

**ARTIFICIAL INTELLIGENCE MODELS IN LINGUISTIC RESEARCH,
CHALLENGES AND OPPORTUNITIES**

ხელოვნური ინტელექტის მოდელები ლინგვისტურ კვლევაში, გამოწვევები და შესაძლებლობები

Tinatin Mshvidobadze

Doctor of Technical Sciences

Associate Professor of Gori State University

Gori, Chavchavadze st. #53, 1400, Georgia

+995555118379 tinikomshvidobadze@gmail.com

<https://orcid.org/0000-0003-3721-9252>

Abstract. Artificial Intelligence (AI) has emerged as a transformative force in linguistic research, offering innovative methodologies and tools for language learning and analysis. This paper provides a comprehensive overview of the potential of AI in linguistic research, with a particular focus on its role in language analysis.

By describing natural language processing (NLP), machine learning (ML), and deep learning (DL) techniques, this paper explores the role of AI in linguistic phenomena in our consciousness.

While AI has revolutionized linguistic research, it also addresses challenges such as the relationship to annotated data, biases in AI models, and ethical considerations.

To address these challenges, two AI technologies are worth considering: computer vision (CV) and natural language processing (NLP). Computer vision can identify text, refine data, and analyze images. NLP translates and interprets multilingual texts.

This paper explores the potential and challenges of using AI technologies in linguistic landscape (LL) research and discusses methods for improving data collection.

Keywords: artificial intelligence, linguistic research, natural language processing, computer vision, language models, machine learning.

თინათინ მშვიდობაძე

ტექნიკურ მეცნიერებათა დოქტორი

გორის სახელმწიფო უნივერსიტეტის ასოცირებული პროფესორი

ქ. გორი, ჭავჭავაძის ქ. = 53, 1400, საქართველო

+995555118379 tinikomshvidobadze@gmail.com

<https://orcid.org/0000-0003-3721-9252>

აბსტრაქტი. ხელოვნური ინტელექტი (AI) ლინგვისტურ კვლევებში ჩამოყალიბდა ტრანსფორმაციულ ძალად, რომელიც გვთავაზობს ენის შესწავლისა და ანალიზის ინოვაციურ მეთოდოლოგიებსა და ინსტრუმენტებს. ნაშრომი წარმოადგენს ხელოვნური ინტელექტის პოტენციალის გამოყენების ყოვლისმომცველ მიმოხილვას ლინგვისტურ კვლევებში, განსაკუთრებული ყურადღება გამახვილებულია მის როლზე ენის ანალიზის პროცესში.

ბუნებრივი ენის დამუშავების (NLP), მანქანური სწავლების (ML) და ღრმა სწავლების (DL) ტექნიკის აღწერის მეშვეობით, ნაშრომში გამოკვლეულია, ხელოვნური ინტელექტის ლინგვისტური ფენომენების როლი ჩვენს ცნობიერებაში.

მიუხედავად იმისა, რომ ხელოვნურმა ინტელექტმა რევოლუცია მოახდინა ლინგვისტიკის კვლევებში, ის ასევე პასუხობს ისეთ გამოწვევებს, როგორცაა დამოკიდებულება ანოტირებულ მონაცემებთან, ხელოვნური ინტელექტის მოდელებში არსებული მიკერძოებები და ეთიკური მოსაზრებები.

ამ გამოწვევების გადასაჭრელად გასათვალისწინებელია ხელოვნური ინტელექტის ორი ტექნოლოგია: კომპიუტერული ხედვა (CV) და ბუნებრივი ენის დამუშავების მოდელები (LLM). კომპიუტერული ხედვის საშუალებით შესაძლებელია ტექსტის იდენტიფიკაცია, მონაცემთა დაზუსტება და ანალიზი. ბუნებრივი ენის დამუშავების მოდელები კი თარგმნა და ინტერპრეტაციას უკეთებენ მრავალენოვან ტექსტებს.

წინამდებარე ნაშრომი იკვლევს ხელოვნური ინტელექტის ტექნოლოგიების გამოყენების შესაძლებლობას და გამოწვევებს ლინგვისტური ლანდშაფტის (LL) კვლევაში და განიხილავს მეთოდებს მონაცემთა შეგროვების გასაუმჯობესებლად.

