equivalent in force to traditional contracts signed with ink and paper. The act of a user clicking “I agree” to accept the terms of a contract – often referred to as a “clickwrap” agreement (Papp 2017: 189) – has consistently been upheld by courts as a form of enforceable contract. The legality and enforceability of clickwrap licenses have been examined and affirmed by judicial bodies, establishing a precedent that the mere act of clicking to agree constitutes a legitimate form of consent to the terms outlined within the agreement. This form of assent, although perhaps less direct than traditional methods of signing a contract, is nonetheless recognized as a valid expression of agreement to the contractual terms presented. This notion of blanket assent – whereby a user agrees to a set of terms and conditions en masse through a single action – still constitutes a form of agreement under the law. (Lemley 2006: 466). The terms of service agreements, often lengthy and complex, are a core component of this framework. The terms often include clauses that waive users’ rights while simultaneously bestowing enforceable rights upon the service providers, rights that can even be upheld in court. As such, there is a pressing need for these documents to be made more accessible and understandable to the average user. Ensuring clarity and fairness in these agreements is not just a matter of legal compliance; it is essential for maintaining trust and transparency (Park 2013: 905).

2. Understanding the complexity of the terms

Navigating the complex terrain of online platform Terms of Use presents a formidable challenge for the average user. One could convincingly argue that the obligation of reading and comprehending the terms of use falls squarely on the users. When users fail to thoroughly review these agreements, it is reasonable to hold them accountable for consenting to conditions that may not be entirely clear to them. While it appears that users are presented with options, in reality, their choice is often reduced to a binary decision: accept the terms as they are or forego the service altogether. This ‘take-it-or-leave-it’ scenario is further compounded by findings from research indicating that users, even when directly faced with the details, consent to having their messages reviewed by the service provider (Perry 2016: 54). This widespread acquiescence to the terms of service is not solely due to their inevitable acceptance as a condition of use. Another significant factor contributing

to this phenomenon is the complexity, opacity, and incomprehensibility of these terms. These documents, often lengthy and laden with legal jargon, are designed more with compliance and protection of the platform in mind rather than clarity or accessibility for users. The expectation for users to read, comprehend, and consent to extensive and complex legal documents before accessing a platform's services is both unrealistic and impractical, often leading to acceptance without true understanding. It is very difficult for users, especially those not in the habit of meticulously reviewing legal documents, to stay informed about changes and their potential impact. The result is a significant disconnect between the platforms and their users, with the latter frequently left in the dark about their rights, obligations, and the full scope of how their data is used. This erode the foundational principle of informed consent, crucial for maintaining ethical standards and adhering to legal statutes designed to protect consumers.

In a study focused on analyzing the Terms of Service for YouTube, Facebook, and Twitter, findings highlighted a distinct contrast. Unlike the content moderation guidelines, and privacy or data policies, which are tailored documents reflecting the unique attributes of each platform, the Terms of Service were found to be fundamentally boilerplate. This suggests that, contrary to the bespoke nature of other policy documents designed to cater to the specific operational and ethical nuances of each site, the Terms of Service adhere to a more standardized format, lacking customization across the platforms: "not meant to be read, still less to be understood." (Karanicolas 2021: 21) Another fundamental issue lies in how companies, including various platforms, have capitalized on the widespread lack of consumer literacy and the inevitable fatigue stemming from the relentless flood of lengthy and obscure agreements. This exploitation allows them to embed provisions that consumers would not naturally associate with the key terms of their transactions. Such practices undermine fundamental principles of consumer protection, as they sneak in terms that subvert the rights and expectations of the consumer under the guise of legal and necessary formalities (Karanicolas 2021: 25).

3. The importance of transparency

Legally mandated transparency could play a crucial role in addressing the issues surrounding the comprehension of terms of use on online platforms. By requiring service providers to present their terms in clear, straightforward language and to structure these documents in a user-friendly manner, users

would be better equipped to understand their rights and obligations (Arató 2022; Arató and Balázs 2022). This could include the use of plain language summaries, clear headings, and the highlighting of key terms related to data privacy, content ownership, and dispute resolution. Furthermore, transparency mandates could enforce the disclosure of critical information in an accessible format, ensuring that users are fully informed about how their data will be used, shared, and protected. Such legal requirements would not only facilitate informed consent but also foster trust between users and service providers, by demonstrating a commitment to fairness and user empowerment in the digital environment.

Benedikt Schmitz and Charlotte Pavillon categorize transparency into two distinct kinds: formal transparency and substantive transparency. Formal transparency refers to the accessibility and clarity of the contract's presentation, such as its layout and the comprehensibility of its language. On the other hand, substantive transparency delves deeper, focusing on whether the consumer, upon reading the contract, can grasp the essence and implications of its content. This type of transparency assesses the contract's understandability, ensuring that the consumer not only reads but also comprehends the terms, their rights, and obligations under the agreement, thereby achieving a genuine understanding of the contract's substance (Schmitz and Pavillon 2020: 191). Also, according to Loos, the mandate for transparency encompasses at least three distinct dimensions. First, the manner in which information is displayed is imperative. It must be laid out in a manner ensuring it cannot be overlooked by the consumer prior to, or at the moment of, contract agreement. This ensures that crucial details are directly in the consumer's line of sight, eliminating any possibility of missing out on key facts. Second, the clarity and understandability of the information are crucial. The wording of the information must be such that it facilitates easy comprehension by the consumer, presenting the information in straightforward, accessible language that allows the consumer to grasp the essence of what is being communicated without undue effort. Finally, the linguistic accessibility of the information is paramount. The information must be provided in a language or dialect that the consumer is proficient in. This principle guarantees that language barriers do not hinder the consumer's understanding of the information, ensuring that they can fully comprehend the details of what they are agreeing to. By adhering to these principles, the transparency requirement aims to safeguard consumers, ensuring that they are fully informed and able to make decisions based on a clear understanding of the terms and conditions they are entering into (Loss 2017: 54).

4. EU Laws on transparency

The Directive on Unfair Contract Terms, under Article 5, mandates that “in the case of contracts where all or certain terms offered to the consumer are in writing, these terms must always be drafted in plain, intelligible language. Where there is doubt about the meaning of a term, the interpretation most favourable to the consumer shall prevail” (Directive 93/13/EEC). This is not unprecedented in law (Arató 2023). In various instances, the Court of Justice of the European Union has clarified the concept of “plain intelligible language.” It is crucial to highlight that a key consideration for the Court is to ascertain whether the mandate for plain and intelligible language is upheld, even in scenarios where this requirement has not been explicitly transcribed into national legislation. This determination underscores the importance of ensuring that contractual terms are accessible and understandable to consumers, irrespective of the specificities of national legal frameworks (Case C-26/13). In another case, the Court emphasized the importance of consumers having the opportunity to thoroughly review all terms and their implications in the General Business Conditions (GBC) (Case C-472/10). Assessing the fairness of a term under Article 3 hinges on whether consumers can predict changes made by a seller or supplier to the GBC, especially regarding service fees, based on transparent and understandable criteria. In another case it also stated that is really important that consumers are able to foresee the charges that they might incur in their contracts (Case C-92/11). The transparency requirement for contractual terms, outlined in Article 4(2) and Article 5 of Directive 93/13, goes beyond just making them grammatically clear. This directive recognizes that consumers often find themselves at a disadvantage compared to sellers or suppliers, especially in terms of knowledge. Therefore, the directive demands that contractual terms be written in simple, clear language to ensure transparency. This requirement is meant to be interpreted broadly, acknowledging the consumer’s need for clear understanding to bridge the knowledge gap (Case C-125/18).

The directive demands that terms must be clear and simple enough for consumers to not only read and understand the language but also grasp how the terms work and the major financial impacts these could have on them (Joined Cases C-776/19 and C-782/19). Because it is a directive, there are many different interpretations in national laws, some Member States, like Germany, permit contracts to be executed in a foreign language if agreed upon in that language. Conversely, other Member States, such as France, stipulate that contracts made with individuals residing in their territory must be written in the official language(s) of that Member State (Loos 2017: 59).

Speaking of another major piece of EU legislation in this area, Article 14 of the Digital Services Act (DSA) focuses on ensuring the clarity and understandability of Terms of Use for providers of intermediary services. It mandates that these providers must clearly outline any restrictions related to the use of their services, including content moderation practices, algorithmic decisions, human review processes, and internal complaint handling procedures, in their terms and conditions. This information must be presented in clear, plain, and user-friendly language, and be easily accessible and available in a machine-readable format. Additionally, any significant changes to the terms and conditions must be communicated to the service recipients. For services primarily aimed at or used by minors, the terms must be explained in a way that is understandable to them. The DSA also requires that these terms and enforcement actions are applied diligently, objectively, and proportionately, taking into account the rights and interests of all parties, including fundamental freedoms and rights. For very large online platforms and search engines, the DSA goes a step further by requiring a concise summary of terms and conditions, including remedies and redress mechanisms, in clear language. Moreover, these platforms must publish their terms and conditions in the official languages of all Member States where they offer services, ensuring broader accessibility and understanding.

Also brought to life by the DSA, the Digital Services Terms and Conditions Database is an important step forward in making online terms and conditions more transparent and understandable for users. Hosted on a specialized platform, it offers a straightforward way for people to access and review the terms of service for various EU digital services, including updates and previous versions. The creation of this database tackles the common problem of complex and frequently changing online terms, which often confuse consumers about their rights and obligations. By providing these documents in a machine-readable format, the database not only makes it easier for users to access this information but also ensures that online platforms meet legal standards for clear and comprehensive contracts. The key feature of the database is its machine-readability, which simplifies the analysis and understanding of these often complex documents. This functionality could lead to the development of tools that help users better grasp the terms of use on online platforms, making the digital environment more user-friendly and transparent (URL1).

The introduction of these legal requirements and the revision of terms to improve transparency are significant steps toward demystifying the legal language that underpins the digital contracts encountered by users daily, but in

itself they are not enough. The easy availability of terms and conditions, along with their presentation in simpler language, are indeed essential and should be viewed as the absolute basic requirements for platforms. However, these aspects merely scratch the surface of what is needed to ensure transparency and user-friendliness in the digital space. Also, empirical research into the readability of legal text indicates that it remains inaccessible to a significant portion of the population, proving to be very difficult or even incomprehensible for many citizens, so the adoption of plain language, while beneficial, does not substantially change this situation (Curtotti and McCreath 2013: 5).

5. AI to the rescue?

While AI holds the potential for misuse on the internet, such as disseminating fake news or deepfake content (Gosztonyi and Lendvai 2023), the emergence of machine readability requirements and Artificial Intelligence (AI) is promising to bridge the gap between the convoluted legal texts and the average user's comprehension. Machine readability converts documents into formats effortlessly processed by computers, setting the stage for AI to analyze and understand this information. This fusion of technologies, when applied to terms of use and similar legal documents, can drastically enhance user understanding and interaction (Call and Wang and Weng 2023). AI has the potential to demystify these documents, fostering a more transparent, accessible, and user-friendly approach to understanding the rules, rights, and responsibilities associated with social media use. AI tools, utilizing natural language processing, can sift through these documents, highlight essential details, and present them in a straightforward, concise manner. AI systems can translate complex legal jargon into plain language, making the content more accessible to people without a legal background and can aid in analyzing reports, enhancing data utility and identifying discrepancies (Bauguess 2018). This not only helps users better understand the agreements they are entering into but also empowers them to make more informed decisions about their online behavior. AI-driven visualization tools can also play a crucial role in enhancing understanding. By converting text into infographics, timelines, or flowcharts, AI can help users grasp complex concepts at a glance. Visual representations of how data is collected, used, and shared, for instance, can make abstract privacy concerns more tangible and comprehensible.

AI's ability to tailor the presentation of terms based on a user's profile and previous platform interactions adds another layer of personalization. For example, content creators on a social media site might see terms related

to copyright and monetization emphasized, ensuring users are aware of conditions most relevant to their activities. Interactive AI tools represent another innovative approach. Chatbots and virtual assistants, powered by AI, can offer on-demand explanations of specific terms or sections. Users can ask questions in natural language and receive instant clarifications, making the process of navigating terms and conditions more engaging and less daunting. This interactivity not only aids comprehension but also helps users feel more confident about their rights and obligations.

Furthermore, AI can monitor changes to terms and conditions over time, alerting users to updates and summarizing the implications of these changes. Given the dynamic nature of social media platforms and privacy regulations, keeping users informed in real-time can significantly enhance transparency and trust. Leveraging this technology benefits service providers as well, since presenting terms and conditions correctly can preempt legal disputes stemming from their inadequate representation (URL2).

While AI lacks a legal background, rendering it incapable of delivering unerring interpretations of legal documents, this limitation is mitigated by several factors. First, technological advancements are continually enhancing the efficiency of legal text analysis. Second, the application of AI in analyzing other legal documents, such as judgments, is gaining traction (Eliot 2021: 3). Most crucially, legal mandates for transparency and clarity are effectively pre-adapting legal content for AI analysis, significantly diminishing the likelihood of errors. This proactive alignment between legal standards and AI capabilities ensures a progressively reliable interpretation of legal documents. Of course, simplicity in reading does not always translate to simplicity in comprehension. A study in this field suggests that poor readability, as measured by the usual formulas used, does not serve as a reliable measure for lack of understandability (Schmitz and Pavillon 2020: 199). A qualitative study in this matter concluded, that although a small fraction of participants raised privacy concerns, the overwhelming majority commended the convenience of quick-join clickwrap options for allowing them to skip over the notices. The issue isn't merely that users find privacy and Terms of Service policies dull or irrelevant; rather, it's that individuals turn to social networking sites with specific goals in mind, such as connecting with friends and family online, along with leveraging other features provided by the platforms (Obaraand and Oeldorf-Hirsch 2020: 142).

Another very important aspect to be evaluated is that traditional views on standard-form contracts falter when applied to social networking sites. Typically, these provisions are based on the belief that users logically

decide to accept these contracts or terms of use. They also assume that a competitive market leads companies to create fair terms, with the idea that competition promotes fairness and reasonableness. Yet, social networks operate differently. The way people decide to join a social network and accept its terms of use often lacks full rationality. Additionally, the concept of a competitive market doesn't fit well with social networks, which tend to move towards monopolistic control rather than offering a variety of choices. This tendency towards monopoly limits the ability of market forces to naturally ensure the fairness of terms through competition. Therefore, the unique aspects of social networks – like the irrationality of consumer decisions and the lack of competitive pressures – indicate that the traditional reliance on legal assumptions and market dynamics for ensuring fair terms of use might be flawed (Chiu 2011: 167–8).

6. Conclusion

In conclusion, the challenge of navigating and understanding the Terms of Use on social networking platforms is a significant issue, with users frequently binding themselves to complex legal agreements without true awareness. This situation underscores the urgent need for clearer, more accessible terms that empower users to make informed decisions. The potential of Artificial Intelligence and machine readability in transforming these dense legal texts into understandable, user-friendly formats represents a promising step forward. By embracing these technologies, we can move towards a digital ecosystem where transparency, fairness, and informed consent are not merely ideals, but realities. Achieving this will require concerted efforts from platform providers, legal experts, and policymakers to ensure that the digital space is as welcoming and safe as it is vast and uncharted.

References

- Arató, Balázs 2022. Norm clarity in the light of Hungarian case law. *Magyar Nyelvőr*, 46 (special issue), 81–90. <https://doi.org/10.38143/Nyr.2022.5.81>
- Arató, Balázs – Balázs, Géza 2022. The linguistic norm and norm of legal language. *Magyar Nyelvőr*, 46 (special issue), 91–103. <https://doi.org/10.38143/Nyr.2022.5.91>
- Arató, Balázs 2023. A végrendeletek értelmezésének egyes kérdései. *Magyar Nyelvőr*, 147(1) 78–92. <https://doi.org/10.38143/Nyr.2023.1.78>
- Balázs, Géza 2023. *Az internet népe*. Ludovika Egyetemi Kiadó. Budapest.

