
**ARTIFICIAL INTELLIGENCE FROM A LINGUISTIC PERSPECTIVE –
OPPORTUNITIES, FUNCTIONS AND CHALLENGES IN NATURAL LANGUAGE
PROCESSING**

Biljana Ivanovska

Faculty of Philology, Goce Delcev University, Stip, North Macedonia,

biljana.ivanovska@ugd.edu.mk

ORCID ID: 0000-0002-6528-699X

Abstract: Artificial intelligence (AI) represents one of the most significant technological achievements of the twenty-first century. Its rapid development has created new opportunities while simultaneously raising numerous linguistic, communicative, and interdisciplinary questions. This paper examines the role of systems based on Large Language Models (LLMs) in natural language processing, with particular emphasis on their ability to analyze, interpret, and generate linguistic structures. The theoretical section explores the relationship between linguistic form, meaning, and communicative context, as well as the extent to which these systems can develop competencies comparable to human linguistic and pragmatic competence. Particular attention is devoted to the capabilities of contemporary models, such as ChatGPT, Gemini, and Copilot, in processing syntactic, semantic, pragmatic, and discourse-related aspects of language. Furthermore, the paper proposes a systematic framework for classifying their functions based on two criteria: (1) linguistic-communicative function and (2) the relationship to knowledge and information sources. Within the first criterion, four principal categories are identified: text transformation, content generation, analysis and categorization of linguistic data, and interaction and assistance through intelligent dialogue systems. The analysis demonstrates that contemporary AI tools are capable of producing coherent and grammatically accurate texts and support a wide range of language-related applications, including machine translation, text summarization, paraphrasing, and information extraction. Nevertheless, fundamental challenges remain concerning genuine semantic understanding, implicit meanings, speech acts, culturally conditioned patterns of communication, and context-sensitive interpretation. The paper emphasizes the need for a critical evaluation of both the capabilities and the limitations of artificial intelligence, as well as its application in contemporary linguistic research, education, and academic communication.

Keywords: artificial intelligence, Large Language Models (LLMs), natural language processing, pragmatic competence, linguistic-communicative functions, academic communication.

**ВЕШТАЧКАТА ИНТЕЛИГЕНЦИЈА ОД ЛИНГВИСТИЧКА ПЕРСПЕКТИВА-
МОЖНОСТИ, ФУНКЦИИ И ПРЕДИЗВИЦИ ВО ОБРАБОТКАТА
НА ПРИРОДНИОТ ЈАЗИК**

Билјана Ивановска

Филолошки факултет, Универзитет „Гоце Делчев“, Штип, С. Македонија,

biljana.ivanovska@ugd.edu.mk

ORCID ID: 0000-0002-6528-699X

Апстракт: Вештачката интелигенција претставува едно од најзначајните технолошки достигнувања на 21 век, чиј забрзан развој отвора нови можности, но и бројни прашања од лингвистичка, комуникациска и интердисциплинарна перспектива. Овој труд ја анализира улогата на системите базирани на големи јазични модели (*Large Language Models – LLMs*) во обработката на природниот јазик, со посебен акцент на нивната способност за анализа, интерпретација и генерирање јазични структури. Во теоретскиот дел се разгледува односот помеѓу јазичната форма, значењето и комуникативниот контекст, како и прашањето до кој степен овие системи можат да развијат компетенции споредливи со човечката јазична и прагматичка компетенција. Посебно внимание се посветува на можностите на современите модели, како ChatGPT, Gemini и Copilot, во обработката на синтаксички, семантички, прагматички и дискурсни аспекти на јазикот. Истовремено се предлага систематска рамка за класификација на нивните функции врз основа на два критериума: јазично-комуникативната функција и односот кон знаењето и изворите. Во рамките на првиот критериум се издвојуваат четири основни категории: трансформација на текст, генерирање нови содржини, анализа и категоризација на јазични податоци, како и интеракција и асистенција преку интелигентни дијалогски системи. Анализата покажува дека современите алатки со вештачка интелигенција успешно продуцираат кохерентен и граматички коректен текст и овозможуваат широк спектар на јазични апликации, вклучувајќи машински превод, сумирање, парафразирање и

екстракција на информации. Сепак, остануваат отворени прашањата поврзани со вистинското разбирање на значењето, имплицитните пораки, говорните чинови, културно-условените модели на комуникација и интерпретацијата на контекстот. Трудот ја нагласува потребата од критичка евалуација на можностите и ограничувањата на вештачката интелигенција и нејзината примена во современите лингвистички истражувања, образованието и академската комуникација.