საკვანძო სიტყვები: ხელოვნური ინტელექტი, ლინგვისტური კვლევა, ბუნებრივი ენის დამუშავება, კომპიუტერული ხედვა, ენობრივი მოდელები, მანქანური სწავლება.

შესავალი. ენა, როგორც ადამიანის კომუნიკაციისა და შემეცნების ფუნდამენტური ასპექტი, დიდი ხანია ინტერესისა და კვლევის საგანია სხვადასხვა აკადემიურ დისციპლინებში. ტრადიციულად, ენის სტრუქტურის შესწავლისას ლინგვისტური კვლევა ეყრდნობოდა თეორიულ ჩარჩოებსა და ემპირიულ მონაცემებს (Elman, 1990: 179-211). თუმცა, ხელოვნური ინტელექტის (AI) გამოჩენასთან ერთად, მოხდა პარადიგმის ცვლილება კვლევაში გამოყენებულ მეთოდოლოგიებსა და ინსტრუმენტებში, რამაც გახსნა ახალი გზები აღმოჩენებისათვის.

ხელოვნური ინტელექტი, რომელიც მოიცავს ადამიანის კოგნიტური ფუნქციების იმიტაციისაკენ მიმართული ტექნოლოგიებისა და მეთოდოლოგიების მრავალფეროვან ერთობლიობას, იქცა ლინგვისტურ კვლევაში ტრანსფორმაციულ ძალად (Graves, 2006:545-552). ისეთი ტექნოლოგიების ინტეგრაციის გზით, როგორცაა ბუნებრივი ენის დამუშავება (NLP), მანქანური სწავლება (ML) და ღრმა სწავლება (DL), მკვლევრებს შეუძლიათ ენის სხვადასხვა ასპექტის ანალიზი უპრეცედენტო სიზუსტითა და ეფექტურობით.

ხელოვნური ინტელექტის NLP ტექნიკური მახასიათებლები კომპიუტერებს საშუალებას აძლევს ინტერპრეტაცია გაუკეთონ და შექმნან ბუნებრივი ენის ტექსტები, რაც ხელს უწყობს ისეთი ამოცანების შესრულებას, როგორცაა მეტყველების ნაწილების მონიშვნა, დამუშავება, დასახელებული ერთეულების ამოცნობა და განწყობის ანალიზი.

მეორეს მხრივ, მანქანური სწავლების ალგორითმები მკვლევრებს საშუალებას აძლევს შეიმუშაონ ენის შესწავლისა და გამოყენების პროგნოზირებადი მოდელები დიდი მოცულობის ენობრივი მონაცემებიდან შესწავლის გზით. გარდა ამისა, მანქანური სწავლების მოდელები, როგორცაა ნეირონული ქსელები, აჩვენებს შესანიშნავ შესაძლებლობებს რთული ენობრივი ნიმუშებისა და სემანტიკური სტრუქტურების აღსაქმელად.

ხელოვნური ინტელექტი მნიშვნელოვან როლს ასრულებს ენის განვითარების პროცესების მოდელირებაში. (Jurafsky, 2019: 361-365). გამოთვლითი მოდელირებისა და სიმულაციის ტექნიკის გამოყენებით მკვლევრებს შეუძლიათ ენის შესწავლაში ჩართული კოგნიტური პროცესების აღწერა და თეორიული ჩარჩოების ტესტირება ემპირიულ მონაცემებთან შედარების გზით.

ენობრივი ლანდშაფტის კვლევა, ასევე ეყრდნობა საჯარო სივრცეებში ენის გამოყენების ხილვადობასა და მნიშვნელობას, რათა გავიგოთ ენის რთული დინამიკა

საზოგადოებრივ ცხოვრებაში. შესაბამისად, საკომუნიკაციო ენის კვლევის მეთოდები „ისეთივე მრავალფეროვანია, როგორც მისი დისციპლინური წარმოშობა და მიმდევრობა“ (Malinowski, 2018: 869–885).