- Bauguess, Scott W. 2018. *The Role of Machine Readability in an AI World*. Available at SSRN: <https://ssrn.com/abstract=3226506> or <https://doi.org/10.2139/ssrn.3226506>
- Call, Andrew C. and Wang, Ben and Weng, Liwei and Wu, Qiang 2023. *Human Readability of Disclosures in a Machine-Readable World* Available at SSRN: <https://ssrn.com/abstract=4561569> or <https://doi.org/10.2139/ssrn.4561569>
- Case C-472/10. Nemzeti Fogyasztóvédelmi Hatóság v Invitel Távközlési Zrt. ECLI:EU:C:2012:242.
- Case C-92/11 RWE Vertrieb [2013] ECR. ECLI:EU:C:2013:180.
- Case C-125/18 Gómez del Moral Guasch ECLI:EU:C:2020:138.
- Case C-26/13. Árpád Kásler and Hajnalka Káslerné Rábai v OTP Jelzálogbank Zrt. ECLI:EU:C:2014:85.
- Cases C-776/19 to C-782/19. VB and Others v BNP Paribas Personal Finance SA and AV and Others v BNP Paribas Personal Finance SA and Procureur de la République. ECLI:EU:C:2021:470.
- Chiu, Aaron T. 2011. Irrationally Bound: Terms of Use Licenses and the Breakdown of Consumer Rationality in the Market for Social Network Sites. *Southern California Interdisciplinary Law Journal*, vol. 21, no. 1.
- Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts.
- Curtotti, Michael and McCreath, Eric 2013. A Right to Access Implies A Right to Know: An Open Online Platform for Research on the Readability of Law. *Journal of Open Access to Law*, 1.
- Eliot, Lance 2021. *AI In The Law Impeded Due to Machine Readability Of Judicial Decisions*. Available at SSRN: <https://ssrn.com/abstract=3990474> or <https://doi.org/10.2139/ssrn.3990474>
- Gosztonyi, Gergely and Lendvai, Gergely Ferenc 2023. *Deepfake: A Multifaceted Dilemma in Ethics and Law*. *Journal of Information Ethics*, vol. 2 no. 2. pp. 109–21.
- Karanicolas, Michael 2021. Too Long; Didn't Read: Finding Meaning in Platforms' Terms of Service Agreements. *University of Toledo Law Review*, vol. 52, no. 1.
- Lemley, Mark A. 2006. Terms of Use. *Minnesota Law Review*, vol. 91, no. 2.
- Loos, Marco 2017. Double Dutch – On the Role of the Transparency Requirement with Regard to the Language in Which Standard Contract Terms for B2C-Contracts Must Be Drafted. *Journal of European Consumer and Market Law*, vol. 6, no. 2.
- Obaraand, Jonathan A. and Oeldorf-Hirsch, Anne 2020. The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, vol. 23, no. 1.
- Papp, János Tamás 2017. A közösségi oldalak felhasználási feltételeinek jogi természete. In: Koltay, András – Török, Bernát (szerk.): *Sajtószabadság és médiajog a 21. század elején*. 4. kötet, Wolters Kluwer, Budapest, 187–212.

- Park, Kevin 2013. Facebook Used Takedown and It Was Super Effective! Finding A Framework for Protecting User Rights of Expression on Socialnetworking Sites. NYU. *Annual Survey of American Law*, vol 68, no. 891.
<https://doi.org/10.2139/ssrn.2252252>
- Perry, Amanda C. 2016. Click Yes to Accept Surveillance. *Business & Bankruptcy Law Journal*, vol. 4, no. 1.
- Schmitz, Benedikt and Pavillon, Charlotte 2020. Measuring Transparency in Consumer Contracts: The Usefulness of Readability Formulas Empirically Assessed. *Journal of European Consumer and Market Law*, vol. 9, no. 5.
- Szűts, Zoltán and Gerencsér, Péter 2020. A társadalmi nyilvánosság webkettes szerkezetváltozása. *Jel-kép*, 2. <https://doi.org/10.20520/JEL-KEP.2020.2.39>
- URL1. <https://platform-contracts.digital-strategy.ec.europa.eu>
- URL2. <https://www.jdsupra.com/legalnews/thoughtful-presentations-of-terms-of-68845/>

János Tamás Papp

PhD Assistant Professor

Pázmány Péter Catholic University

Faculty of Law and Political Sciences

Department of Civil Law

E-mail: papp.janos.tamas@jak.ppke.hu

<https://orcid.org/0000-0001-8682-6900>

**THE CONNECTION POINTS BETWEEN SMART CONTRACTS AND ARTIFICIAL
INTELLIGENCE – OPPORTUNITIES AND CHALLENGES**

Abstract

This study explores the opportunities and challenges presented by smart contracts and their technological foundation, the blockchain. It details the history and mechanics of blockchain technology, emphasizing its decentralized, immutable, and tamper-proof nature. The paper delves into the concept of smart contracts, tracing their origins to Nick Szabo and their evolution through the Ethereum blockchain. It highlights the role of oracles in enabling smart contracts to interact with the physical world and addresses various practical issues such as immutability, code errors, and the legal implications of smart contracts. The study also examines the integration of artificial intelligence (AI) in smart contracts, discussing how AI can function as an oracle to provide reliable information and support the contractual process. By examining different types of blockchains and smart contracts, the study provides insights into their potential applications and the inherent limitations and risks associated with their use, particularly in terms of legal enforcement and jurisdiction.

Keywords: smart contracts, blockchain technology, distributed ledger, immutability, oracles, AI oracles, Ethereum, legal implications, legal nature of smart contracts, code errors

**I. Technological foundation of smart contracts – the blockchain
technology**

a) Brief history of blockchain

In 2008, Satoshi Nakamoto issued the so-called Bitcoin Whitepaper (URL1), a technology and business plan describing in detail how the first and one of the most important blockchain networks, the Bitcoin blockchain works.

b) Definition and technological side of blockchain

A blockchain is a decentralised, distributed electronic database ("ledger"), which is structured as a chain of consecutive – as a rule – tamper-proof and unalterable blocks.

The obvious question arises as to what is the basis on which the transactions on the blockchain are carried out. The answer to this is the protocol governing the blockchain, which is accepted by the participants

when from the outset, and which is able to ensure the operation of the system in a quasi-self-regulating way.

The participants are called “*nodes*”, which are in the physical sense computers, but would not be meaningful in themselves without related human decision-making, so nodes can also be determined as the people operating them. In the case of so-called public blockchains (see later), a node can be anyone who downloads and installs the blockchain software and starts running it on their computer. The blockchain requires a high amount of electronic storage space, which leads to distinguishing two main types of nodes. Full nodes store the full version of the blockchain starting from the first, so-called genesis block, while lightweight nodes store only a simplified version of the blockchain (block headers) on their computers, saving significant space on the storage device (URL2).

The third category of nodes are the so-called “miners”, which, in addition to storing the entire blockchain, are also actively involved in its expansion. Put simply, miners validate subsequent blocks with the transactions in them (a block usually contains several transactions), continuously extend the chain by generating new blocks (URL3). While transactions are awaiting validation, miners compete to be the first to find the solution to a mathematical problem in order to create the next block and to submit it to and through the nodes of the system, thereby extending the blockchain. This is the so-called proof-of-work consensus mechanism used by the Bitcoin blockchain, whereby one party (the miner) has to perform a complex calculation, while sacrificing resources to receive something in return, and the other parties (the nodes) can easily and quickly verify that the reward has been earned (URL4). The “winning” miner is rewarded with newly minted bitcoins,¹ plus a transaction fee paid by the parties to the transaction (URL5).

How does the blockchain look like? As mentioned above, the first block is the so-called genesis block. Subsequent blocks, which can be extended indefinitely, contain, among other things, the details of the transaction and a hash, which can be defined as the unique fingerprint of the block. The hash is a security function, which is used to verify whether the content of the block has been tampered with or changed in any possible way. In addition to its own hash, a block also contains the hash of the block immediately preceding it (URL6).

The immutability and tamper-proofness of blocks and the transactions they contain are made possible by a combination of the above technological solutions. Any attempt to change a block would require not only the specific

¹ From April 2024, the reward of the miners is 3,125 BTC per block.

block to be changed, but also the contents of the blocks that follow it in the chain, which are often large in number. Changing an existing block would also change the hash associated with that block. When checking the validity of a given block, it would be immediately apparent that the hash is not the one related to the data stored in the block. As the hash of each block is used to generate the hash of all subsequent blocks, changing a block will also change the hash of subsequent blocks.

c) Different kinds of blockchain

In terms of participation in the blockchain network, we distinguish between three main categories, based on the right to read the content of blocks and the right to add and control them.

This classification distinguishes between public, private and consortium blockchains. A typical example of a public blockchain is the above-mentioned Bitcoin blockchain, in which anyone can participate as a node and even as a miner once the software has been installed. In the case of private and consortium blockchains the principle of decentralisation is inadvertently undermined, as the granting of participation and the necessary rights presupposes central actor(s) with additional rights. In the private blockchain, a closed circle, typically members of an organisation or a profession, participates. A consortium blockchain differs in that the participants are generally composed of members of several organisations (URL7).

A second classification distinguishes between permissionless and permissioned blockchains. The former is often associated with the public blockchain and the latter with the private and consortium networks, but this is not necessarily true in all cases, although the pairing most often occurs in such constellations. In practice, public, private and consortium blockchains can all be created in a permissioned or permissionless form. In the case of a permissioned blockchain, the ability to control and/or add transactions to the blockchain is restricted for certain nodes or a certain set of nodes, whereas in the case of a permissionless blockchain, there are no such restrictions.

We can see that in the case of private and consortium blockchains, not all of the elements/principles indicated above are fully implemented, especially the principle of decentralisation, it is therefore questionable whether the blocks in them are truly immutable. However, this does not and should not diminish the practical benefits of these systems.

In addition, it should be borne in mind that the use of blockchain – which is also true for smart contracts – is not always the most optimal solution.

Often, the novelty and modernity of the technology is an excellent marketing tool ("hype"), emphasising the innovative approach of the economic operator concerned, but it is always necessary to consider whether the objective can be achieved more effectively with classical centralised solutions.

II. Basics of smart contracts

a) The beginning of smart contracts and definition

Smart contracts have become widespread thanks to the second largest blockchain, the Ethereum. Like the Bitcoin, this term does not only cover a kind of blockchain, but also currently the second most significant cryptocurrency, the Ether (URL8). In order to further exploit the potential of the blockchain technology, it became necessary to create an additional "layer" within the system. Thanks to this layer, in Ethereum, it is possible to create smart contracts (URL9).

However, the inventor of smart contracts was not Buterin, but a Hungarian-born American computer scientist-cryptographer, Nick Szabo, who first created the concept of smart contracts in the mid-1990s (URL10). According to Szabo, smart contracts are computer codes that fulfil contractual conditions. He envisaged that this would reduce the need to rely on trusted third parties, and that automatic performance will reduce the risk of abuse and will result in lower costs of performance.

Similarly to blockchain, there is no officially recognised definition of smart contracts, which can also be defined by their properties, mostly based on Nick Szabo's concept presented above. In other words, a smart contract is a computer code that is created between two or more parties, and when the event(s) specified in the code occur, the contractual obligations will be automatically fulfilled. However, a significant difference with Szabo's definition is that smart contracts in the 21st century are running on the blockchain, which ensures that they are immutable and tamper-proof. While immutability is certainly to be welcomed, since it is in the interest of the contracting parties that their agreement should not be falsified, it virtually precludes the possibility of amending the contract in the event of a change in the circumstances and the parties' relations, which is far from being positive.

b) Oracles – Communication of smart contracts with the outside world

Although an increasing part of our lives takes place in cyberspace, the physical world still remains the most relevant area which cannot be ruled out.

To this end, it is essential to ensure that smart contracts are able to communicate effectively and securely with the physical space.

The means of doing this are the so-called oracles which/who provide proof that the conditions programmed into the smart contract have been fulfilled (URL11). The oracles can be natural persons, for instance notaries, who, by virtue of their official nature and obligations of neutrality and impartiality, fit perfectly into the oracle role. Besides, software, websites, databases and registries can also function as oracles, just as hardware (e.g. sensors) that are able to communicate with each other and with the specific smart contracts thanks to the Internet of Things (IoT) phenomenon (URL12).

The essential characteristics of oracles are the ability to credibly attest the fulfilment of conditions off-chain (i.e. outside the blockchain) and to connect to the smart contract on the blockchain. In case of use of centralised oracles, there is a risk of failure if they get hacked, manipulated, or shut down, leading to the circumvention of the contractual conditions by any of the parties. In order to tackle this issue it is possible to programme several oracles (decentralised oracles) at once into the smart contract and also to specify a “yes/no ratio” for conflicting information to be accepted (URL13).

In addition to all this, it is important to underline that the – often voiced – presumption of authenticity of the data registered in blockchain and smart contracts is in itself wrong, and any recognition of this in law or in practice could lead to significant harm. Technology alone does not make data authentic. In order for this data to be considered authentic, it is essential that oracles are involved, whereby data from the physical world is passed through a reliable “filter” into the on-chain space.

In the case of smart contracts, the role of the oracle can even be fulfilled by artificial intelligence. Generally, artificial intelligence as an oracle can be extremely useful in many aspects throughout the entire contractual process. Primarily, the “if-then function” is notable, which can assist the parties in cases where a legal consequence is tied to the occurrence of an objective condition (e.g., weather parameters, exchange rate changes, price developments, inflation, election or sports results, etc.) under the contract. In such cases, the artificial intelligence oracle provides reliable information about the occurrence of these objective conditions to the blockchain platform, thus triggering the legal consequence tied to the condition in the contract.

An example worth mentioning is when the parties are negotiating off-chain, and the AI oracle supports the formation of the smart contract by translating natural language into computer language. The AI oracle can also function to oversee contractual performance, establishing, for instance, non-

performance by the obligor and, within certain limits, the reason for it. If, for example, the subject of the contract is not delivered, the AI oracle examines whether the non-performance occurred within the obligor's sphere of interest or, for instance, was due to force majeure, and then forwards the result of the examination to the blockchain platform. This information either triggers the contractual consequence of non-performance (e.g., a penalty) recorded in the smart contract or prevents it from taking effect, considering the unavoidable external cause. Finally, the AI oracle can also serve to modify the smart contract in cases where unforeseen changes in circumstances occur that necessitate or justify modifying the contract. In such cases, the AI oracle sends this information to the blockchain platform, enabling the smart contract to be modified before self-execution. Of course, other uses of the AI oracle are conceivable; only the most typical ones have been mentioned above.

We can differentiate between software and hardware oracles. While the former can relay any online accessible information to the blockchain system, the latter are connected to and gather information from the physically perceptible world. The literature also distinguishes between outbound and inbound oracles, depending on whether the oracle relays information from an external source to the blockchain platform or vice versa, from the blockchain platform to the physically perceptible world. Lastly, it is worth mentioning the distinction based on whether the oracle merely relays information from the physically perceptible world or also performs calculations and transmits their results to the blockchain platform. The former are referred to in the literature as data carrier or automated oracles, while the latter are known as computation oracles. A good example of a calculation-performing oracle is when someone's creditworthiness must be assessed based on various criteria, and if the individual is deemed creditworthy, this automatically triggers the loan disbursement.

Generally, regardless of how artificial intelligence is used, the so-called black box effect arises, meaning that in some cases, the decision-making processes or operational mechanisms of AI systems are not transparent or comprehensible to human users, and often not even to developers. If artificial intelligence serves as an oracle in a smart contract, this phenomenon poses considerable risks, as the oracle may provide inaccurate or false information to the blockchain platform, unjustifiably triggering or preventing a contractual consequence. This phenomenon, referred to in the literature as the oracle problem, can be addressed by so-called consensus oracles. This refers to a decentralised network of oracles that work together to achieve the same goal, meaning they reach a result on the same issue, but not the individual results, for example, their average, is the information relayed to the blockchain platform

and triggers a contractual consequence in the smart contract. However, it is emphasised that the use of consensus oracles does not eliminate the problem described above but only reduces the likelihood of severe functional disorders. Examples of this include the unjustified fulfilment (self-execution or self-performance) of the smart contract or its non-fulfilment due to AI error (the failure of self-execution or self-performance), as well as the formation or non-formation of the smart contract if the contract is created by AI. An example of an AI-created contract is the AI's assessment of creditworthiness, depending on the result of which the loan or credit agreement is either formed or not.

Thus, the improper functioning of the artificial intelligence oracle can cause numerous problems and raises the issue of legal liability from multiple perspectives. The literature has proposed various views and approaches regarding who bears the risk of damage and who should be liable if artificial intelligence acts as an oracle in a smart contract and a malfunction occurs, resulting in harm to one of the parties. Without delving into the possible forms of legal liability in detail here, it is merely noted that analogies based on the liability of parents and animal keepers have even appeared in connection with the AI service provider's or AI owner's liability. (Papadoulis and Papakonstantinou: 2023) If a single guiding principle were to be established concerning the complex liability and risk-bearing issues associated with artificial intelligence oracles, it would likely be the emerging principle that the party fundamentally responsible for any disruptions endangering the proper execution of a smart contract is the one who requested the involvement of the artificial intelligence as an oracle, or the one obligated to ensure the correct and flawless operation of the artificial intelligence. In summary, the liability issues related to smart contracts are particularly unique because the legal relationships involve entities that, although not parties to the contracts themselves, can significantly influence them through their actions. These entities either provide the platform for the execution (performance) of the contract or facilitate or hinder its completion by transmitting information. Therefore, legal science cannot approach these liability issues from traditional perspectives; instead, innovative thinking and highly multifaceted analysis are required to develop clear and fair solutions for these entirely new legal disputes. It is also possible that artificial intelligence will play a key role in mapping out all possible variations and potential scenarios requiring resolution.

c) Is a smart contract really a contract?