Клучни зборови: Вештачка интелигенција, големи јазични модели (LLM), обработка на природен јазик, прагматичка компетенција, јазично-комуникативни функции, академска комуникација.

1. КАКО ЛУЃЕТО КОМУНИЦИРААТ МЕЃУ СЕБЕ?

Пред да се осврнам на прашањето што е „вештачка интелигенција“, би сакала да започнам од фактот како луѓето меѓу себе комуницираат: Кои се сознанијата од лингвистиката и сродните дисциплини за тоа како луѓето зборуваат и општат меѓу себе?

Јазикот понекогаш се смета за исклучително креативен инструмент за пренесување пораки. Тоа несомнено не е сосема погрешно, но сепак е премногу поедноставено. Бидејќи јазикот е одбележан со рутини и формули, а со тоа е и многу предвидлив: Ова се гледа, на пример, кај видовите текстови кои можеме да ги препознаеме врз основа на само неколку типични (шаблонски) елементи. Такви елементи се јазичните формули, како *Почитувани дами и господа...* (обраќање во писмо или говор) или *повремено ќе има...* (најава за временска прогноза) и сл. Овие јазични елементи често се јавуваат и во комбинација со пропратни јазични информации како што се акцентот и интонацијата, како на пример: *ја знаеш ли ОВАА?* (со значење на одредена шега). Притоа, јазикот е вграден во медиумската и мултимодалната средина: Секвенцата на поздравување е придружена со невербални знаци како контакт со очи, мавтање (од далечина), ракување, поклонување итн. — сложена комбинација од различни знаци што сочинува еден „комуникативен жанр“. Знаците се употребуваат рутински и решаваат социјален проблем: менаџмент на контактот, статусот на говорителите, одредување на позицијата на авторитет на соговорникот, степенот на наметнување на еден од говорителите и итн.

Зборовите се сфаќаат како симболи кои имаат значење што е конвенционално утврдено. Во комбинација на различни зборови се добива значење на реченицата, кое понекогаш отстапува од поединечното значење или го конкретизира (значењето на *слеп* како одделна лексема во споредба со *има слеп патник во возот* /во поширок контекст-некој што патува без билет/, или пак, со значење *тој е слепец*/не го бидува за ништо/). Оваа перспектива кон јазикот е многу корисна кога треба да се анализира зошто една реченица не изгледа „добро оформена“, односно ги прекршува правилата, станува забележителна и отстапува од нормата. Исто така, за да се научи и усвои втор јазик, од голема помош е соодветното разбирање на лексиката и граматиката, бидејќи се работи за дејство кога треба брзо да се научат сложени рутини и алгоритми од јазичната употреба. Ова се случува, на пример, кога различни конјугациски форми на глаголите се сфаќаат како парадигма што треба да се научи, која покажува варијанти во други глаголски времиња (на пример табелата на глаголи во трите форми што вообичаено се учи напамет од учениците/студентите/изучувачите на странскиот јазик). За разлика од излагањата за перформативноста на јазичната употреба, со оваа „системско-лингвистичка“ перспектива, сепак, важни аспекти паѓаат во втор план или потполно се заборават.

Овие две контрастни сфаќања за јазикот ја сочинуваат точната позадина за историјата на вештачката интелигенција. Тие се застапени кај пристапите засновани на правила (системска лингвистика) и статистичките пристапи (јазична употреба, перформанс). Со тоа, во понатамошниот текст треба да стане јасно зошто јазично ориентираната вештачка интелигенција би можела да функционира толку добро, но и каде лежат нејзините граници.