ბუნებრივი ენის მოდელებს, რომლებიც შექმნილია ადამიანური ენის დასამუშავებლად და გენერირებისათვის, შეუძლიათ ენის გამოყენების ნიმუშების აღმოჩენა, სხვადასხვა ენის გავრცელების იდენტიფიცირება და მრავალენოვანი ტექსტების თარგმნა ან ინტერპრეტაცია.

ერიკ ვოსი გვთავაზობს ხელოვნური ინტელექტის (AI) ტექნოლოგიების მკვლევების წინაშე არსებული ზოგიერთი გამოწვევის პოტენციურ გადაწყვეტილებებს. ესენია ორი ხელოვნური ინტელექტის ტექნოლოგია კომპიუტერული ხედვა (CV) და ბუნებრივი ენის მოდელები (LLM). (Voss, 2024: 112–125).

სტატიაში განვიხილავთ ხელოვნური ინტელექტის ტექნოლოგიების გამოყენების შესაძლებლობებსა და გამოწვევებს ლინგვისტური ლანდშაფტის კვლევაში და მეთოდებს მონაცემთა შეგროვების სიზუსტის გასაზრდელად, მისი რთული, მრავალენოვანი გარემოს ანალიზის მხარდაჭერის მიზნით.

მეთოდები. მოცემულ სტატიაში გამოყენებულია აღწერითი, ანალიზისა და განმარტების მეთოდები, რის საფუძველზეც გამოკვეთილია აღნიშნული კვლევის მნიშვნელოვანი საკითხები. კონცეფციის ჩამოსაყალიბებლად დავიმოწმეთ სხვადასხვა მკვლევარების შეხედულებები, მათ საფუძველზე მოვახდინეთ მსჯელობისა და დასკვნების წარმოდგენა.

ნაშრომში მოცემულია ყველაზე ხშირად გამოყენებული ხელოვნური ინტელექტის მეთოდების მიმოხილვა ლინგვისტური ლანდშაფტის კვლევისას, როგორცაა ბუნებრივი ენის დამუშავების (NLP), მანქანური სწავლების (ML) და ღრმა სწავლების (DL) მოდელები.

შედეგები და მსჯელობა. ხელოვნურმა ინტელექტმა რევოლუცია მოახდინა სინტაქსური, სემანტიკური და პრაგმატული ასპექტების ანალიზის მეთოდებში. (Pennington, 2014: 1532 -1543).

მრავალმხრივი კვლევით შესწავლილი იქნა ხელოვნური ინტელექტის, განსაკუთრებით ბუნებრივი ენის დამუშავების (NLP) როლი ენობრივი მოდელების შემუშავებისას. მაგალითად, ჩომსკის ტრანსფორმაციულ-გენერაციული გრამატიკის თეორია (Chomsky, 1957: 117), რომელმაც საფუძველი ჩაუყარა ენის შესწავლის გამოთვლით მიდგომებს, რამაც განაპირობა ხელოვნურ ინტელექტზე დაფუძნებულ მოდელებზე გადასვლა.

გამოთვლითი ლინგვისტიკის პიონერებმა, როგორცაა ტერი ვინოგრადი (Winograd, 1972: 1-191). და როჯერ შანკი (Schank & Abelson, 1977:47-50), შეიმუშავეს ადრეული ხელოვნური ინტელექტის სისტემები, რომლებსაც შეეძლოთ ბუნებრივი ენის ტექსტის გაგება და გენერირება.