The subtitle above raises the following two important questions: 1. Are smart contracts really smart? 2. Are they contracts at all?

As determined above, a smart contract is nothing more than computer code written by humans that performs predetermined actions when predetermined conditions are met. Both the conditions and the actions are therefore determined by the human programmer, not by the contract/code. The smart contract therefore has no intelligence or creativity to create for itself the obligations to be fulfilled and the related conditions, nor does it shape the code that has already been programmed.

It is much more difficult to give a clear answer to the second question. In the case of smart contracts, we are dealing with a global phenomenon which did not emerge in a particular country or its legal system and become cross-border in nature, but which developed in a cross-border environment from the outset. With smart contracts, individual states and communities of states (including the European Union) have been given a ready-made product that is an integral part of everyday reality and which they have had to understand and define how it fits into their legal system. Nevertheless, with rare exceptions, individual states have not been able to develop a specific regulatory environment quickly and effectively. As a consequence, it is necessary to assess whether smart contracts constitute a contract in a given country on the basis of the current rules of contract law and, if so, whether they fulfil the legal requirements of a written form which is a condition of validity for several transactions.

As smart contracts are relatively new and still evolving developments in law and technology, there is no consensus regarding their legal nature. Some opinions hold that smart contracts cannot be considered contracts in the legal sense but are merely computer protocols that facilitate the execution (performance) of already established agreements. In contrast, another viewpoint asserts that smart contracts are indeed contracts capable of entirely replacing traditional agreements. There is also an opinion that questions the contractual nature and character of smart contracts primarily based on whether the coding and execution of the smart contract accurately and completely reflect the parties' intentions, which is an indispensable conceptual element of contracts as per the Civil Code. (Werbach and Cornell: 2017) This issue may particularly arise for contracting parties who are unfamiliar with the technology used in smart contracts or do not understand the programming language involved.

d) Self-enforcement

In addition, it is crucial to clarify how enforcement relates to smart contracts, mainly for terminological reasons. Smart contracts, by virtue of their

automated nature, their immutability and unforgeability, conceptually exclude the possibility of the breach of contract. Enforcement can take place in case of breach of contract according to Hungarian (and many other national) legal terminology. Nevertheless, in several cases we hear/read about the automatic enforcement or self-enforcement of smart contracts. The root of the problem lies in the English terminology commonly used to describe the characteristics of smart contracts. In English, often the term “self-enforcing or self-executing contract” is used. In reality, however, we are talking about self-performing agreements. Therefore, in the English terminology, it would be more appropriate to use the term “self-performing contract” or “self-fulfilling contract” which could lead to the avoidance of misunderstandings.

III. Key practical issues related to smart contracts

It should be noted at the outset that the questions that arise cannot always be answered in a clear-cut way. With the exception of a few countries where regulation has already taken concrete shape, the answers are mostly theories developed by academics and researchers, which may however serve as perfect basis for future regulation.

It may be worth comparing the smart contract with the concept of a contract under the Hungarian Civil Code (Ptk.) and the recognised methods of contract formation in the Civil Code to get closer to answering the question of whether a smart contract is a contract under Hungarian law. According to Section 6:58 of the Civil Code, a contract is a mutual and concurrent declaration of will by the parties, creating an obligation to perform a service and a right to claim the service. A contract is formed through the mutual and concurrent expression of the parties’ will, either orally, in writing, or through conclusive conduct. For certain types of contracts, the Civil Code requires written formality and validity (e.g., contracts affecting the ownership rights of real estate). Implicitly, the Civil Code also defines the requirement of written form by clarifying that a declaration must be considered in writing if it is made in a form suitable for recalling the content of the declaration unchanged, identifying the person making the declaration, and the time of making the declaration. Consistent judicial practice considers electronic signatures that meet the legal requirements as such, while simple email exchanges are regarded merely as conclusive conduct. (In summary, it can be stated that an email does not meet the written requirements necessary for unequivocally determining the identity of the declarant, and thus a legal statement made via electronic mail cannot be considered as being in written

form. This is because, instead of verifying the identity of the sender, it can at most be established from whose electronic mailbox the email originated. This remains true even if the email includes a so-called embedded signature. The relevant judicial practice is now considered uniform: contracts are formed through conclusive conduct as a result of an offer and acceptance made via email. Compared to other forms of behaviour reflecting the intent to enter into a contract, the sole advantage of conclusive conduct conveyed by electronic mail is that the content of the agreement can be more easily proven later.) In line with the Civil Code's spirit and technological neutrality, in principle, the smart contract can meet the requirement of written form, meaning it is conceivable that smart contracts will play a role in simpler real estate transactions in the future, initially in ensuring the execution of contractual content. However, as long as it is necessary to submit paper-based copies of contracts to the land registry authority, smart contracts may only play a supplementary supporting role in real estate transactions, but this could undoubtedly increase security.

a) Possible areas of application of smart contracts

The first question to be asked is which transactions can be covered by smart contracts at all. To answer this question in concrete terms, we must examine the contract law of the country in question, both in respect of material and formal validity requirements. In case if the content of a smart contract is in conformity with the law of a given country, it can be a valid contract (URL14). However, numerous states provide for formal validity rules for specific transactions.

If we move away from the national legal requirements that may stand in the way of the validity of smart contracts, we need to take into account the essential differences between human language and code language. Human language is characterised by its flexibility, which allows us to express our ideas in a nuanced way. In the context of contract law, this means that we can put almost any provision into a contract which is in conformity with the applicable law. The wording of our contracts often allows for multiple interpretations, precisely because of the diversity of human language. In contrast, computer code is a rigid, exact "language" where interpretation has no role, or if it does have some, it is extremely limited.

Due to the rigidity of the code language, smart contracts are best suited to agreements that are highly constrained even when written in human language, and where the parties have little or no desire to leave room for different interpretations. This is one of the reasons why smart contracts first appeared in financial transactions, which traditionally contain rigid provisions.

One of the main limitations of smart contracts is their inability to handle abstractions such as good faith, which often require human interpretation, context-dependent decision-making, and flexible judgement that code-based systems currently cannot provide. Smart contracts, functioning as program code, automatically execute based on pre-defined conditions and events. Consequently, these contracts rigidly follow the code and cannot consider factors such as the parties' intentions, the context of the situation, or the behavioural requirements posed at a fundamental level by civil law.

b) Jurisdiction and applicable law

As mentioned above, both the blockchain technology and smart contracts are a global phenomenon, with “dispersed”, in most cases anonymous, nodes around the world, and with the possibility that the contracting parties are not located in the same country. The legislators are therefore faced with the task of classifying and/or regulating a phenomenon which, due to its inherently cross-border nature, is not subject to purely domestic experience.

One of the first steps in a judicial procedure is to establish whether we are facing a cross-border agreement at all and, if so, to determine the jurisdiction and the applicable law. It is important to stress, however, that the mere fact that a given smart contract runs on the computers of nodes all over the world does not in itself mean that the given smart contract is a cross-border agreement [ELI Principles on Blockchain Technology 4 c)]. In this context, it can be useful for the contracting parties to consider choosing the applicable law if there is no mandatory legal provision which excludes this option (e.g. for real estate transactions). The choice of law can help avoiding difficulties that may arise in the absence of choice, especially when the parties choose the laws of a legal system having introduced specific and clear rules for smart contracts.

c) The immutability of smart contracts

As explained above, smart contracts are based on blockchain technology, which conceptually precludes the subsequent reversal or modification of blocks and transactions. On the one hand, this is positive, as immutability also leads to tamper-proofness, contributing to the security of transactions through technological means.

However, one also has to consider the negative side to his kind immutability. Relations between the parties are in numerous cases dynamic, just as the parties to a transaction and the circumstances. In longer-term relationships, the need to modify the contract is common, but may not materialise if the

parties have agreed in a smart contract. For foreseeable events, parties can program changes into their smart contracts from a fixed date which will be carried out automatically, but in unforeseeable situations requiring quick actions, smart contracts are not the most convenient choice. The solution to this issue may be the use of such underlying blockchain technology which allows the subsequent modification of the blocks or the implementation of technically feasible reverse transactions (ELI ELI Principles on Blockchain Technology, Principle 10.).

Immutability can also lead to conflicts with consumer protection rules. Just to highlight one of these rules: in the case of distance contracts, the 14-day right of withdrawal granted to the consumer by Article 9 of the above-mentioned Consumer Rights Directive (Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council Text with EEA relevance). It is precisely the immutability of smart contracts that could prevent the consumer's right of withdrawal from being exercised, thus lead to the situation when the smart contract does not comply with the relevant legislation on consumer protection.

d) Code errors

Smart contracts may be affected by code errors or bugs. A code bug can result in significant financial losses for the parties to the contract and raises serious liability issues. One of the most serious code bugs in the past concerned the Ethereum system, more precisely the so-called The DAO (Decentralised autonomous organisation, details of which will be described later). In the case of The DAO, the smart contract contained a code bug which was exploited by some participants to cause a loss of 3.6 million Ethereum (USD 50 million at the time) to the community. Crucially, there was no consensus among the participants in The DAO on how to proceed in such a case. The majority of those who adhered to the principles of smart contracts believed that the status quo should not be touched, as "the code is law". The majority, however, was of the opinion that the damage had to be remedied, and in order to do so, breached the principle of immutability by consensus (URL15). This led to the so-called hard fork of the Ethereum blockchain, resulting in the genesis of the Ethereum classic and Ethereum. The former is effectively a renaming of the cryptocurrency still used in the previous system, while Ethereum is used by the majority who accept the changed state (which includes the refunded damage) (URL16).

In the event of a code error, it is essential that it is fixed quickly to ensure the reliability of the system and to avoid user abandonment. The question is: who is responsible for these errors and the resulting damage: the programmer, the nodes, or the contracting parties? The programmers are certainly the ones who should bear the responsibility, but their actual liability and the obligation to pay damages is questionable because of their frequent anonymity and the problem of establishing jurisdiction and applicable law in the event of a breach of anonymity. This results in that in many cases the only means of remedying the problems is to take a decision similar to that in The DAO case.

Another important factor is that most of the software used to create smart contracts is open source, which means that – within the limits set by its creator – anyone is free to develop it further. In the course of such development, a code error in the program further complicates the assessment of liability.

e) A Linguistic Perspective on Smart Contracts, Some Specific Issues of Interpreting Smart Contracts

The examination of smart contracts from a linguistic perspective can be particularly enlightening, as their emergence introduces new contexts for the long-established civil law principles developed over centuries regarding the interpretation of contracts written in natural human language. According to the current Hungarian Civil Code, a legal statement, in case of a dispute, should be interpreted in a manner that considers the intended meaning that the recipient, given the declarant's presumed intent and the circumstances of the case, was expected to understand based on the generally accepted meaning of the words. Judicial practice has further refined this main rule by considering factors such as what an average educated consumer might have understood when concluding a particular contract or reading its text, or what meaning one party might have attributed to a particular expression depending on their own level of education.

One might think that such issues do not arise in connection with smart contracts since they are written not in human language but using programming codes. Nonetheless, there can be situations where interpretation becomes necessary, although smart contracts fundamentally aim to execute agreements between parties in an automatic and unambiguous manner. Smart contracts are written in code, which must be executed according to precisely defined rules. However, coding errors, misinterpretations, or a lack of proper understanding of business logic can cause problems.

Although smart contracts operate independently from a technical standpoint, they can still be part of traditional contracts and fall under local

legal regulations. Consequently, in case of a dispute, courts might need to interpret what the code actually meant to the parties. Did the parties understand what specific operation the code performed? When drafting smart contracts, it is not always clear what (business) intention the different parts of the code reflect. If a dispute arises over the meaning of certain parts of a smart contract, the documents related to the contract preparation, emails, and other communications can be relevant during interpretation.

Errors made during the drafting of a smart contract or unforeseen events occurring in the meantime can also raise interpretive issues, especially if it is not clearly defined what should happen in such cases. It is evident that a different dimension of interpretation opens up in the case of smart contracts, and this interpretation's starting point will not be the linguistic understanding of the contract text, but rather inferring the parties' intentions based on other specific aspects.

In conclusion, although the purpose of smart contracts is to minimize interpretive issues through automatic execution, in reality, situations can still arise where communication and intention between the parties play a crucial role in the precise interpretation of the contract. If linguistic interpretation based on the generally accepted meaning of words is necessary at all, it is limited to the statements and declarations of intent made during the negotiations preceding the contract conclusion.

IV. Categorisation of smart contracts

Generally, we can distinguish four main variants of smart contracts, however, this classification is not official and there may be many more situations and perspectives for the smart contract classification.

The first variant is the smart contract created purely as a code, without the use of human language. In this case, the question arises: is the code itself law?

In the case of the second type, the code is only a means of automatic fulfilment of a legal agreement off-chain and written in human language. In this case, the smart contract is not a contract, but a tool of performance alongside the human language contract.

In the third version, the contract written in human language is converted into a smart contract, i.e. a code. In this case, the parties aim to take advantage of blockchain technology and automatic fulfilment. The risk of this variant lies in the differences between the two (rigid code and flexible human) languages, as presented above. Not every nuance of the human language

can be clearly transformed into the code language, so there is a real risk that the code will contain something other than the parties' original intentions as expressed in human language. Theoretically, it is possible that the parties take the human-language contract as the prevailing one if this were to be discovered, but the unstoppable self-fulfilment of a smart contract makes it virtually impossible to stop the performance process.

The last variant is the hybrid smart contract, in which some provisions appear off-chain as a contract written in human language, and other provisions are on-chain in code, intended for self-fulfilment. This solution allows the parties to capture those concepts and provisions that leave open multiple interpretations and are thus difficult to convert into code (e.g. the force majeure is clearly one of them) in the off-chain world, and to place those that are suitable for adoption in self-fulfilling code in the on-chain environment.

Another but in many respects similar approach distinguishes three current forms of smart contracts. The first group includes smart contracts that serve to execute the performance of agreements established through other means, effectively fulfilling a security function. The second category encompasses hybrid contracts, where the contract is partially formed as a smart contract. The third form is the standalone smart contract, where all elements of the parties' legal relationship are contained within it. (Papadouli–Papakonstantinou 2023)

V. Some examples of smart contracts from the real world

a) Non-fungible tokens

Non-fungible tokens (or NFTs) are assets tokenized within the blockchain through a process called minting. Tokens are unique identification codes, stored on the blockchain with the corresponding assets elsewhere. NFTs many times take the form of pieces of art but their use is much broader than the field of art (e.g. also real estate can be tokenized). By having private keys to the given NFT, the rights joint to it can be exercised. The NFTs have a unique identifier associated with a blockchain address (URL17).

As shown by their name, NFTs are not fungible, irreplaceable, but are definitely tradeable on the blockchain through special marketplaces like the OpenSea (URL18). Often, the question arises, what the sense of an NFT artwork is in the era where digital files can be easily copied. The value of the NFTs is largely attributed to the originality, i.e. the related private key implies that its owner is the owner of the original digital file. This is not different in the physical world either: for instance, a famous painting of Rembrandt can

also be easily copied but the copy – even if of perfect quality – will not give any value close to that of the original one. However, compared to the physical world, NFTs include a digital and blockchain-based certificate attesting their originality, and as we saw it above, the blockchain technology provides the advantage of tamper-proofness.

The process of minting several times entails smart contracts which assign ownership and make possible the secure transfer of NFTs. There exist several smart contract blockchains with NFT creation tools (e.g. TRON, Tezos). The advantages of smart contracts for minting and trading NFTs are multifold. Smart contracts can automatically check the time and place of creation of NFTs and they can help prove ownership and its change through the use of the above-mentioned digital certificate (URL19). In some cases, smart contracts can be useful to guarantee the collection of resale royalties (within the field of copyright, resale rights entitle artists to a share of the sale price when their artwork is resold) after the subsequent sales of their NFT's (URL20).

b) Initial Coin Offering

In the financial-investment area, smart contracts are a common manifestation of so-called Initial Coin Offerings (ICOs). In an ICO, the initiating company raises a fixed amount of cryptocurrency from investors in order to launch its own token platform. This platform is set up to achieve a specific goal, just like a classical company.