2. КОРИСТЕЊЕТО НА КОМПЈУТЕРОТ И ЕУФОРИЈАТА НА ВЕШТАЧКАТА ИНТЕЛИГЕНЦИЈА (ВИ)

Додека во почетниот период на развојот на компјутерите доминираа аритметичките операции, уште многу рано се согледа потенцијалот компјутерите да се сфатат како машини за обработка на општи симболи. Уште во 1840-тите години, Ада Лавлејс го препозна овој потенцијал во „Analytical Engine“ на Чарлс Бебиџ. Според неа, речиси сите симболи – било да се букви или графички елементи – можат да бидат претставени со бинарно кодирани броеви (Кремер, 2015, стр. 75-90).

Во 1950-тите години, математичари како Алан Тјуринг и Џон фон Нојман покажаа голем интерес за компјутерите и во нив гледаа машини за обработка на симболи, кои се блиски по својата функција до човечкиот мозок (Митчел, 2020, стр. 3). Во 1956 година, младиот математичар Џон Мекарти ја организираше познатата работилница за вештачка интелигенција (Artificial Intelligence Workshop) на колеџот Дартмут, со цел основниот принцип на вештачката интелигенција да се имплементира со помош на компјутери.

Овие пристапи беа интердисциплинарни, но силно обележани од симболизмот и формализмот. Во основа, оваа идеја е многу слична на системско-лингвистичкото сфаќање на јазикот: постојат симболи

што имаат значење и со кои може да се оперира според одредени правила. При тоа не станува збор само за граматички операции, туку и за семантички и логички операции. Еден пример е формалниот јазик на предикатната логика, со кој реченица како:

„Сите автомобили се бучни“ се изразува на следниот начин: $\forall x (\text{Auto}(x) \rightarrow \text{laut}(x))$

Значењето на овој израз е: „За секое x важи: ако x е автомобил, тогаш x е бучен“ (Бубенхофер, Краселт, Принц и Като, 2025, стр. 51).

Доколку информациите, исказите или текстовите воопшто се кодираат според оваа шема, тие лесно можат да се сведат на варијанти на јазичната површинска структура (на пример, „возила“ наместо „автомобили“), а сепак да го задржат истото значење. Исто така, системот може да ги пренесе и на друг јазик, при што е потребно само лексичките единици да се приспособат на друга јазична форма – останатите формално-логички дефинирани правила остануваат непроменети.

Големи беа надежите што се полагаа во овие системи за обработка на симболи. Во 1966 година, Џозеф Вајценбаум ја објави чет-програмата “Eliza”, која требаше да овозможи „разговор на природен јазик со компјутер“ (*natural language conversation with a computer*) /Вајценбаум, 1966, стр. 36–45/. Речникот и збирката правила функционираа речиси како и денешните системи. Оригиналниот програмски код денес повторно е достапен и постојат бројни нови имплементации на оваа програма (Бери, Кистон, Хеи, и сор., 2025).

Овие симболички пристапи и пристапи засновани на правила долго време ја обликуваа автоматската обработка на јазикот, како на пример во машинското преведување: беше потребно синтаксичката структура на една реченица машински да се анализира и според одредени правила да се пренесе во синтаксичката структура на целиот јазик. Покрај тоа, требаше соодветно да се пренесе и зборовниот материјал со пребарување во соодветни речници. Основа за овие операции најчесто е таканареченото „Part-of-Speech-Tagging“ (означување на деловите од говорот), односно (по можност автоматизираното) додавање на информации за видот на зборовите, што овде е прикажано на примерот со TreeTagger, еден автоматски означувач на видови зборови (Schmid, 1994).