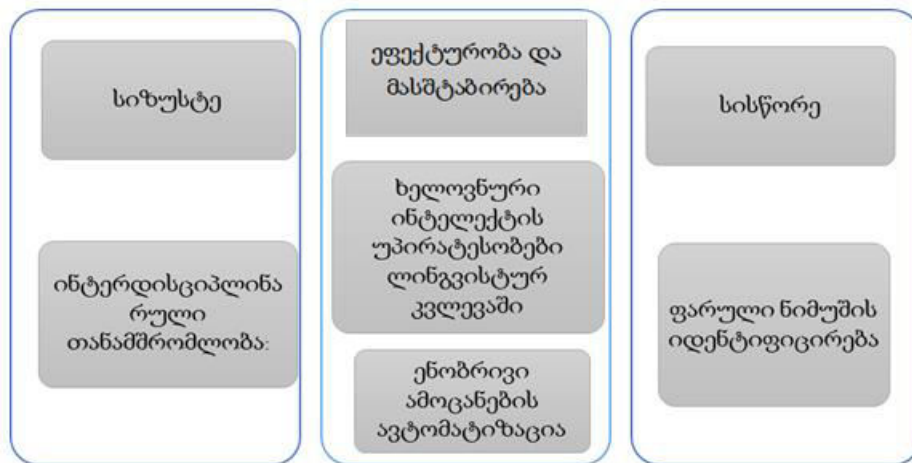
ხელოვნური ინტელექტი გამოიყენება ენის შესწავლის სფეროში, ენის განვითარების პროცესების მოდელირებისა და მისი შესწავლის ტრაექტორიების ანალიზისათვის.

მაგალითად, ასეთია ელმანის მარტივი რეკურენტული ქსელების (SRNs) და კონექციონისტური დროითი კლასიფიკაციის (CTC) მოდელები. ისინი გამოყენებულია ენის შესწავლაში ჩართული კოგნიტური პროცესების სიმულირებისა და მისი განვითარების შედეგების პროგნოზირებისათვის.

ხელოვნური ინტელექტის ტექნიკურმა საშუალებებმა შეცვალა ლინგვისტური ანალიზი ბუნებრივი ენის ტექსტის დამუშავების, ინტერპრეტაციისა და გენერირების მძლავრი ინსტრუმენტების უზრუნველყოფით.

განვიხილავთ ხელოვნური ინტელექტის უპირატესობებს ლინგვისტური ლანდშაფტის კვლევისას. (სურ.1.).

ეფექტურობა და მასშტაბირება: ხელოვნური ინტელექტის ტექნიკა საშუალებას აძლევს მკვლევრებს ეფექტურად დაამუშაონ და გააანალიზონ დიდი მოცულობის ტექსტები. ბუნებრივი ენის დამუშავების (NLP) ალგორითმები, ავტომატურად ახდენენ ტექსტიდან ენობრივი მახასიათებლების ამოღებას, რაც მათ საშუალებას აძლევს გააანალიზონ წერილობითი და სალაპარაკო ენის უზარმაზარი კორპუსები დროის მცირე მონაკვეთში.



სურათი 1. ხელოვნური ინტელექტის უპირატესობები ლინგვისტურ კვლევაში

სიზუსტე: ლინგვისტურ კვლევაში ხელოვნური ინტელექტით მართული მიდგომები ხშირად მაღალი დონის სიზუსტით იძლევა შედეგებს. მანქანური სწავლების (ML) ალგორითმები მუშავდება დიდი მოცულობის ანოტირებულ ლინგვისტურ მონაცემთა ნაკრებზე, რაც მათ საშუალებას აძლევს შეისწავლონ ენაში არსებული რთული ნიმუშები და ურთიერთობები (Devlin, 2019:4171-4186). შედეგად, ხელოვნური ინტელექტის მოდელებს შეუძლიათ შეასრულონ ისეთი დავალებები, როგორცაა მეტყველების ნაწილების მონიშვნა, სინტაქსური დამუშავება და სემანტიკური ანალიზი, რაც იძლევა სანდო ინფორმაციას ენობრივი სტრუქტურებისა და მნიშვნელობების შესახებ.

დაფარული ნიმუშების გამოვლენა: ხელოვნური ინტელექტის ტექნიკურ ინსტრუმენტებს აქვს უნარი გამოავლინოს დახვეწილი ლინგვისტური ნიმუშები და ურთიერთობები, რომლებიც შეიძლება ადვილი არ იყოს ხელით ანალიზის დროს. მანქანური (ML) და ღრმა სწავლების (DL) ალგორითმების ლინგვისტურ მონაცემებზე გამოყენებით, მკვლევრებს შეუძლიათ დაადგინონ დაფარული კორელაციები და ტენდენციები, რაც იწვევს ახალ აღმოჩენებსა და ხედვებს ისეთ სფეროებში, როგორცაა ენის ვარიაცია, ცვლილება და დისკურსის ანალიზი.