Tokens can be divided into several types. Utility tokens can be used to access an online service and/or product developed or designed by the initiators of ICO projects. In addition, tokens can also be used as an investment and the profits generated can be distributed as a quasi-dividend to the token holders (security tokens).

The first ICO took place in 2013 and was announced by the then Mastercoin (URL21). However, in their current form, they have spread through the Ethereum blockchain, thanks in part to the system's ability to create smart contracts. A smart contract collects the required amount of cryptocurrency, automatically issues the corresponding number of tokens in exchange, later "manages" them, and is able to pay commissions and dividends.

It should be stressed that most ICO projects do not have a real product or service behind them, just an idea. Their advantages are speed, avoiding the administrative and bureaucratic burdens of the stock exchange and the transparency provided by blockchain technology. As Ethereum, for example, is a public blockchain, investors can track the use of the cryptocurrency they collect.

c) Decentralised autonomous organisations

Another practical example of smart contracts worth mentioning is the DAO, or decentralised autonomous organisation. The DAO scandal, which was the result of a code error and led to a hard fork of the Ethereum blockchain, was mentioned above. But what is the DAO as a phenomenon and a legal institution?

A DAO is a complex network of smart framework contracts and smart contracts into which the structure and operation of a virtual organisation is programmed. The “life” of a pure DAO takes place entirely in the on-chain world, without, for example, a physical seat or registration in national company/legal entity registers although their State registration is – depending on the fulfilment of legal conditions, like the form – possible under some jurisdictions (e.g. Wyoming, Gibraltar, Switzerland) (URL22). By joining, members receive tokens which grant them voting rights. The decisions are automatically implemented by the smart contract according to the rules it sets out.

In practice, transparency is a positive feature of DAOs, but the lack of hierarchy can lead to slow decision-making processes due to the large number of actors involved. The main risk, however, is legal uncertainty due to the lack of legal regulation in most of the countries globally in respect of the legal form, legal personality and liability.

References

- ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection, Report of the European Law Institute, Principle 4 c).
- ELI Principles on Blockchain Technology, Smart Contracts and Consumer Protection, Report of the European Law Institute, Principle 10.
- Kevin Werbach – Nicolas Cornell 2017. Contracts Ex Machina, *Duke Law Journal*, Volume 67, pp. 313–82.
- Vasiliki Papadouli – Vagelis Papakonstantinou 2003. A preliminary study on artificial intelligence oracles and smart contracts: A legal approach to the interaction of two novel technological breakthroughs. *Computer Law and Security Review*, Volume 51.

Internet Resources

- URL1. Nakamoto, S. (2008) Bitcoin: A Peer-to-Peer Electronic Cash System. <https://bitcoin.org/bitcoin.pdf> (Downloaded: 1 April 2024)
- URL2. Dr. Ravi Chamria, What are lightweight nodes in blockchain? Advantages and Disadvantages, <https://www.zeeve.io/blog/what-are-lightweight-nodes-in-blockchain/> (Downloaded: 1 April 2024)

- URL3. What are the Types of Nodes in Blockchain? <https://utimaco.com/service/knowledge-base/blockchain/what-are-the-types-of-nodes-in-blockchain> (Downloaded: 1 April 2024)
- URL4. Scott Nevil, What Is Proof of Work (PoW) in Blockchain?, <https://www.investopedia.com/terms/p/proof-work.asp> (Downloaded: 1 April 2024)
- URL5. Bitoin.com, What are Bitcoin network fees? <https://www.bitcoin.com/get-started/what-are-bitcoin-network-fees/> (Downloaded: 1 April 2024)
- URL6. Bitpanda, What is a hash function in a blockchain transaction? <https://www.bitpanda.com/academy/en/lessons/what-is-a-hash-function-in-a-blockchain-transaction/> (Downloaded: 1 April 2024)
- URL7. Shobhit Seth, Public, Private, Permissioned Blockchains Compared, <https://www.investopedia.com/news/public-private-permissioned-blockchains-compared/> (Downloaded: 1 April 2024)
- URL8. <https://ethereum.org/en/whitepaper/> (Downloaded: 1 April 2024)
- URL9. The most popular programming languages for the creation of smart contracts are: Solidity and Vyper, <https://chain.link/education-hub/smart-contract-programming-languages> (Downloaded: 5 April 2024)
- URL10. Nick Szabo “Smart Contracts” (1994) and Smart Contracts: Building Blocks for Digital Markets (1996)
- URL11. <https://ethereum.org/en/developers/docs/oracles/> (Downloaded: 5 April 2024)
- URL12. Smart contracts and IoT, <https://www.geeksforgeeks.org/smart-contracts-and-iot/> (Downloaded: 5 April 2024)
- URL13. Cem Dilmegani, Guide to Oracles in 2024: What Are They, Types & Use Cases, <https://research.aimultiple.com/blockchain-oracle/> (Downloaded: 5 April 2024)
- URL14. Example of Liechtenstein. <https://www.mondaq.com/technology/935298/blockchain-comparative-guide> (Downloaded: 5 April 2024)
- URL15. Samuel Falcon, The Story of the DAO – Its History and Consequences, <https://medium.com/swlh/the-story-of-the-dao-its-history-and-consequences-71e6a8a551ee> (Downloaded: 8 April 2024)
- URL16. Ethereum Classic and the Ethereum Hard Fork, <https://help.coinbase.com/en/coinbase/getting-started/crypto-education/eth-hard-fork> (Downloaded: 8 April 2024)
- URL17. Rakesh Sharma, Non-Fungible Token (NFT): What It Means and How It Works, <https://www.investopedia.com/non-fungible-tokens-nft-5115211> (Downloaded: 8 April 2024)
- URL18. <https://opensea.io/>
- URL19. All You Need to Know About NFT Smart Contracts, Binance Blog, <https://www.binance.com/en/blog/nft/all-you-need-to-know-about-nft-smart-contracts-568745413587703085> (Downloaded: 8 April 2024)
- URL20. Crypto Council for Innovation, How NFT Royalties Work – and Sometimes Don’t, How NFT Royalties Work – and Sometimes Don’t, <https://>

cryptoforinnovation.org/how-nft-royalties-work-and-sometimes-dont/
(Downloaded: 8 April 2024)

URL21. Ruben Merre, ICO 101 — History of Initial Coin Offerings (ICOs), <https://medium.com/hackernoon/ico-101-history-of-initial-coin-offerings-icos-part-1-from-mastercoin-to-ethereum-4689b7c2326b> (Downloaded: 8 April 2024)

URL22. Nestor Dubnevich, The Best Entities and Countries for DAO Registration in 2024, <https://legalnodes.com/article/choose-a-crypto-friendly-country-for-dao>
(Downloaded: 8 April 2024)

Balázs Arató

Habilitated Associate Professor
Károli Gáspár University of the Reformed Church
in Hungary, Faculty of Law, Department of
Commercial and Financial Law/Department of
Civil Law

E-mail: arato.balazs@kre.hu
<https://orcid.org/0000-0001-5854-8018>

Tamás Sajben

Head of the Brussels Representation at
Hungarian National Chamber of Notaries
E-mail: sajben.tamas@kamara.mokk.hu

NLP-ENHANCED CLASSIFICATION OF REMOTE EMPLOYMENT OPPORTUNITIES

Abstract

In the rapidly evolving digital era, the labour market is undergoing significant transformations, particularly in online job searching. This progress, however, presents the challenge of efficiently filtering a large amount of information, making artificial intelligence (AI) tools, especially those using Natural Language Processing (NLP), increasingly vital.

In this study, which is aimed at identifying remote work opportunities, an automated text classification system that merges linguistics with AI has been developed. This system goes beyond merely sorting job offers; it focuses on understanding and interpreting the language of job advertisements for effective computer comprehension, while an attempt is made to preserve the semantics of the text.

The research commences with data preparation, where texts from online job portals are subjected to technical filtering and normalization, along with linguistic analysis for key feature extraction. Subsequently, the exploration of applying rule-based classification and supervised machine learning algorithms to this textual data is undertaken, which serves to demonstrate AI's proficiency in linguistic categorization.

The accuracy of these AI-driven methods in job advertisement classification is underscored by a comparative analysis of model performance, highlighting the synergy between linguistic principles and AI. The study concludes that the efficiency of online job searching for remote positions is significantly improved by this AI and NLP-enhanced system, illustrating the growing intersection of linguistics and AI in data analysis.

Keywords: NLP-based text classification, machine learning models, rule-based learning methods, text processing

Introduction

In the contemporary digital era, the job market has experienced a paradigm shift, particularly in the realm of remote work. This shift is mainly due to digitalization and the COVID-19 pandemic, especially in fields requiring high online connectivity, like IT. Developers and other IT professionals increasingly prefer remote work for its broader opportunities. This trend is mirrored on various online job advertisement platforms, where remote positions are prominently featured.

Given the vast number of listings, artificial intelligence, specifically Natural Language Processing, becomes crucial for efficiently classifying and finding suitable opportunities. This shift has necessitated the development of sophisticated tools to efficiently navigate and categorize the influx of online job opportunities. NLP presents a robust framework for addressing these needs by analysing the language of job advertisements to discern remote positions from traditional ones.

This paper investigates the application of NLP in enhancing the classification of remote employment opportunities. By integrating linguistic analysis with artificial intelligence, the research develops an automated classification system designed to shift through online job postings, selecting those that offer remote work.

The significance of NLP in this context is manifold, drawing upon its proven efficacy in text analysis and machine learning to adapt to the nuances of job advertisements. The methodologies employed, including data extraction, preprocessing, and vectorization, are fundamental to transforming raw textual information into a structured form suitable for classification.

A case study will provide the empirical foundation for the research, applying the discussed NLP techniques to actual job advertisements and evaluating their effectiveness. The study aims to demonstrate the practical usefulness of NLP-driven systems in enhancing the searchability and classification of remote jobs, reflecting the evolving requirements of the job market.

Theoretical framework and Methodological approach

Essence of NLP

In the field of artificial intelligence, NLP plays a crucial role as an advanced channel between human communication and machine interpretation. This technology is indispensable across a wide range of applications, significantly improves the ability of machines to understand human language. From basic translation services to complex sentiment and semantic analysis, NLP capabilities are essential for analysing large datasets, automating interactions, and aiding decision-making processes in a wide range of industries.

A compelling example of NLP's application is highlighted in related work focusing on customer satisfaction (Tusar & Islam 2021) – a critical factor for business success in today's competitive landscape. Many organizations, recognizing the importance of understanding and meeting customer needs, invest heavily in various strategies. However, traditional manual analysis often

falls short in accurately addressing the complex needs of customers, leading to decreased satisfaction levels, loss of loyalty, and increased marketing costs to counter these effects. An innovative solution to these challenges is found in the implementation of Sentiment Analysis, which utilizes the combined strengths of NLP and Machine Learning (ML) techniques. Sentiment Analysis is extensively applied to extract insights from the public opinion on topics, products, and services, utilizing publicly available online data. In this vein, this research explores the effectiveness of NLP techniques, such as Bag-of-Words and TF-IDF, alongside a variety of ML classification algorithms (Support Vector Machine, Logistic Regression, Multinomial Naive Bayes, Random Forest) to identify the most efficient approach for Sentiment Analysis on a large, imbalanced, and multi-class dataset. Remarkably, this study achieves an accuracy of 77% using Support Vector Machine and Logistic Regression with the Bag-of-Words technique. This related work serves as an illustrative example of how NLP can be leveraged to enhance business strategies by accurately analysing and responding to customer sentiments. The insights from such related work underscore the transformative potential of NLP in business intelligence and customer relationship management. By integrating the lessons learned from these applications of NLP, it becomes possible to not only improve customer satisfaction but also to innovate and streamline decision-making processes across various sectors.

Thus, current research builds upon this foundation, emphasizing the power of NLP to offer solutions that are not just reactive but also proactive in understanding and catering to the nuanced landscape of human language and interaction.

NLP in text classification

Text classification is a foundational aspect of NLP, utilizing both knowledge-based systems and machine learning techniques to effectively decode and categorize textual data. Knowledge-based systems are built on explicit, expert-defined rules, facilitating applications like spam detection through their clarity and simplicity. On the other hand, machine learning approaches – including Logistic Regression, Support Vector Machines (SVM), and Random Forests – are adept at pattern recognition, rendering them ideal for tasks requiring adaptability, such as sentiment analysis and opinion mining.

One example study in the field of NLP focuses on Opinion Mining, specifically targeting Bangla text and employing data from various social media platforms (Taher–Azharul Hasan–Afsana Akhter 2018). This research

leverages both linear and nonlinear SVM configurations, alongside the N-gram method, to classify documents more effectively. Unlike traditional approaches that consider single words as vectors, this work utilizes N-grams, sequences of 'n' words, as a unified vector, leading to better classification results for different 'n' values. This study underscores the utility of combining machine learning techniques with NLP methodologies, such as N-grams, to enhance the processing and understanding of large-scale textual data.

A closely related work (Hansen et al. 2023) examines the shift towards remote work catalysed by the pandemic, analysing over 250 million job vacancy postings across five English-speaking countries. This research employs a sophisticated language-processing framework, refined through 30,000 human classifications, achieving 99% accuracy in identifying postings that offer hybrid or fully remote positions. This notable precision in classifying job postings based on remote work opportunities showcases the practical application of NLP in monitoring significant societal trends.

All these related works illustrate the broad applicability and impact of NLP and machine learning in understanding and interpreting the complexities of human language and social changes. They highlight the progression of NLP systems from basic text categorization to sophisticated analyses capable of uncovering nuanced insights within vast datasets.

NLP methodologies for data analysis

A methodical approach to data analysis is the basis for successful NLP applications. It starts with data extraction and cleansing, removing extraneous elements that could cloud analysis. Subsequent preprocessing stages, including tokenization and lemmatization, further refine the text to its analytical essence. Annotation assigns critical markers to data, paving the way for machine learning algorithms to learn efficiently. Finally, vectorization translates the curated text into a numerical form, laying the foundation for complex computations and classifications.

Synthesizing theory and practice

The theoretical concepts of NLP play an importance role in the development of systems that allow effective classification of texts. These principles form the backbone of practical applications, such as organizing unstructured job advertisements and facilitating the rapid identification of remote work opportunities. The following case study demonstrates how these NLP

methodologies are employed to dissect, understand, and categorize job advertisements. This application of NLP not only illustrates the practical utility of theoretical knowledge, but also demonstrates the transformative potential of NLP in reshaping the job market in the digital age.

Case study: NLP-driven classification of remote jobs

NLP technique implementation

This section outlines the implementation of NLP techniques to identify remote employment opportunities within online job advertisements. The process began with data extraction, where a total of 324 job advertisements for 'full-stack web developer' positions were collected from three major online portals using Octoparse, a user-friendly web scraping tool. This facilitated the rapid acquisition of high-quality, relevant data, which was then prepared and cleaned for analysis.

Python (Van Rossum–Drake 2009) and its related data science modules and libraries served as the main platform of the research. Pandas framework (McKinney 2010) was used for many of the data manipulation tasks, which provided the bases to securely and efficiently transform and send data down the pipeline. The GPT-3.5 Turbo model (OpenAI developer platform, 2023) was employed to correct text inconsistencies, such as merged words, within the job descriptions. This step involved using the OpenAI API for text correction, significantly enhancing the quality of the dataset for subsequent processing stages. Preprocessing included normalization activities like lowercase conversion, punctuation removal, and tokenization, followed by more advanced procedures such as stop word removal and lemmatization (Jurafsky–Martin 2023). These were carried out utilizing the NLTK (Bird, Loper–Klein 2009) Python library which is a standard tool for data scientists to process textual data.

Annotation of the dataset was carefully carried out to tag listings as 'Remote' or 'Not Remote,' using criteria developed in collaboration with project leadership. This binary classification was crucial for the study's focus on remote job identification. Additionally, vectorization of the dataset was performed using both TF-IDF and OpenAI's Ada-002 model embeddings to prepare the data for machine learning analysis, leveraging these methods' ability to capture semantic nuances and importance within the text.