Табела 1: Ознаки за делови од говорот

Збор/збороформа (Wort/Wortform)	Tag (Tag / POS)	Лема (Lemma)
Zudem	ADV (Adverb)	Zudem
Musste	VMFIN (Finites Modalverb)	Müssen
das	ART (Artikel)	Das
Wortmaterial	NN (Nomen+Nomen)	Wortmaterial
ebenfalls	ADV (Adverb)	ebenfalls
in	APPR (Präposition)	in
die	ART (Artikel)	die
Zielsprache	NN (Nomen+Nomen)	Zielsprache
überführt	VVPP (Vollverb Partizip Perfekt)	überführen
werden	VAINF (Hilfsverb Infinitiv)	werden
indem	KOUS (Unterordnende Konjunktion mit Satz)	indem
entsprechende	ADJA (Attributives Adjektiv)	entsprechend
Wörterbücher	NN (Nomen+Nomen)	Wörterbuch
abgefragt	VVPP (Vollverb Partizip Perfekt)	abfragen
Worden	VAPP (Hilfsverb Partizip Perfekt)	werden
Sind	VAFIN (finitives Hilfsverb)	sein

Извор: Шмид, 1994

3. ВИ-СИСТЕМИ ЗА ГЕНЕРИРАЊЕ ТЕКСТ ДЕНЕС

Најпознатите алатки за вештачка интелигенција (ВИ), како ChatGPT (OpenAI), Gemini (Google) или Copilot (Microsoft), во одредена мера го „попречуваат пристапот“ кон многуте други примени на големите јазични модели (*Large Language Models*). Во нив спаѓа и машинското преведување, но секако и генерирањето слики, екстракцијата на информации или услугите за асистенција. Поради тоа, во понатамошниот текст општо ќе пишувам за ВИ-системи кои поврзуваат еден или повеќе големи јазични модели со дополнителни елементи, обучени за сосема специфични функции. Бидејќи опкружувањето на овие алатки е многу динамично, на ова место нема да се даде преглед на сите алатки, туку рамка со цел бројните примени да можат систематизирано да се опфатат.

За таа цел се земаат предвид принципите на функционирање и се дефинираат две рамки: а) јазично-комуникативна функција и б) референца на знаење/извори. Притоа, една поединечна алатка, се разбира, може да преземе различни функции од двете рамки.

А) Јазично-комуникативни функции

- Трансформација: Влезниот текст се претвора во друг текст со помош на ВИ. *Примери:* машински превод, парафразирање, промена на стил, корекција, поедноставување, сумирање (резимирање), преформулирање итн.
- Генерирање: Зададената задача (барање) до ВИ резултира со текст. *Примери:* пишување текстови, собирање идеи, аргументи, изработување структури/планови итн.
- Анализа, категоризација, анотација: Влезниот текст се анализира, категоризира или анотира од страна на ВИ според одредени зададени упатства. *Примери:* категоризација на текстови според теми, видови текстови или слично (на пр. препознавање спам), анотација на лингвистички информации (видови зборови, текстуални функции, сопствени имиња итн.).
- Интеракција и асистенција: Дијалоска интеракција со ВИ, каде што резултатите се текстови или дејства (на пр. управување со софтвер). *Примери:* четбот, систем за туторство, прелистувач поддржан од ВИ итн.

Б) Поврзаност со знаењето / упатување на извори

- Интерен модел (*Modell-intern*): ВИ нема можност за пристап до надворешни извори и затоа може да го репродуцира само она што е моделирано во самиот јазичен модел.
- Поддржана со пребарување (*Retrieval-gestützt*): ВИ се надополнува со дополнителни извори кои, исто така, се достапни како вгнездувања и можат да се користат за генерирање на одговорот. Примери: *Retrieval-Augmented Generation (RAG)*, на пример во науката, кај што документираните архиви се прават достапни за некој LLM (голем јазичен модел).
- Поддржана со алатки (*Tool-gestützt*): ВИ користи надворешни алатки како калкулатори, бази на податоци, програмски интерфејси (API) или самостојно креиран код за понатамошна обработка на резултатите. Примери: Прелистувачи поддржани од ВИ кои самостојно можат да повикуваат и читаат веб-страници, комплексни четботови итн.
- Систем поддржан со работни процеси (*Workflow-gestützt*): Системите за ВИ се интегрирани во работните процеси. Примери: Работен процес со „човек во циклусот на одлучување“ (*Human in the Loop*), во кој ВИ презема (деумно) одредени задачи, но луѓето, сепак, го надгледуваат конечниот резултат.