ენობრივი ამოცანების ავტომატიზაცია: ხელოვნურ ინტელექტზე დაფუძნებული ინსტრუმენტები და აპლიკაციები ახდენს სხვადასხვა ენობრივი ამოცანების ავტომატიზაციას, რაც ზრდის პროდუქტიულობას და ეფექტურობას ლინგვისტურ კვლევაში. მაგალითად, მეტყველების ამოცნობის ავტომატიზებული სისტემები ცვლიან სალაპარაკო ენას ტექსტად, სალაპარაკო კორპუსების ანალიზისათვის. ანალოგიურად, მანქანური თარგმანის სისტემები საშუალებას იძლევა ტექსტის ავტომატური თარგმნისა სხვადასხვა ენიდან, რაც ხელს უწყობს ენობრივ კვლევებსა და ურთიერთობას.

ინტერდისციპლინარული თანამშრომლობა: ხელოვნური ინტელექტის ინტეგრაცია ლინგვისტურ კვლევაში ხელს უწყობს ინტერდისციპლინარულ თანამშრომლობას

ლინგვისტიკის, კომპიუტერული მეცნიერებების, ფსიქოლოგიებისა და სხვა სფეროების წარმომადგენლებს შორის.

მიუხედავად იმისა, რომ ხელოვნურმა ინტელექტმა მნიშვნელოვანი პოზიტიური ცვლილებები შეიტანა ლინგვისტურ კვლევაში, მას თან ახლავს რამდენიმე გასათვალისწინებელი შეზღუდვა:

- ხელოვნური ინტელექტის ტექნიკური საშუალებები, განსაკუთრებით მანქანური სწავლებისა და ღრმა სწავლების მოდელები, კვლევისათვის ხშირად მოითხოვს დიდი რაოდენობის ანოტირებულ მონაცემებს.
- ლინგვისტური მონაცემების ანოტირება შეიძლება იყოს შრომატევადი პროცესი, განსაკუთრებით ისეთი ამოცანებისთვის, როგორცაა სინტაქსური ან სემანტიკური დამუშავება. (Pang, 2008: 1–135).
- გარდა ამისა, ანოტირებული მონაცემების ხელმისაწვდომობა შეიძლება შეზღუდული იყოს გარკვეული ლინგვისტური ერთეულებისათვის, რაც იწვევს ხელოვნური ინტელექტის მოდელებში მიკერძოებებს და ხელს უშლის დასკვნების განზოგადებას.

მოდელის შემუშავება ენის კვლევისას ხელოვნური ინტელექტის გამოყენებით. ენის კვლევის მოდელირება ხელოვნური ინტელექტის (AI) გამოყენებით წარმოადგენს იმ მრავალმხრივ პროცესს, რომელიც მიზნად ისახავს ენის შესწავლის პროცესში ჩართული კოგნიტური პროცესების სიმულირებას. კვლევის ეს სფერო ენის განვითარების თეორიული ჩარჩოებისა და პროგნოზირებადი მოდელების შესაქმნელად იყენებს ხელოვნური ინტელექტის ტექნიკურ საშუალებებს, როგორცაა გამოთვლითი მოდელირება, მანქანური სწავლება და ბუნებრივი ენის დამუშავება.

ხელოვნური ინტელექტის გამოყენებით ენის შესწავლის მოდელირების ერთ-ერთი მთავარი მიზანია მის საფუძვლად მყოფი კოგნიტური მექანიზმების წარმოდგენა. მაგალითად, ფონოლოგიური უნარების შეძენისა და სიმულირებისათვის გამოყენებულია კონექციონისტური დროითი კლასიფიკაციის (CTC) მოდელები და მარტივი რეკურენტული ქსელები (SRN).

სემანტიკური ანალიზი: სემანტიკური ანალიზი წარმოადგენს კიდევ ერთი სფეროს, სადაც ხელოვნურმა ინტელექტმა მნიშვნელოვანი წვლილი შეიტანა. მანქანური სწავლების და მეტყველების დილემის ტექნიკის გამოყენებით, მკვლევრებს შეუძლიათ ტექსტიდან სემანტიკური ინფორმაციის ამოღება, ენობრივი გამოთქმებიდან მნიშვნელობის გამოტანა და სიტყვებსა და ფრაზებს შორის სემანტიკური ურთიერთობების განსაზღვრა.