The text processing steps are shown in Figure 1–4.

description
Power Platform Developer Manchester, Ruddington or London – Hybrid working model (2 days office based, 3 days remote) Very competitive day rate Fantastic opportunity to secure and initial 6 months day rate contract inside IR35 as a Power Platform Developer with Smart DCC. The Power Platform Developer role will be working with the Enterprise and End to End Architecture teams, this is a hands-on role working within the DCC’s EIT function. The Power Platform Developer is the DCC subject matter expert for automating many of the DCC’s manual processes. Power Platform Developer is fundamental to the success of reducing aoperation complexity and introducing a standards-based approach to using Microsoft 365 developement platforms. What will you be doing? • Process

Figure 1. Raw extracted text from a job advertisement.

cleaned_descriptipon
Power Platform Developer Manchester, Ruddington or London – Hybrid working model (2 days office based, 3 days remote) Very competitive day rate Fantastic opportunity to secure and initial 6 months day rate contract inside IR35 as a Power Platform Developer with Smart DCC. The Power Platform Developer role will be working with the Enterprise and End to End Architecture teams, this is a hands-on role working within the DCC’s EIT function. The Power Platform Developer in the DCC subjec matter expert for automating many of the DCC’s manual processes. Power Platform Developer is fundamental to the success of reducing operational complexity and introducing a standards-based approach to using Microsoft 365 development platforms.

Figure 2. Text after cleaning using GPT model.

power platform developer manchester ruddington london – hybrid working model 2 day office based 3 day remote competitive day rate fantastic opportunity secure initial 6 month day rate contract inside ir35 power platform developer smart dcc power platform developer role working enterprise end end architecture team handson role working within dcc’s eit function power platform developer dcc subject matter expert automating many dcc’s manual process power platform developer fundamental success reducing operational complexity introducing standardsbases approach using microsoft 365 development platform • process automation reporting ensure business function able replace historic process auditable automation utilising microsoft power platform
--

Figure 3. Text after tokenization.

```

[-0.021912286058068275, -0.0009061856544576585, 0.007242659106850624,
-0.03044510819017887, -0.028424536809325218, 0.024902187287807465,
-0.01160463783890009, -0.02001458778977394, -0.017024686560034752,
-0.03877314180135727, 0.018622029572725296, 0.005689685698598623,
-0.003751028561964631, 0.00536885190901613, -0.02300448715686798,
0.006136805284768343, 0.00948167126657763, -0.015440993942320347,
-0.0078024123795330524, -0.009563586674630642, -0.005484897643327713,
-0.005413222126662731, -0.0020052131731063128, -0.012649054639041424,
-0.029489431530237198, 0.01417131070047617, 0.006850149482488632,
-0.01697007566690445, 0.008539647795259953, -0.003559893299762964,
0.013044977560639381, -0.0199190191924572, 0.009078922681510448,
-0.013946044258773327, -0.0159870944917202, -9.519429295323789e-05,
-0.011229193769395351, 0.015604824759066105, 0.01899064891040325,
-0.01154320128262043, 0.024724705144762993, 0.0017799466149881482,
-0.01915447786450386, -0.009092574939131737, -0.0037203102838248014,
0.019482139497995377, 0.010212081484496593, -0.016260145232081413,
0.0099868150279375, 0.039182718843221664, -0.00016607005090918392,
0.020792780444025993, -0.016533195972442627, -0.02350963093340397,

```

Figure 4. Section of the text after embedding vectorization.

AI method evaluation

The study employed both rule-based and machine-learning classification methods to categorize job advertisements. The rule-based approach utilized tokenized document scoring to differentiate between 'Remote' and 'Not Remote' classifications based on predefined expressions and their accumulated scores. This method, while straightforward, provided a foundational understanding of the dataset's characteristics and the effectiveness of simple linguistic rules in preliminary classification tasks. Table 1 showcases a rule-based classification example where a job advertisement for a Hybris Developer is correctly classified as 'Remote' based on a comparison of tokenized text scores - 15 for 'Remote' and 4 for 'Not remote'. This demonstrates the rule-based system's ability to categorize job advertisements effectively using predefined linguistic expressions.

Table 1. Rule-based classification example

Original text	Tokenized text	Class label	Prediction	Remote score	Not remote score
Hybris Developer (100% remote) We offer : Interesting and challenging job in one of the biggest company in Polish automotive industry A lot of independence in action and the opportunity to implement your own ideas Private medical care Company discounts 100% remote work or if you prefer hybrid work – we have office in convenient location in Warsaw	offer interesting challenging job one biggest company polish automotive industry lot independence action opportunity implement idea private medical care company discount 100 remote work prefer hybrid work office convenient location warsaw	Remote	Remote	15	4

For a more sophisticated analysis, supervised machine learning models were trained using the pre-processed dataset, with both TF-IDF vectors and embedding vectors serving as inputs. For this purpose, Scikit learn (Pedregosa, et al., 2011) Python library was used, which contains the toolset to prepare, train, test and deploy machine learning models. Among the supervised machine learning methods, three classification algorithms were selected based on the literature guidelines. The chosen models proved to be effective for classifying textual data in several cases. The models were as follows:

- Logistic Regression (Uddin, Khan, Hossain & Moni 2019)
- Support Vector Machine (Hassan, Ahamed & Ahmad 2022)
- Random Forest (Occhipinti, Rogers & Angione 2022)

The training process involved hyperparameter tuning with GridSearchCV and ten-fold cross-validation (Gareth, Witten, Hastie & Tibshirani 2013) to optimize model performance and mitigate overfitting. Model evaluation was comprehensive, incorporating metrics such as accuracy, precision, recall, F1-score, and AUC-ROC to assess classification effectiveness rigorously. These metrics are widely used by experts in the field of machine learning (Brown 2018).

Comparative performance analysis

The comparative analysis of the classification methods used in this study underlines the distinct advantages and limitations inherent to both rule-based and machine-learning approaches in the context of remote job advertisement classification. A summary of the model metrics is shown in Table 2.

Table 2. Evaluation of trained models

Model	Accuracy (%)	Recall (%)	Precision (%)	F1 score (%)	ROC AUC (%)
Rule-based	96.92	94.29	100	97.06	N/A
TF-IDF LR	98.46	100	97.22	98.59	100
TF-IDF SVM	76.92	82.86	76.32	79.45	86.95
TF-IDF RF	78.46	74.29	83.87	78.79	87.14
Embedding LR	76.92	74.29	81.25	77.61	86.1
Embedding SVM	76.92	71.43	83.33	76.92	86.57
Embedding RF	81.54	80	84.85	82.35	91.43

Rule-based classification, as a method, offered the benefit of providing immediate insights with relatively minimal computational resources, acting as an effective preliminary filter for identifying remote opportunities. The inherent simplicity of rule-based systems, relying on manually defined rules and expressions, makes them particularly transparent and easy to understand. Legislation has a similar requirement for norm clarity (Arató 2022; Arató–Balázs 2022). However, this reliance also introduces constraints on their adaptability and scalability, as the manual adjustment of rules to accommodate new data or variations in job advertisement formats can be labour-intensive and less dynamic.

In contrast, machine learning models, especially the TF-IDF Logistic Regression method (TF-IDF LR), demonstrated superior performance

in accurately identifying remote job opportunities. The TF-IDF Logistic Regression stood out with an overall accuracy of 98.46% and a perfect recall rate of 100%, showcasing its exceptional reliability. This method's capability to flawlessly differentiate between remote and non-remote listings, as indicated by its perfect ROC AUC score of 100% (Fig. 5), underscores the significant potential of machine learning techniques in enhancing the precision of job classification systems. Such models are adept at handling the nuanced linguistic features present in job descriptions and adapting effectively to the dataset's complexity, offering a scalable solution for the evolving demands of the job market.

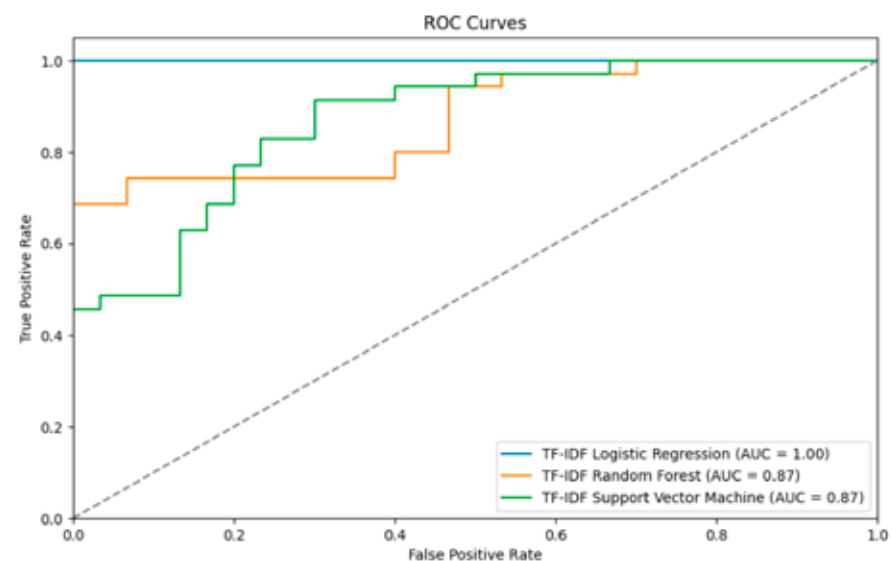


Figure 5. ROC curves of TF-IDF models.

Following closely, the Rule-based classification method demonstrated considerable efficacy with an accuracy of 96.92% and an F1 score of 97.06%. Its precision ensured that identified remote job advertisements were genuinely pertinent, providing a high degree of confidence in the classified results. Despite its slightly lower recall compared to the logistic regression model, the rule-based approach remains valuable in contexts where avoiding false positives is critical, highlighting the importance of selecting the right method based on specific project needs.

The evaluation of these models based on comprehensive performance metrics illuminated the most effective techniques for classifying remote job opportunities, revealing the essential role of AI and NLP in refining the job search process. While machine learning models like TF-IDF Logistic Regression exhibit a clear advantage in terms of accuracy and adaptability, the rule-based approach maintains its relevance through its simplicity and the immediate clarity it provides. The machine learning models utilizing embedding vectors offer a detailed linguistic analysis, as demonstrated by the Embedding Logistic Regression (Embedding LR), which provides a robust framework for understanding

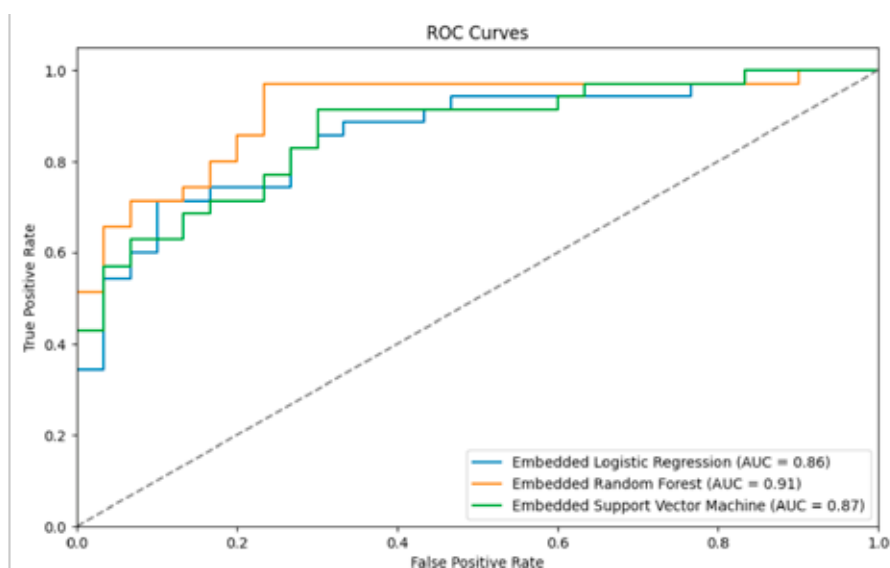


Figure 6. ROC curves of Embedded models.

contextual language in job advertisements. The Embedding Random Forest (Embedding RF) model, with an accuracy of 81.54% and a ROC AUC of 91.43%, highlights the practical application of embeddings in dealing with complex language patterns, marking a substantial advancement from traditional rule-based and TF-IDF methodologies. The analysis reveals that while rule-based and TF-IDF approaches yield higher performance metrics, embedding-based models also demonstrate considerable effectiveness in classifying job advertisements. The comparative advantage of rule-based and TF-IDF methods lies in their higher accuracy and recall rates. However, the

embedding models, with their contextual language processing capabilities, still deliver robust results, indicating their potential for application in more complex linguistic scenarios. This underscores the merit of incorporating a diverse array of methodologies in developing comprehensive classification systems for remote job advertisements.

This analysis confirms the importance of leveraging the strengths of all classification strategies to meet the varying requirements of remote job advertisement identification. As the digital job market continues to expand, the integration of these advanced NLP techniques into classification systems is key, promising a more efficient and accessible way for job seekers and employers alike to navigate the remote work landscape.

Conclusions and future directions

This paper contributes valuable insights into the application of NLP techniques for classifying job descriptions aimed at remote work opportunities, demonstrating that even with a relatively modest dataset, our results are both impressive and encouraging. An investigation of the effectiveness of rule-based classification systems and machine learning methods, particularly the TF-IDF logistic regression model, highlights an efficient path forward in automating the classification process. Remarkably, this model distinguishes itself by delivering high accuracy and reliability, showcasing its potential in settings where minimizing false negatives is essential.

This research reveals an important finding: effective and precise models can be developed without the exhaustive need for large datasets. This aspect of our research is particularly noteworthy as it suggests a strategic advantage in the use of NLP for job classification. By demonstrating the feasibility of achieving significant results with smaller datasets, our approach sheds light on a more accessible and sustainable methodology for model development and deployment. This efficiency not only alleviates the burden of gathering, annotating, and maintaining voluminous data sources but also reduces the computational demands typically associated with training and implementing sophisticated models.

The challenges associated with the embedding methods are due to the scale of the data set relative to the complexity of the vector spaces, which highlight the potential for future improvements and optimisation of the approach. The limitations observed with embedding methods encourage further research and the search for innovative solutions that can circumvent these limitations.

Looking ahead, expanding the dataset size emerges as a promising direction, potentially enabling more refined model tuning and improved generalization capabilities. Furthermore, the exploration of advanced deep learning techniques, such as recurrent neural networks (RNNs) and transformer-based models, could unlock new dimensions in semantic analysis, enhancing the ability of classification systems to interpret job descriptions with greater nuance and accuracy.

In summary, this paper makes a compelling case for the strategic use of NLP techniques in job classification, especially in contexts constrained by dataset size or computational resources.

References

- Arató, Balázs 2022. Norm clarity in the light of Hungarian case law. *Magyar Nyelvőr*, 46 (special issue), 81–90. <https://doi.org/10.38143/Nyr.2022.5.81>
- Arató, Balázs – Balázs, Géza 2022. The linguistic norm and norm of legal language. *Magyar Nyelvőr*, 46 (special issue), 91–103. <https://doi.org/10.38143/Nyr.2022.5.91>
- Bird, S. – Loper, E. – Klein, E. 2009. *Natural Language Processing with Python*. O'Reilly Media Inc.
- Brown, J. 2018. Classifiers and their Metrics Quantified. *Molecular informatics*, 37(1–2), 1700127. <https://doi.org/10.1002/minf.201700127>
- Gareth, J. – Witten, D. – Hastie, T. – Tibshirani, R. 2013. *An Introduction to Statistical Learning: with Applications in R*. New York: Springer.
- Hansen, S. – Lambert, P. J. – Bloom, N. – Davis, S. J. – Sadun, R. – Taska, B. 2023, March. Remote Work across Jobs, Companies, and Space. Cambridge: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w31007>
- Hassan, S. – Ahamed, J. – Ahmad, K. 2022. Analytics of machine learning-based algorithms for text classification. *Sustainable Operations and Computers*, 3(2666–4127), 238–48.
- Jurafsky, D. – Martin, J. H. 2023, January 7. *Speech and Language Processing*. Retrieved from <https://web.stanford.edu/~jurafsky/slp3/ed3book.pdf>
- McKinney, W. 2010. Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*. Austin TX. <https://doi.org/10.25080/Majora-92bf1922-00a>
- Occhipinti, A. – Rogers, L. – Angione, C. 2022. A pipeline and comparative study of 12 machine learning models for text classification. *Expert Systems with Applications*, 201(0957–4174), 117193.
- OpenAI developer platform 2023. Retrieved 10 30, 2023, from <https://platform.openai.com/docs/introduction>.