4. ЗАКЛУЧОК

Токму во областа на науката за јазикот (лингвистиката) веќе се покажаа како неопходни горенаведените системи и се јавува потреба од развој на тенденција кон изнаоѓање на сè повеќе нови начини на работа. Ова води до прашања за тоа како задачите треба смислено да се поделат помеѓу ВИ и човекот и кои цели се поврзани со соодветната задача. Со ВИ што генерира текст, доминантен станува едностран поглед на пишувањето и комуникацијата. Притоа се заборава дека примарно се важни процесите на пишување или комуникација, а не само конечниот текст. Токму во училишните или универзитетските средини се појавуваат нови предизвици: мора да се дискутира за тоа кои вештини се од централно значење за учениците и студентите, како тие треба да се пренесат и како ќе се проверува нивното успешно стекнување.

Важно е да се познаваат не само техничките концепти на системите за вештачка интелигенција (ВИ), туку пред сè и нивните основни епистемолошки претпоставки. Во случајот со генеративните јазични модели, тоа се јазично-теоретски основни претпоставки, односно дистрибуциската хипотеза, кои ги објаснуваат начинот на функционирање, но и границите на генеративните јазични модели.

Општо земено, интересно е да се констатира дека јазикот — попрецизно употребата на јазикот — доживува вистинска ренесанса во машинската обработка на податоци. Долго време во електронската обработка на податоци, текстот се сметаше за тешко обработлив „неструктуриран податок“, па затоа им се даваше предност на „структурираните податоци“ во форма на записи во бази на податоци, нумерички вредности и слично (сп. Гугерли, 2009).

КОРИСТЕНА ЛИТЕРАТУРА

- Bender, E. M., & Koller, A. (2023). *Climbing towards NLU: On meaning, form, and understanding in the age of large language models*. *Computational Linguistics*, 49(1), 1–33. https://doi.org/10.1162/coli_a_00484
- Berry, D., Ciston, S., Hay, C. A., Marino, M. & Millican, P. (2025). Arthur I. Schwarz et. al., ELIZA Archaeology, <https://sites.google.com/view/elizaarchaeology/home> (Stand: 19.12.2025).

- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et.al. (2022). *On the opportunities and risks of foundation models*. Stanford Center for Research on Foundation Models. <https://arxiv.org/abs/2108.07258>
- Bubenhofner, N., Krasselt, J., Prinz, M. & Kato, H. (2025). *Studienbuch Linguistik: Band 1: Gegenwartssprache*, Berlin, S. 51.
- Gugerli, D. (2009). *Suchmaschinen: die Welt als Datenbank*. Frankfurt am Main 2009 (= Edition Unselld 19).
- Ivanovska, B., & Xhaferri, G. (2025). INTERCULTURAL PRAGMATICS AND POLITENESS IN LANGUAGE EDUCATION: A CASE STUDY OF STUDENT COMMUNICATION IN A MULTILINGUAL HIGHER EDUCATION CONTEXT. *KNOWLEDGE - International Journal* , 72(5), 641–648. Retrieved from <https://ojs.ikm.mk/index.php/kij/article/view/7825>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et.al. (2023). *ChatGPT for good? On opportunities and challenges of large language models for education*. **Learning and Individual Differences**, **103**, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Krämer, S. (2015). „Wieso gilt Ada Lovelace als die erste Programmiererin und was bedeutet überhaupt ‘programmieren’“, in: Dies. (Hg.), *Ada Lovelace. Die Pionierin der Computertechnik und ihre Nachfolgerinnen*, Paderborn, S. 75–90.
- Miljković, J. (2025). SLOGANS AND COLOURS AS MEANS OF COMMUNICATION. *SCIENCE International Journal*, 4(1), 93-100. <https://doi.org/10.35120/>
- Mitchell, M. (2020). *Artificial Intelligence: A Guide for Thinking Humans*. First paperback edition, London, S. 3.
- Schmid, H., (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. *Proceedings International Conference on New Methods in Language Processing*, Manchester, UK.
- Weizenbaum, J. (1966). ELIZA – a Computer Program for the Study of Natural Language Communication between Man and Machine, in: *Communications of the ACM*, 9(1), S. 36–45.