სიტყვების ჩასმის მოდელები, როგორცაა Word2Vec და GloVe, სწავლობენ სიტყვების განაწილებულ წარმოდგენას მათი კონტექსტუალური გამოყენების საფუძველზე დიდ ტექსტურ კორპუსებში, რაც მკვლევრებს საშუალებას აძლევს, დააფიქსირონ სემანტიკური მსგავსებები სიტყვებს შორის. (Mikolov et al., 2013).

დისკურსის ანალიზი: დისკურსის ანალიზი წარმოადგენს კიდევ ერთ სფეროს, სადაც ხელოვნური ინტელექტის აპლიკაციებმა ხელი შეუწყო ლინგვისტურ კვლევას. (Attali & Burstein, 2006:127).

ბუნებრივი ენის დამუშავების (NLP) ტექნოლოგია საშუალებას აძლევს მკვლევარებს გაანალიზონ დისკურსის სტრუქტურა, თანმიმდევრულობა წერილობით და სალაპარაკო ტექსტებში. დისკურსის სემანტიკა და დამუშავება გავრცელებული ამოცანებია, რომლებშიც ხელოვნური ინტელექტის ტექნიკური საშუალებები გამოიყენება დისკურსის სტრუქტურებსა და ერთეულებს შორის ურთიერთობების გამოსავლენად. (Jurafsky & Martin, 2019).

ხელოვნური ინტელექტის ტექნოლოგიები. ბუნებრივი ენის მოდელები ამუშავებენ მონაცემებს და იყენებენ ბუნებრივი ენის დამუშავების ქვესფეროს. ეს მოდელები, როგორც წესი, დაფუძნებულია ღრმა სწავლების არქიტექტურებზე და მოქმედებს ტექსტური

მონაცემების უკიდურესად დიდ რაოდენობაზე ტექსტში სემანტიკური და სინტაქსური თვისებების გამოსავლენად. ენობრივი მოდელების ყველაზე ცნობილი მაგალითებია GPT (გენერაციული წინასწარ მომზადებული ტრანსფორმატორი) მოდელები და GPT მოდელებზე დაფუძნებული საჯარო ინსტრუმენტები, როგორცაა ChatGPT, რომელიც შემუშავებულია ტექნოლოგიური კომპანია Open AI-ის მიერ.

ბუნებრივი ენის მოდელები შექმნილია ენობრივი ნიმუშების შესასწავლად და შემდეგ ტექსტის ან მეტყველების სახით ხშირად წარმოქმნილი ნიმუშების გენერირებისთვის, საძიებო წინადადებების მსგავსად (OpenAI, 2024b).

ლინგვისტური კორპუსის მოდელები მუშავდება, რათა ამოიციონ მომდევნო სიტყვა, ფრაზა ან წინადადება კორპუსში პროგნოზირებული ტექსტის ალბათობის საფუძველზე.

ტექსტის იდენტიფიცირება და კოდირება: ლინგვისტური ლანდშაფტის კვლევებში შეგროვებული მონაცემები, როგორც წესი, ანოტირებული და კლასიფიცირებულია, რათა გამოვლინდეს მონაცემებს შორის არსებული კანონზომიერებები და ურთიერთობები.

ხელოვნური ინტელექტის ტექნოლოგიებს შეუძლიათ თავიდან აიცილონ ზოგიერთი შეცდომა, რომელსაც ადამიანები უშვებენ ვიზუალური მონაცემების შეგროვების პროცესში. მონაცემთა შეგროვების ავტომატიზაციით (მაგ., ენის ავტომატური ამოცნობა, ავტომატური მონიშვნა ან გამოსახულების კლასიფიკაცია), ხელოვნური ინტელექტის ტექნოლოგიას შეუძლია შეამციროს მომხმარებლების ძალისხმევა და ხელი შეუწყოს მონაცემთა სწრაფ შეგროვებას (Gilles & Ziegler, 2021: 65–86).