- Pedregosa, F. – Varoquaux, G. – Gramfort, A. – Michel, V. – Thirion, B. – Grisel, O. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825–30.
- Taher, A. – Azharul Hasan, K. – Afsana Akhter, K. 2018. N-Gram Based Sentiment Mining for Bangla Text Using Support Vector Machine. In *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)* (pp. 1–5). IEEE.
- Tusar, M. – Islam, M. 2021. A Comparative Study of Sentiment Analysis Using NLP and Different Machine Learning Techniques on US Airline Twitter Data. In *2021 International Conference on Electronics, Communications and Information Technology (ICECIT)* (pp. 1–4). IEEE.
- Uddin, S. – Khan, A. – Hossain, M. – Moni, M. 2019. Comparing different supervised machine learning algorithms for disease prediction. *BMC Medical Informatics and Decision Making*, 19(1), 281.
- Van Rossum, G. – Drake, F. L. 2009. *Python 3 Reference Manual*. CreateSpace.

Benjamin Szilágyi

BSc Student

Óbuda University, Hungary,

Alba Regia Technical Faculty

E-mail: szilben1997@gmail.com

<https://orcid.org/0009-0005-9690-4382>

Rozália Lakner

Associate Professor

Óbuda University, Hungary,

Alba Regia Technical Faculty

E-mail: lakner.rozalia@amk.uni-obuda.hu

<https://orcid.org/0000-0001-5665-479X>

BÉLA POKOL

THE LAYERS OF BEING AND THE QUESTIONS OF ROBOT ETHICS

Abstract

The paper seeks to analyze the new ethical dilemmas that arise in the social contexts of the robot world. It is based on the theoretical foundation of the ontology of Nicolai Hartmann, which finds the place of ever-increasing artificial intelligence in reality among the layers of being. From this starting point, it examines the summative studies of the robotics analysis already developed in English and looks at their correction that needs to be made in the theory of four-layered human existence in comparison with the analyzes so far.

Keywords: layers of being, robot ethics, artificial intelligence, Nicolai Hartmann

Human existence and the life of human communities are based on the cumulative regularities of the layers of being that are built upon each other through evolution, according to the theses of Nicolai Hartmann's ontology (Hartmann 1962). The accelerated development and increasing use of artificial intelligence (AI) in recent years in this structure directly affects the top layer of the four (physical, biological, spiritual and intellectual) layers of being, increasing its strength to the detriment of the lower ones. And with the later development of artificial intelligence, eventually breaking away from human control and gaining independence, it can be perceived as an evolutionarily created new layer of being. Unlike the three previous evolutionary leaps, however, it would not require all the lower layers of being. Taking into account the robots that are the physical incarnations of AI today, AI only needs the physical layer of being. (Pokol 2017). Against this theoretical backdrop, the analyses in this study seek to explore the emerging moral and related legal dilemmas within the mechanisms of contemporary societies that are increasingly permeated by artificial intelligence, while at the same time considering the extent to which the analytical framework changes when the multi-layered nature of human lives, and thus society, is constantly kept in mind.

1. The preliminary questions of robot ethics

In his study of the etic problems of the robot world, *Keith Abney* identifies three areas to group the problems: 1) the field of requirements and prohibitions for robot makers and programmers (such as medical ethics);

2) secondly, the field of requirements to be programmed into robots, first formulated by Assimov under the heading “Three Laws of Robotics”; and 3) finally, perspectivistically for the future, the question of the moral demands and “human rights” that robots might have at that time in possession of self-awareness emerges (Abney 2011: 35). A common dilemma for all three areas is the choice between the main starting points of moral theory already elaborated in the various moral philosophical schools of the comprehensive moral philosophical communities. One such school can be identified as the *deontological starting point* (the rule is the rule, and these must be followed), for which Kant’s moral philosophy is best known, and the polarising opposition school, which considers consideration of the *consequences of action as the basis for moral decision-making*. Finally, thirdly, mention can be made of the *virtue ethics* school, which focuses not on the requirements to be considered in every situation in defining morality (like the two previous schools, albeit of the opposite direction), but on the enduring dispositions of the human personality, more simply, on socialised moral values. Here the person does not ask what the moral rule is in a situation, because in the increasingly complex modern world there are often no clear rules, but how a brave, just, faithful, true man decides (Abney 2011: 37).

Of the three schools, the school of deontology is only possible for robots that are used in the narrowest domain and follow the exact rules without being able to weigh the rules, because all situations can only be calculated and controlled in such a narrow domain, but even here unforeseen situations can arise and steer the robot decision in the wrong direction. For example, it could in principle be fed into the decision algorithm of a combat robot to “never kill a child!”. However, in the case of child soldiers in African wars, this would mean a predetermined liquidation of the combat robot (Abney 2011: 42). In the case of general-purpose robots, the deontological approach is completely inapplicable. However, the consequentialist school of moral philosophy, which is also tied to the consideration of individual situations, also seems better only because of its life-like nature. Here the guiding premise is to “increase, not decrease, the happiness of as many people as possible with the chosen decision!” and this is impracticable because it would require the processing of a huge amount of information, most of which could not be done in a timely manner even with the greatest capacity of computer data storage. Keith Abney’s position, therefore, is that with respect to the second area of robotic morality (i.e. moral decision premises programmed into the robotic algorithm), there is a mixture of deontology and virtue ethics that gives the best perspective, and a mixture of these can create the best built-in robotic moral version.

According to this, the more abstract moral norms (moral virtues) form the decision framework, and the built-in goals and decision contexts always specify the determinants of the decision chosen by the robot in the given situations:

“The hybrid approach of hypothetical rather than categorical imperatives (within a deliberately restricted, not universal, frame) coming from virtue ethics appear the best bet for near-term robotic morals (in sense two). [...] The emphasis on being able to perform excellently in a particular role, and the corresponding specificity of the hypothetical imperatives of virtue ethics to the programming goals, restricted contexts, and learning capabilities of non-Kantian autonomous robots, makes virtue ethics a natural choice as the best approach to robot ethics.” (Abeny 2011: 51)

The connections between layers of being and morality are touched on indirectly by Abney, where he opposes emotivism, which identifies morality with moral emotions, and the cognitive perception of morality, which opposes it. It shows that if morality is tied to emotions because of the emotivists’ viewpoint, then primates with emotions cannot be excluded from morality either, which is absurd:

“Such views, in addition to being unable to explain why nonhuman animals lack morality, also have struggled to explain the apparent cognitive meaningfulness of ethical claims and especially ethical disagreement. (They also naturally have severe difficulties accounting for the ethics of emotionless robots)” (Abney 2001: 46).¹

In contrast, he sees the position of evolutionary psychology, which emphasises the new decision-making mechanism of human evolution as an explanation for morality, meaning that humans have increasingly also developed a system of cognitive decision-making that reshapes current decisions in such a way that the instinctual-emotional first step of thinking is always followed by a second cognitive consideration, thus correcting the first:

“Evolutionary psychology’s claim there are not one but two of decision-making systems within most humans. The first is an instinctual, emotionally laden system that serves as the default for much human activity, particularly when stressed or under pressure. Many other animals share this noncognitive

¹ It is worth pointing out that this view held by Hegel in legal philosophy long before the advent of scientific psychology in 1820. See Georg Wilhelm Friedrich Hegel, “Grundlinie der Philosophie des Rechts. (Werke Band 7.) Suhrkamp, Frankfurt am Main. 1979. 301. p.

decision-making system, in which (quite literally) we »know not what we do« – or quite why we do it. [...] But this »ghost in the machine« does not exhaust human agency; Libet and others found we also have a »veto« ability that can, after its subconscious initiation, still alter our action, in accord with a decision by a second, conscious cognitive system” (Abney 2011: 46).

Abney recalls the mutually shaping effect of the two overlapping layers and almost recalls Nicolai Hartmann:

“In humans, this deliberative system overlays the ancestral instinctual, emotional (and faster) decision-making system and so reason is quite often trumped by our instinctual drives” (Abney 2011: 46).

Having concluded that the upper (cognitive-rational) layer of the two-layered human decision-making mechanism is responsible for moral decision-making, Abney poses the question of whether moral decision-making is in principle possible without a reconfigured, overridden lower layer? After all, the answer to this question also depends on whether a moral decision is possible for robots without an emotional layer. In this question, he then decides exactly the opposite way as Hartmann did earlier. It is very possible – he says – that a rational decision-making mechanism is sufficient for moral decision-making even without an emotional layer of being:

“Hence, deliberative system capable of agency necessary for the existence of morality, and so for moral personhood. But is the ancestral emotional system needed as well? [...] In other words – could (emotionless) robots be moral person? [...] The key to moral responsibility and personhood is the possession of moral agency, which requires the capacity for rational deliberation – but not capacity for functional emotional states, therefore, robots may well qualify” (Abney 2011: 47).

Based on Hartmann, there are two problems with these analyses. On the one hand, given the three layers of being above the physical layer of being, it can be seen as flawed that Abney combines biological stimuli with the determinants of the emotional layer. Already here there is building on each other and transformation, and an instinct of the raw instinctual world is supplemented by the emotions of the higher layer of mental existence. For example, the ferocity of a biological sex drive is informed by feelings of togetherness, not to mention the intellectual-symbolic overrides that still build on it, and

the sex-altering aspects of the sublimated love relationships they produce.² That is, it is not a double but a triple decision-making mechanism that must be analytically separated in human decision-making, and in addition to the most basic instinctual reactions and determinants, their emotionally reshaped manifestations are still under more rational considerations at the intellectual level. However, a decision and the instinct that directly determines it, respectively its emotional transformation and its intellectual overwriting, are embedded in the interdependent laws of all three upper layers of being. Thus, human morality in societies of all human civilisation requires, in order to survive as a race, that men and women live together in some form of permanent community in order to have children and be educated. A larger community is necessary for the successful struggle and survival of the struggle with the forces of nature and other groups of people, and within these larger communities they must interact in more or less harmonious relationships to organise common activities. Moral virtues (norms and values) are therefore tailored to and maintained by the laws of the specific physical, biological, spiritual-emotional and intellectual layers of being of humans and their communities, and it is only because of the narrowing of moral theories in recent decades that conscious moral choices have become the focus of moral philosophy. Hegel in the early 1800s or Rudolf von Jhering in the 1870s and then Nicolai Hartmann in the 1920s still saw clearly that each person in his socialisation only takes on the accumulated moral norms and values, virtues of many generations, from which the broader communities are maintained, without which individuals could not be fit to live.

From this follows another problem with Abney's analysis, and that is that moral choice seems to consist only in following norms according to an intellectual-rational calculation, but does not require the lower psycho-emotional layer of being. Moreover, as we have seen, the laws of the biological layer of being and the instinct that imparts this to every human being are important for the moral decision. But also in view of this, it can be said that moral norms, moral virtues exist only in human communities (and thus are socialized in the people of the next generations), because only in this way is possible a lasting and harmonious human existence in human communities defined by all four human layers of being. Thus, if an artificially intelligent being can exist with the spiritual layer of being alone, and at most needs only a physical-mechanical body to have self-consciousness and to exercise conscious activity, or to be able to reproduce itself permanently in time, then

² See Luhmann's work, which analyzes this process historically: Niklas Luhmann, "Liebe als Passion: Zur Codierung von Intimität", Suhrkamp, Frankfurt am Main 1994.

the moral norms of human existence based on the biological-psyhic layers of life have no function. The moral norms would mean only external things for such being. Thus, if such a robotic being can constantly rebuild its program and even its hardware with the Deep Learning algorithms – as it does for the most part today – then the erosion of moral norms, which are external and functionless for it, is almost inevitable. That is, although it is possible to program instructions that mimic emotions into robots, and they can still account for the decision-making aspects (prohibitions, decision priorities) required by moral norms in today's robots, which are still essentially under human control, but when they reach some level of self-learning capability, it may be uncertain whether the inference of those norms will remain. In the distant future (but in the case of exponential progress in even twenty to thirty years), it would be wrong to assume, in the case of robots in the robotic world, freed from human control and self-aware, the survival of the norms of the human world in the robotic world.

2. Operational morality, functional morality and full moral personality

To better analyze the moral dilemmas and problems of the robot world, the three-way division used by *Colin Allen* and *Wendell Wallach* in their joint study seems useful. Based on different degrees of decision autonomy, they denote the degree of *operational morality* for robots that can only perform the actions determined by the programmers who created their algorithm and possibly by their specific users, and fully fed into them. On the other hand are those that have reached the level of *functional morality*, and this means that they choose the specific action in each situation based on the information provided by their sensors among the action frames fed into their algorithm. Finally, the most autonomous level of morality is seen in robots that reach the level of *full moral personality* with the cessation of human influence, although this type cannot be considered probable now and in the near future, but later their creation can be assumed:

“System with very limited autonomy and sensitivity have only ‘operational morality’, meaning that their moral significance is entirely in the hands of designer and users. As machines become more sophisticated, a kind of ‘functional morality’ is possible, where the machines themselves have the capacity for assigning and responding to moral challenges. The creators of functional morality in machines face many constraints due to the limits of present technology. This framework can be compared to the categories of artificial ethical agents described by James Moor (2006: 18) which range from agents

whose actions have ethical impact (implicit ethical agents) to agents that are explicit ethical reasoners (explicit ethical agents.) As does Moor, we emphasize the near-term development of explicit or functional moral agents. However, we do recognize that, at least in theory, artificial agents might eventually attain genuine moral agency with responsibilities and rights, comparable to those of humans” (Allen–Walach 2011: 57–58).

Without going into the possible criticism of whether it is worthwhile to use the degree of morality for robots that have already been fully defined by programmers under the name of operational morality, the robots of functional morality are really interesting in today’s stage of development in the robotic world. Self-driving cars, self-propelled combat robots, and to a lesser extent robotic nurses in elder care and health care facilities that have already achieved this autonomy, have such robots and self-driving cars slowly rolling between us or transporting us (mostly only in Japan and the United States today), and the moral decision problems they raise give practical significance to their analysis. The authors go through the possibilities of choice among the trends in moral theory already seen above, and they see the virtue ethics direction as suitable for creating the functional morality of robots. According to their analysis, the moral values (virtues) fed in this way can give the decision framework, which is clarified by training through neural learning mechanisms, and in this way the more general viewpoints of virtues become practical moral decision factors:

“The virtue-based conception of morality can be traced to Aristotle. Virtues constitute a hybrid between top-down and bottom-up approaches, in that the virtues themselves can be explicitly described (at least to some reasonable approximation), but their acquisition as moral character traits seems essentially to be a bottom-up process. Placing this approach in a computational framework, neural network models provided by connectionism seem especially well suited for training (ro)bots to distinguish right from wrong” (Allen–Walach 2011: 59–60).

Structurally, this is broadly analogous to the decisions that people are used to making in their daily lives, which are based on abstract moral reasoning and adapted to particular situations, and which are driven less consciously than with mere moral sense. But with the important difference that because of the lack of consciousness and self-awareness of today’s advanced robots, the hybrid determinants subtly tuned by programmers (framing virtues plus their training-concretized memory without consciousness) give the more or

less accepted moral norms of today's human societies for the appropriate or approximate decisions. As for the third version, whether it is really possible to theoretically accept robots according to the degree of full moral personality according to human morality, can only be judged skeptically in the case of a robotic world theoretically beyond human control and having achieved full autonomy, according to the above explanation.

3. The devaluation of the physical-biological environment as a moral problem?

The four-layered human existence and the growing weight of the upper, intellectual layer of being and the devaluation of the lower layers have characterized the human evolution so far, but the increasing adoption by robots of the various kinds of work and environmental perceptions will lead in the future to a major change in human socialization, of which the direction of paying attention to reality and turning the details of the real world into experience for him will be largely changed. In a study, *David Zoller* analyzes the increasingly widespread takeover of the work of humans by robots in terms of how this process deteriorates the perception of everyday reality in human consciousness and how the skills and observation capabilities that still exist today are disappearing. The fact that this can already be observed by anyone is enough to recall the telephone numbers already stored in cell phones and thus largely erased from consciousness, or the spatial orientation information that is disappearing from our consciousness due to GPS, and the partial death of this ability. (A recent brain research also found that in case cab drivers in London could claim that the tiny part of the brain in which a group of brain neurons specialized for this purpose to store the vast amount of information about the streets of London disappeared with the spread of GPS, and this brain sector shifted to another function instead).

Zoller brings this problem closer to moral issues by basing moral decision making on the perception of the whole of reality and, in this case, on the formation of human identity and on the detailed perceptual knowledge acquired by an adult from childhood. If future generations socialize themselves since childhood to be surrounded by robots and to have their immediate environment perceived by robots without performing perceptual activities and tasks instead, they will not only be disenfranchised but also lack detailed knowledge of today's adults. As responsible beings, they also cannot grow up to make moral decisions, in other words, they become childish:

“My own argument is premised on the way that skill opens up corners of reality, so to speak, that are inaccessible to the unskilled. [...] The maturity or adulthood we earn by adjusting ourselves to the ‘real world’, of course, has a certain moral and personal appeal: a world of lazy psychological infants is, we might think a worse world on a variety of spectra” (Zoller 2017: 81, 86).