მაგალითად, Lingscape პროექტი ენის ავტომატური ამოცნობისას უკვე დიდი ხანია ექსპერიმენტებს ატარებს ხელოვნური ინტელექტის ინსტრუმენტების დამატებაზე, რაც ხელს შეუწყობს ენის იდენტიფიკაციას და კოდირებას ტექსტის შემცველი სურათების გამოქვეყნებისას. (Purschke, 2021: 9).

დიდი მონაცემების შეგროვება და ანალიზი. ბოლო წლებში შეგროვებული ციფრული მონაცემების რაოდენობა სტაბილურად იზრდება, რაც იწვევს მასშტაბური კორპუსის მონაცემების ანალიზს. „დიდი მონაცემების“ შეგროვების ტენდენცია, რომელიც განიხილება, როგორც „ემპირიულად ობიექტური ინფორმაცია“ (Lyons, 2019: 179–197), გამოყენებულია პროექტებში, რომლებიც ქმნიან „კრაუდსორსინგის“ პლატფორმებს, რათა შეაგროვონ მონაცემები დიდი გეოგრაფიული არეალიდან.

ბუნებრივი ენის მოდელებზე დაფუძნებული გენერაციული ხელოვნური ინტელექტის ინსტრუმენტები, როგორცაა ChatGPT-4.0, შეიძლება გამოყენებულ იქნას მონაცემთა ანალიზში ენების იდენტიფიცირებისას, ტექსტური აზრების, ბილბორდების და სხვა წერილობითი მასალების კონტექსტისა და კონოტაციების აღქმისათვის. მოწინავე ბუნებრივი ენის დამუშავების შესაძლებლობები მათ საშუალებას აძლევს ეფექტურად დაამუშაონ და ინტერპრეტაცია გაუწიონ კომპლექსურ ენობრივ მონაცემებს. დიდი მონაცემთა ანალიზის მაგალითებია ენობრივი სირთულეების შესწავლა ონლაინ და ოფლაინ სივრცეებში (Rosendal et al, 2023: 181–210).

ტექნოლოგიური შეზღუდვები. ამჟამინდელი ენობრივი მოდელების შეზღუდვაა მათი უუნარობა, ამოიციონ არადომინანტური ენების მახასიათებლები. ისინი კარგად არ მუშაობენ დაბალი რესურსების მქონე ენებზე. ასევე, ამ მოდელებს შეიძლება ჰქონდეთ პრობლემები გაურკვეველ ტექსტებთან, რომელთა წაკითხვა ან ამოცნობა შეუძლებელია. უფრო მეტიც, ინდივიდუალური ლექსიკური ერთეულები შეიძლება ამოიციონ მხოლოდ დამოუკიდებლად და არა როგორც ფრაზის ან წინადადების ნაწილი.

დასკვნა. ხელოვნური ინტელექტის ლინგვისტურ კვლევაში ინტეგრაციით კვლევისა და აღმოჩენების ახალი ერა დაიწყო, რამაც შეცვალა ენის შესწავლისა და ანალიზისადმი მიდგომის წესი. ხელოვნური ინტელექტის ისეთი ტექნიკური საშუალებებით, როგორცაა ბუნებრივი ენის დამუშავება (NLP), მანქანური სწავლება (ML) და ღრმა სწავლება (DL),

მკვლევრებს შეუძლიათ ლინგვისტური მონაცემების ანალიზი უპრეცედენტო სიზუსტითა და ეფექტურობით, ფარული ნიმუშების და სემანტიკური სტრუქტურების გამოვლენით, რომლებიც შესაძლოა ხელით ანალიზის დროს არ იყოს ხილული.

ხელოვნური ინტელექტი ენის განვითარების პროცესების მოდელირების ძლიერ ინსტრუმენტად იქცა. გამოთვლითი მოდელირებისა და სიმულაციის ტექნიკის გამოყენებით, მკვლევარებმა ღირებული ინფორმაცია მოიპოვეს ენის შესწავლის შედეგებზე მოქმედი ფაქტორების, ასევე ენის შესწავლის ეტაპებისა და თანმიმდევრობების შესახებ.