The fact that these realms of reality go beyond our perception, and there comes instead the mechanical information processing of these robots, allows us to adapt in ways that are now unconscious because of this change, and this also shakes our moral identity, Zoller says:

“Given that automating a skilled activity means agreeing that we will exit some niche of perceptual reality, and maybe exit it forever [...]. The more suddenly, broadly, and pervasively we hand our perceptual facility over to the robots, the more likely we will make mistakes and simply ‘lose data’ that were surprisingly integral to our moral and social lives” (Zoller 2017: 86).

While it must be acknowledged that Zoller, in contrast to previous analyses that focus on robots taking over jobs as a unilateral human facilitation – apart from the already discussed socially negative consequences of unemployment (see, e.g., Ford 2014) – went deeper by looking more closely at the change in human perceptual capacity, it must be criticized for unconsciously placing too much emphasis on the layers of being in the physical-biological environment. Looking at Hartmann’s layers of being, this change can be read completely differently. The changes outlined by Zoller do not mean the loss of the perception of the whole reality and the ability to do so, but only the ability to perceive the physical-biological layers of being and to pass them on to robots and software bots. In this way, man’s liberated perceptual abilities and brain sectors can be more reconstructed to process information about his spiritual-emotional layer of being and his intellectual layer of being, respectively. His moral decisions will therefore be made in the future with less physical and biological environmental information – these will be shut down by robots in mechanical processes – and these decisions can instead be based more on the information from the spiritual-emotional and rational-intellectual layers of being. The diminishing importance of the two lower layers of being, and instead the greater expansion of the two upper layers of meaning for human existence, can of course significantly reshape the foundations of our moral decisions and the incentives that play a role in them. For example, the implantation of dozens of body sensors and their connection to information bases collected in the clouds, as well as automatic

diagnosis by robots of health software and automatic activation of specific doses of drugs implanted in the body, may make the alarms provided by pain genes in our cells largely obsolete (see Kelly 2016: 34–56). Prenatal genetic engineering therefore makes it possible to minimize this, and the conditions of painless human life may redefine the moral obligations and incentives involved today. Overall, therefore, we do not share Zoller’s concerns about moral infantilization.

4. Moral dilemmas and responsibilities in hybrid and networked systems

In a study, the authors *Wulf Loh* and *Janina Loh* examined the issues of moral and legal responsibility that arise in currently developed self-driving cars (Loh–Loh 2017: 35–48). They assume that today’s self-driving cars are only at the stage of operational morality, so they do not even achieve functional moral autonomy vis-à-vis their manufacturers and programmers. The authors have taken this position based on a moral decision-making structure developed by *Stephen Darwall*, which is divided into four aspects and aims to separate the aspects of autonomy necessary for moral decision-making. The aspect of autonomy required for the level of overall moral personality is called *personal autonomy*, i.e., the ability to possess and choose between personal values, goals, and ultimate aims in life. *Moral autonomy* is the other aspect, and this means that their values and goals include moral principles and ethical beliefs, and along with these they always consider alternatives when making decisions. These two do not exist in today’s robots, and only humans are capable of such autonomy, but the aspect of *rational autonomy* is already available to robots at the level of functional morality. This means that the robot can weigh reasons of different weights when making a decision. Their algorithm can already enable this by incorporating pure abstract decision frameworks – leaving some freedom – in which the weighting between possible decision directions is done in light of specific data constantly recorded by their sensors, and they decide based on that data. Finally, the fourth aspect of autonomy is *decision autonomy*, and this means the robot’s ability to make decisions not only by external data – continuously concretizing the built-in framework determinants – but also its internal decision priorities without changing them.

Based on the authors’ examples – two types of robots already in use (*Kismer* and *Cog*) – it seems possible to achieve this degree of autonomy based on their self-learning mechanisms, which are integrated into the robot’s algorithm and are no longer externally controlled:

“Cog the first robot that can interact with its surroundings due to its embodiment, might pass as an example of a weak functional responsible agent, since its ability to communicate as well as judgments has been improved over that of Kismet. Even more importantly, Cog’s overall autonomy has evolved, since it includes an “unsupervised learning algorithm” (Loh–Loh 2017: 40).

Since the current algorithm for self-driving cars does not yet include such an unsupervised self-learning mechanism, they are only at the level of operational morality, and this moral and legal responsibility lies entirely with their developers (designers, manufacturers, and programmers) and car dealers or owners, and respectively between the occupants of the car.

But even with this level of technology, self-driving cars already surpass humans, leaving them – and especially their programmers – with moral dilemmas not seen in the case of humans in extraordinary and unexpected driving situations. For example, if within the braking distance directly in front of the car a group of children jumps into the road to retrieve a rolled ball, the driver cannot stop or even brake at that average speed, leaving him or her with no moral or legal responsibility in the dire event. But self-driving automation, which can react much faster, may still have to make a decision if it can’t stop but crashes into a pillar – potentially seriously injuring the car’s occupants – or drives and kills children to avoid doing so. But technical capabilities far beyond humans could create a dozen similar new aspects of moral decision-making for self-driving cars in the future. The authors of the study therefore suggest that a separate ID card will soon be created for owners of self-driving cars, in which the final setting of the car software program, the dilemmas left open by manufacturers, must be decided at the time of purchase, so that moral and legal responsibility for the following can be assumed:

“Since these dilemma situations do not allow for on-the-fly-decisions, the driver will have to take them beforehand. This means that the driver will have to fill out a moral profile of some sort, maybe in the form of a questionnaire, maybe in the sense of a setup program much as with today’s electronic devices. For convenience, it seems plausible that these moral settings can be saved to a sort of electronic identification device, like an electronic key or the driver’s smartphone, assuming that issues of data security can be solved” (Loh–Loh 2017: 46).

The development of networked robots and the gradual becoming of “smart objects” (smartphones, smart TVs, etc.) around us have only recently begun,

and as they expand, the Internet of Things (IoT) will become more and more involved in our lives in the future.

Human-robot hybrid systems are thus expanding to include additional aspects, and this creates another set of moral and legal dilemmas. *Adam Henschke* analyzes these in his new study (Henschke 2017: 229–43). Smart things are widely available through multifunctional smartphones, smart televisions, robotic vacuum cleaners, and semi-self-driving automated cars with a variety of sensors, but even in everyday life in much of the world, these other smart things have been developed that are already beyond the research lab stages and have already reached the homes of high-tech users with small-scale production. These, however, as we have already experienced with smartphones, etc., will proliferate in a few years and their mass use raises new moral and legal dilemmas. One example is the smart refrigerator, which contains food with RFID (radio frequency identification) and thus digitally identified quantity, shelf life, etc., and the smart refrigerator constantly reads this data, detects the depletion of each food quantity, and since it is connected to the web-based sales mechanisms of nearby supermarkets on the Internet, it can order food and other household items to be automatically delivered. In Japan's aging society, an increasingly large amount of elderly people can be cared for through the use of care robots, and in fully digitized smart homes, such a robot can also care for helpless elderly people, taking over ordered food deliveries in this way. By observing and communicating with the helpless elderly person entrusted to its care, it can call the family doctor or, if necessary, the hospital by phone if its built-in algorithms make a more serious health problem likely.

This example shows how, in a decade or two, robots will be needed in more and more parts of the world, to solve more and more of the work through omnipotent robots and smartphones that can be used in comprehensive information systems to fulfil their functions. However, this growing indispensability of the Internet of Things also creates new dangers and moral dilemmas compared to simple robots. Adam Henschke points out in his writing that the novelty of the Internet of Things compared to single robots is that the latter mainly raises the problem of physical security and the risks have to be assessed in this dimension. (E.g. a robot hoover recently inflicted serious injuries on an unexpected occupant, but one or two fatal accidents of self-driving Tesla cars can also be cited for this). In contrast, security problems and dangers in the Internet of Things occur in two different dimensions. Here, in addition to physical security, information security issues also play a role, since the aforementioned elderly care robot, which

is connected to the software of hospitals, doctors and other places on the internet, can provide hackers or others with information about data recorded by its built-in camera and other sensors. They can share the continuously collected health data about an elderly caregiver not only to the software of the hospital in charge, but also to those who make malicious intentions and plans. In the same way, our smart TVs with a range of applications can not only fulfil their convenience but, with their built-in cameras and microphones, transmit the entire life of the home to software and information databases that we do not see.

This vulnerability can also lead to a physical vulnerability, such as when a hacked automatic door lock is opened remotely for an intruder by external instructions from smart devices. Or, as has already happened in an elegant beach hotel, the electronic smart locks were blocked by a criminal group from outside the flats, and the hotel guests of the wealthy elite were prisoners until the required ransom was paid. However, Henschke also mentions the possibility of the electronic lock of a billionaire's car being blocked by criminals after he got out and his trapped children in it being released on a sunny day only if he transferred hundreds of thousands (Henschke 2017: 234). Immediately after the incident, the said elegant hotel replaced the electric locks that could be swept from the outside and reinstalled the good old traditional locks. After such an incident, the said billionaire will probably also restrict the internet functions of his car for a while. All this, however, forces choices in moral and legal dilemmas and elections that are broadly worth pondering. In the world of our objects, which is becoming more prevalent in the Internet of Things, the old simple things are already being dropped, and we will not be able to replace the objects we wave into the cloud database at will. Just as we would not give up the internet today despite all the negative aspects that vulnerability brings.

One such dilemma of the Internet of Things embedded in networked and comprehensive cloud databases is which of the conflicting requirements of the two types of security – physical security and information security – should be given priority? For example, making the smart home of an elderly person who is barely able to move around fully remotely monitorable by medical centres through cameras and microphones may be important to some extent, but it may also mean exposing the most intimate manifestations of life beyond what is necessary. If the emphasis is on information autonomy and limited observation and transparency, the information that is still needed in rare cases may not be passed on to the care centre, and the elderly care recipient may die. Henschke points out that there are often typical priorities,

and for example, in a smart TV, information security has a higher priority, and for this purpose, we can easily address the constraints here. However, with thousands of self-driving car applications tied to cloud software, we pay more attention to physical security requirements and only secondarily to information security requirements (Henschke 2017: 239).

5. Self-learning, machine learning and responsibility

As mentioned above, the main problem of the future will be the dilemma of self-driving cars that are detached from humans and can no longer be blocked from the outside in certain unexpected situations, when the algorithm of this self-driving car, built on neural self-learning, has already decided autonomously. As this has been the main direction of artificial intelligence development in recent years, it is almost certain that this will not be circumvented in this area either. Therefore, today it is worth taking a closer look at the dilemmas of moral and legal responsibility of robots with a high degree of neural self-learning and their makers, owners and users. This question is addressed in their joint study by *Trevor N. White* and *Seth D. Baum* (White–Baum 2017: 66–79) and by *Shannon Vallor* and *George A. Bekey* (Valor–Bekey 2017: 338–53) analysed from different angles.

Trevor and Baum's study not only considers designers, builders and users, but also takes into account the 'punishment' of the robot itself in the case of advanced robots, which already have a punishment and reward system built into their programming, and repeated punishments and rewards reinforce in their programming the decision directions (positive or negative) regarding the selection of future robot responses. This also integrates punishment/reward into the learning algorithm. When the situation arises in the future, the robot's decisions are encouraged in the right direction, and the robot does not need to have consciousness and self-awareness to do this. This way of reinforcing self-learning through repetition is also acceptable according to the authors:

“Non-conscious robots could conceivably be punished with some sort of reduced reward or utility as per whatever reward/utility function they might have. Specifically, they could be reprogrammed, deactivated, or destroyed or put into what is known as a 'Box': digital solitary confinement restricting an AI's ability to communicate or function. To make this possible, however, such robots ought to be based (at least in part) on reinforcement learning or similar computing paradigms (except ones based on neural network algorithms)” (Trevor–Baum 2017: 71).

The neural learning system, however, is judged by the authors to be such that designers and programmers already lose control over the robot's reaction to a given situation and should therefore be banned from the outset as a potential source of danger, possibly disaster:

"Designers could be similarly liable for building robots using opaque algorithms, such as neural networks and related deep-learning methods, in which it is difficult to predict whether the robot will cause harm." (ibid.)

In the case of algorithms that allow such opaque robot behaviour, it is no longer enough to prescribe liability after the fact, but the prescriptive prohibition is the appropriate thing to do:

"Hence, instead of liability, a precautionary approach could be used. This would set a default policy of disallowing any activity with any remote chance of causing catastrophe. In effect, people would be held liable not for causing catastrophe but taking actions that could cause catastrophe" (Trevor-Baum 2017: 74).

If one agrees in principle with the authors on the dangerous character of neural deep learning software mechanisms, it only needs to be reiterated that this is a ban on the main way to develop artificial intelligence and therefore it should be considered unlikely in the light of the powers behind industry, military, etc. Therefore, it seems advisable to look for other paths that try to find other solution without banning neural deep learning.

It should of course be emphasised that neural network learning, which mimics the functioning of the central nervous system, can be controlled by involving external human control before a reality-building effect can be triggered. However, this is increasingly falling short for a number of reasons, and this is analysed by Vallor and Bekey in the study cited earlier. One reason is that the advantage of using artificial intelligence instead of humans, the incredibly fast responsiveness would be lost if retrospective human control were introduced. Moreover, ninety-nine percent of the time, the responses are correct, many times higher than human performance. Moreover, the quality of much slower human control may be questionable, as the robot's decision may be more correct than the superior human decision. The latter happened with IBM Watson's drug diagnosis algorithm, and the unusual cure highlighted by artificial intelligence from millions of oncology studies and diagnoses, and later synthesised by them, proved more correct than the oncology decision it overrides:

“Watson’s diagnoses and treatment plans are still vetoed by licensed oncologists. Still, how reliably can a human expert distinguish between a novel, unexpected treatment recommendation by Watson that might save a patient’s life – something that has reportedly already happened in Japan – and the oncological equivalent of ‘Toronto’?” (Vallor–Bekey 2017: 343).³

The dilemma of losing speed and thus eliminating the robot’s advantage is also and not even the weakest competitor would have been lost. So it has become a symbol of wrong decisions made by artificial intelligence, which is rare but causes tragedy in many cases illustrated by the robotic soldiers and decision-making software used in war situations. Here, the question constantly arises whether the robot soldier entering the most dangerous area and building can use the destructive weapons in its possession to destroy those inside without an external human decision, or instead the order of destruction may only be given by remote human confirmation. In the same way, the dilemma arises as to whether a reconnaissance aircraft could be destroyed immediately by the robotic aircraft, or this could only be done with human intervention from the remote command room based on the information it transmits? The constraint of speed requires the robot itself to make and execute a decision, as the loss of time from an external human decision could lead to robot destruction if it broke into a dangerous location. But shooting down friendly fighting machines that have already happened several times, or killing children and women misidentified as enemies in the invaded area, argues against this (Vallor–Bekey 2017: 349).

Detailed neural network learning algorithms, the latest trend in artificial intelligence, already provide self-learning software with multiple depths for the simple computational starting position by continuously feeding in billions and billions of data, thus gradually making the starting position the most advanced. In this technique, between the inputs of the self-learning software and the task-specialised outputs, thousands of intermediate neural layers are found between the massive data, independently finding patterns and regularities and highlighting them for use. By combining billions of pieces of data, they can highlight and use the smallest regularities that are imperceptible to humans when making decisions:

³ “Toronto Mistake” was one of Watson’s fundamental mistakes in a nationwide television quiz when he beat everyone with his answers to the toughest questions. As a final mistake, he made Toronto one of the US cities and not even the weakest competitor would have been lost. So it has become a symbol of wrong decisions made by artificial intelligence, which is rare but causes tragedy in many cases.

“Between the input and output node layers are »hidde«, layers of nodes that function to process the input data, for example, by extracting features that are especially relevant to the desired outputs. Connections between the nodes have numerical ‘weights’ that can be modified with the help of a learning algorithm; the algorithm allows the network to be ‘trained’ with each new input pattern until the network is optimised. [...] The interest in neural network has grown in recent years with the addition of more hidden layers giving depth to such network, as well as feedback or recurrent layers. The adjustment of the connections strengths in these more complex networks belongs to a loosely defined group of techniques known as deep learning” (Vallor–Bekey 2017: 341).

The effects of decision patterns highlighted by these detailed learning algorithms, while often achieving surprisingly good results in practice, may not be understood by designers and programmers, and their decisions may consistently cause surprises, including varying degrees of unpleasant surprises. Who should bear the legal and moral responsibility for this?

6. Identity in the world of artificial intelligence

James DiGiovanna raises thought-provoking questions when he asks in his study how the identity of people with brain implants may change. He discusses this issue together with the question of the identity of robots, which in the future may appear as fully artificial beings and already have self-awareness (DiGiovanna 2017: 307–21). Let us consider the two problem areas separately.

The possibility of memory augmented by brain implants has been developed in recent years in mouse experiments and has been shown to be effective. All hope to mitigate and cure the effects of rapidly spreading Alzheimer’s disease in ageing societies (see Kaku 2014: 132–33). DiGiovanna is exploring the possibility of other developments in the coming years, in addition to diseases spreading en masse, to increase brain capacity. And if a technical solution is found to the problems that remain in this field today, it is almost certain that this will become commonplace, first among the elite, then in society as a whole, to increase the greatest value of human intelligence. This means, however, that the permanent identity of each individual, which is the basis for contacts in communities, may be more or less annulled and it may become uncertain how much we can expect our partners to survive their qualities that we have known and loved so far:

“The ability to rewrite mental content such as ethical values, the capacity for empathy, and general personality traits undermines personhood. [...] A para-person that could experiment with worldviews, completely adopting and deleting values systems, preferences, and bases for judgement, would be largely lacking in what is commonly understood as the most basic element of personal identity” (DiGiovanna, 2017: 311).

This was the basis of our choice in the case of our friends, wife and girlfriend, but in the same way, our closer human relationship with some of our work colleagues is based on the love of their traits, while the relationship with others is only coldly collegial. Therefore, our lives in society and in various small communities within our society are based on our permanent identities, and this can change fundamentally after the addition of heart, hearing and other physical enhancements when the brain is changed with brain implant.

With the gradual changes in one's life, one's consciousness is always being rebuilt in detail, and this leads to small gradual changes in one's identity, which in the modern world are intensified by the information expansions of the last century. In comparison, however, in the future we will be able to plant a whole range of information – the contents of books and studies, smaller libraries – with brain implants in our heads and together with it handle new basic logical and value-processing mechanisms that we did not have in our lives before, and we did not have the skills to do so. Now, this will fundamentally affect the contact between the individual and his communities. In any case, the bases of contact based on the present permanent identities could be eliminated by this change. After such a new content of consciousness – especially if the values of the contact partners have been supplemented and reclassified – I cannot know to what extent my boyfriend, my girlfriend, my wife, my colleague, etc. are the same. Whether the qualities we have loved in them so far are still alive, or in the same way the experiences we have had together so far, which provided the same response in our close relationship even without words, are still relevant to him. This can only be exacerbated by the possibility that the brain implants in the brain, which complement our knowledge, norms and logical abilities acquired with the help of our biological brain, are constantly being updated anew and anew from the outside, as we already know today. Moreover, they can constantly connect to the information bases of their software stored in the clouds. To what extent will our friend, equipped with such, remain familiar, on whom we can rely, because “yet we know him!”?!

This question of identity also extends to legal and moral problems. To what extent can I respect someone for their past behaviour or simply despise

them because after a brain refreshment they can either be a “moral athlete” or just a cold advantage-seeker. Or does legal responsibility for yesterday’s action make sense for someone who has since thought and acted differently? The other side of this is whether, if we can change the consciousness of a sociopath by brain implant and their consciousness is partially erased and a new socially friendly consciousness is introduced, is a system of punishment still necessary? And this raises the question of whether, in addition to voluntary brain implant, its forced installation is acceptable? Or, in part, can it be made compulsory by the state for all children to be screened and tested in childhood, as is currently the case with compulsory vaccinations? DiGiovanna calls para-persons the future humans with such augmented brains – avoiding the name cyborg, which has already been invented for them in science fiction – and given the current state of laboratory research, this future does not mean a distant future at all, and the probability of its realisation is high. Dealing with the legal and moral dilemmas and adapting today’s solutions to the situation of the time therefore require extensive consideration.

Beyond para-persons, in the case of fully artificial and, unlike today, self-aware robotic beings, the likelihood of which cannot be ruled out, even if it is not as great as the former, the question of identity can be addressed by raising new aspects. DiGiovanna places the content of identity at the centre in order to expose the dilemma of identity in relation to robotic beings. Some details of humans and their consciousness are constantly changing, but their enduring characteristics and value preferences change only slightly, even over many years, so that those who always live in their environment can more or less be ascribed an identity that embraces change. It is the slowness of change that enables me, even in today’s fast-moving world, not to be disappointed by my previous experiences with the motivations and characteristics of those who come into contact with me. But that is exactly what disappears for robots, which are thousands and millions of times faster than humans at processing information and learning themselves in a very short time:

“Slow change of character and appearance is part of what makes personal identity [...] But with an artificial person, sudden and radical change in both the physical and mental becomes possible” (DiGiovanna 2017: 311, 307).

Permanent value preferences in information processing and cooperation based on them are already problematic with robots due to the mass and speed of their information acquisition as well as their constant self-learning and self-changing. DiGiovanna’s proposal may also mean that the possibility of “self-awareness” and “ego-awareness” need to be reconsidered even for a future powerful MI

robot. These presuppose the permanent identity of a person, but this relies on the slowness of our changes in consciousness and thus the permanence of our information processing. When an artificial being is freed from human control and switched to independent information processing and from that to self-learning and self-transformation, it can learn thousands of times a day, every hour and even every minute, and can transform itself in its ever-shorter new cycles, then what we call a stable self-consciousness, ego-consciousness, in contemporary humans almost disappears. With this emphasis, DiGiovanna also adds a new question to the much discussed question, i.e. how the question of self-consciousness and ego-consciousness of the future advanced robot consciousness will stand. And how can one imagine moral value without permanent self-consciousness and ego-consciousness?

For this reason alone, the train of thought calculated by the mechanical extension of the current human image that such a robotic being will probably also be “super-ethical” in the case of superintelligence must be considered wrong (see Petersen 2017). In this context, however, it is also necessary to address more comprehensively the explanations and analyses that, in the case of the development of robots with their own consciousness – by human analogy – provide for the recognition of their moral needs and the granting of human rights in their writings. For these analyses conceive of future robots simply as a new kind of human companion and an extension of human existence. Once the robots’ programmes have incorporated emotions into their algorithms, these analyses demand that society pay attention to robots’ emotions and grant them human rights as well:

“It probably needs to be legislated how much pain and danger a robot can be exposed to. [...] It could easily be that this would lead to further ethical debates about other rights of robots. Can robots own property? What happens if someone is accidentally injured? Can they be sued or punished? Who is responsible for them if they are sued? Can a robot own another robot? Such questions then give rise to another question: should robots be given an ethical purpose?” (Kaku 2014: 251).

Our previous explanations answered several questions from these, based on the robotics studies conducted in the intervening period, but the basic problem behind them should also be highlighted, as whole studies and volumes have emerged from similar assumptions, e.g. a new volume in this area edited by *Jason P. Doherty*: “AI Civil Rights: Addressing Artificial Intelligence and Robot Rights.”

Now, the basic problem with this line of thought is that it ignores the fact that rights and ethical requirements can only arise in robots when ego-

consciousness and self-consciousness are created. But it also means that if this really happens in the future, they will simultaneously be freed from human control by the thousandfold development and built up as a separate new layer of being above the previous four layers of being of human societies. From that time on, however, they would be indifferent to the whole biological sphere and the human societies connected with it and would not need “judicial protection”. That is, a robotic world that reached this level would not be part of human society as a “new comrade” in dominion over the world, but as human existence emerged from the primate world and rose above the animal-biological layer of being and became more and more autonomous, so now the artificial machine intellect, detached from biological preconditions, rises above human society. In contrast to the previous construction of ever newer layers of being on the lower layers of being, the new layer of being of artificial intelligence would only need the lowest physical layer of being, and for it the biological and psychological-emotional layers of being would not be necessary. These robotic beings would not need rights and ethical demands, but they will dominate the whole reality, including human societies, as we humans dominate the four-layered earthly world today. In this way, those analyses are rather right that discuss whether, if the robot world really reaches this level, what will happen to humanity?!

7. The moral credo of “Unabomber”, the “mad mathematician”

In the mid-1990s, after many years of a series of bombings and an FBI chase, a secret perpetrator called “*Unabomber*” gave the reason for his actions in a one-and-a-half-hundred- page pamphlet that he spoke out against the inhumanity of the development of technological society since the Industrial Revolution. His peculiar language was recognised by his brother and by notifying the FBI, the long-suffering bomber was captured. It turned out to be *Theodor John Kaczynski*, a mathematician from Harvard. At one point in his university career, he became the enemy of a society dominated by technology and began his series of explosions, targeting the developers and major users of that technology. Several died and more were wounded in the process, and he planned to retaliate even more if they had not been arrested.

Now that the exponentially evolving impact of technology development over the last thirty years has become truly indisputable, and the scale and impact of its further acceleration has already been the subject of several comprehensive analyses, it is worth refocusing on arguments made by Unabomber, the “mad mathematician”. This is what *Jai Galliot* does in his new study. He places the resistance fighter, who has since been busy developing his theses in his prison cell, among theorists and movements of

antitechnology, and an attempt is made to highlight his main theses in light of the current state of the robot world (Galliot 2017: 369–85).

Kaczynski has only drawn the practical conclusions of the earlier theses of Jacques Ellul's 1964 volume *The Technological Society*, which in their own way were also a continuation of Oswald Spengler's 1922 work analysing the decline of Western civilisation. Both authors explained the decline in terms of technological development (Spengler, 1995). The purely pessimistic and resigned tone in Spengler and Ellul then became a moral resistance in the case of Kaczynski, and after seeing that there was no way to reform this development, he believed that only revolutionary violence remained to prevent the destruction of humanity. Decades after his pamphlet, it is now worth considering how the current state of the robotic world and the more radical changes that are already largely visible could mean the endangerment of humanity, or at least a significant deterioration in its condition.

As a starting point for their approach to technological society, it is worth highlighting that both Spengler and Ellul and Kaczynski view human existence as embedded in the physical-biological environment. From this they conclude that human existence is destroyed when, as a result of the industrial revolution, human life becomes more and more technologically mediated and in this way more and more distant from the physical-biological environment:

“Ellul wrote that the machine trends not only to create a new human environment, but also to modify man's very essence and that the milieu in which he lives is no longer his. He must adopt himself, as though the world were new, to a universe for which he was not created. Kaczynski shares this sentiment” (Galliot 2017: 373).

On the other hand, if we take into account Hartmann's thesis, which keeps in mind the four interdependent layers of being of human life (physical, biological, mental and intellectual) and which assumes in the course of evolution the ever stronger transforming effect of the upper layers on the lower layers, the above thesis is exaggerated and without reason and it must be classified as too pessimistic.

Kaczynski and his predecessors see it as a decay of human life when the top intellectual layer of the four strata of being becomes ever more dominant over the lower ones. However, this has been the case, albeit more slowly, over the last two or three thousand years and one can single out the use of metals and especially iron, from which the transformation of the human environment was fundamentally altered. The industrial revolution has only accelerated this, and especially since the 1950s it has become tumultuous to base the various

activities of human communities on intelligence and the technology associated with it. That is, human life is by no means based only on the physical-biological layers of being. So when their share and decisive power in human life diminish and this environment is widely mediated and transformed technologically, it does not mean that human society is destroyed. In all this, only the weight of importance of the four layers of being in reality shifts, making human life more based on the intellectual layer and radically increasing the dominance of this spiritual layer of being over the lower one. This assessment of ours could only be suspended if at some point in the evolution of the robotic world this world were to truly emerge from human control and artificial intelligence were to rise as a new layer of being above the human societies that had hitherto been at the peak of evolution. The elevation of Kaczynski as a prophet would then only be prevented by the fact that under such circumstances and its dangers, the failure to be elevated as a hero would be the least of the problems. However, to the best of our knowledge and belief, this can only be considered an unlikely option today, and rather the growth of human societies characterised by artificial intelligence without a new autonomous layer of being can be considered a realistic vision for the future.

Bibliography

- Abney, Keith 2011. "Robotics, Ethical Theory and Metaethics: A Guide for the Perplexed. In: Patrick Lin – Keith Abney – George A. Bekey (eds.): *Robotethics*. The MIT Press, Cambridge Massachusetts. London. pp. 35–54.
- Allen, Collin – Wendell Wallach 2011. "Moral Machines: Contradiction in Term or Abdication of Human Responsibility?". In: Patrick Lin – Keith Abney – George A. Bekey (eds.): *Robotethics*. The MIT Press, Cambridge Massachusetts. London. pp. 55–68.
- DiGiovanna, James 2017. "Artificial Identity". In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0*. Oxford University Press. New York. pp. 307–21.
- Doherty, Jason P. (ed.) 2016. *AI Civil Rights: Addressing Artificial Intelligence and Robot Rights*. Kindle Edition. Ford, Martin, *The Rise of Robots: Technology and the Threat of a Jobless Future*. Basic Books.
- Galliot, Jai 2017. "The Unabomber on Robots". In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.), *Robot Ethics 2.0*. Oxford University Press. New York. pp. 369–85.
- Hartmann, Nicolai 1962. *Das Problem des geistigen Seins. Zur Grundlegung der Geschichtsphilosophie und der Geisteswissenschaften*. Walter de Gruyter Verlag. Berlin.
- Hegel, Georg Wilhelm Friedrich 1979. *Grundlinie der Philosophie des Rechts*. (Werke Band 7.) Suhrkamp, Frankfurt am Main. S. 301.).
- Henschke, Adam 2017. "The Internet of Things and Dual Layers of Ethical Concern" In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0*. Oxford

- University Press. New York. pp. 229–43.
<https://doi.org/10.1093/oso/9780190652951.003.0015>
- Kaku, Michio 2014. *The future of the mind. The Scietific quest to understand, enhance and empower the mind.* Doubleday. New York.
- Kelly, Kevin 2014. *The Inevitable: Understanding the 12 Technological Forces That Shape Our Future.* Penguin Books.
- Klinewicz, Michal 2017. “Challenges to Engineering Moral Reasoners” In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0.* Oxford University Press, New York. pp. 244–57.
- Loh, Wulf – Janina Loh 2017. Autonomy and Responsibility in Hybrid System. In: Patrick Lin – Ryan Jenkins, Keith Abney (eds.): *Robot Ethics 2.0.* Oxford University Press, New York. pp. 35–50.
<https://doi.org/10.1093/oso/9780190652951.003.0003>
- Luhmann, Niklas 1994. *Liebe als Passion: Zur Codierung von Intimität, Suhrkamp.* Frankfurt am Main.
- Pokol, Béla 2013. *Theoretische Soziologie und Rechtstheorie.* Kritik und Korrigierung der Theorie von Niklas Luhmann. Passau. Schenk Verlag.
- Pokol Béla 2018. *Künstliche Intelligenz: Die Entstehung einer neuen Seinschicht? (Ki – im Spiegel von Nicolai Hartmanns Ontologie.)* Pázmány Law Working Papers No.2018/12. <https://doi.org/10.2139/ssrn.3225107>
- Splengler, Osvald 2007. *Der Ubntergang des Abendlandes.* Albatros Verlag/Patmos Verlag.
- Talbot, Brian – Ryan Jenkins – Duncan Purves 2017. “When Robots Should Do the Wrong Thing” In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0.* Oxford University Press. New York. pp. 258–73.
- Vallor, Shannon – George A. Bekey 2017. “Artificial Intelligfnce and the Ethics of Self-Learning Robots”. In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0.* Oxford University Press. New York. pp. 338–53.
<https://doi.org/10.1093/oso/9780190652951.003.0022>
- White, Trevor N. – Seth D. Baum 2017. “Liability for Present and Future Robotics Technology”. In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0.* Oxford University Press. New York. pp. 66–79.
<https://doi.org/10.1093/oso/9780190652951.003.0005>
- Zoller, David 2017. “Skilled Perception, Authenticity, and the Case Against Automation”. In: Patrick Lin – Ryan Jenkins – Keith Abney (eds.): *Robot Ethics 2.0.* Oxford University Press. New York. pp. 80–92.
<https://doi.org/10.1093/oso/9780190652951.003.0006>

Béla Pokol

Professor Emeritus

University Eötvös Lóránd

E-mail: belapokol1@gmail.com

<https://orcid.org/0000-0002-1170-1764>