ლინგვისტურ კვლევაში ხელოვნური ინტელექტის მიერ მიღწეული მნიშვნელოვანი წინსვლის მიუხედავად, გამოწვევები კვლავ რჩება ისეთ საკითხებში, როგორცაა მონაცემთა ხარისხი, ინტერპრეტაციის შესაძლებლობა, ეთიკა, მრავალენოვნება და თანამშრომლობა. ამ გამოწვევებთან გამკლავება მოითხოვს ინტერდისციპლინურ თანამშრომლობას, ეთიკურ მოსაზრებებს და ინოვაციური მეთოდოლოგიების შემუშავებას.

ხელოვნური ინტელექტის ტექნოლოგიების უწყვეტი განვითარება, ინტერდისციპლინურ თანამშრომლობასთან და კვლევით პრაქტიკასთან ერთად, მკვლევრებს საშუალებას მისცემს, აღმოაჩინონ ახალი ხედვები ენის შესწავლის, ანალიზისა და გამოყენების კუთხით. ხელოვნური ინტელექტის პოტენციალის გამოყენებით, მკვლევრებს შეუძლიათ გააღრმავონ ჩვენი აღქმა ბუნებრივი ენის სირთულეების შესახებ და გაუხსნან გზა ახალ აღმოჩენებსა და ინოვაციებს ლინგვისტიკის სფეროში.

ვინაიდან ხელოვნური ინტელექტის ტექნოლოგიები აგრძელებს განვითარებას და საკუთარი პოტენციალის გამოვლენას, ის, სავარაუდოდ, სულ უფრო მნიშვნელოვან როლს შეასრულებს სამომავლო მეთოდოლოგიების ჩამოყალიბებაში, ენის გამოყენების უფრო ღრმა გაგებისათვის მრავალენოვან საზოგადოებრივ სივრცეებში.

გამოყენებული ლიტერატურა

- Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater. *The Journal of Technology, Learning and Assessment*, 4(3), p. 127;
- Chomsky, N. (1957). Syntactic Structures. *The Hague: Mouton*. p. 117;
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Human language technologies, Volume 1*. (pp. 4171-4186);
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179-211.
- Graves, A., Schmidhuber, J., & Mohamed, A. (2006). Off-line handwriting recognition with multidimensional recurrent neural networks. *Advances in neural information processing systems*, 18, 545-552.
- Graves, A., Schmidhuber, J., & Mohamed, A. (2006). Off-line handwriting recognition with multidimensional recurrent neural networks. *Advances in neural information processing systems*, 18, 545-552;
- Gilles, Peter & Ziegler, Evelyn (2021). Exploring corpus linguistics approaches in Linguistic Landscape research with automatic text recognition software. *In Evelyn Ziegler & Heiko. Berlin: Peter Lan*. (pp. 65-86);
- Lyons, Kate (2019). Let's get phygital: Seeing through the ltered' landscapes of Instagram. *Linguistic Landscape*, 5(2), 179-197;
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *In Advances in neural information processing systems* (pp. 3111-3119);

- Malinowski, David (2018). Linguistic Landscape. In Aek Phakiti, Paul De Costa, Luke Plonsky & Sue Stareld (Eds.) *The Palgrave handbook of applied linguistics research methodology*. London: *Palgrave Macmillan*, (pp. 869–885);
- OpenAI (2024b). How ChatGPT and our language models are developed, *OpenAI*;
- Jurafsky, D., & Martin, J. H. (2019). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed.). *Pearson*. 361-365;
- Purschke, Christoph (2021). Crowdsapes. Participatory research and the collaborative (re)construction of Linguistic Landscapes with Lingscape. *Linguistics vanguard*, 7(s1), 20190032 . p.9;
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135;
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *In Advances in neural information processing systems* (pp. 3111-3119).
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532 -1543);
- Rosendal, Tove, Nielsen, Helle Lykke, Järlehed, Johan, Milani, Tomasso M. & Löfdahl, Maria (2023). Language, translocality and urban change: Online and offline signage in four Gothenburg neighbourhoods. *Linguistic Landscape*, 9 (2), 181–210;
- Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. *Psychology Press*. p.47-50;
- Voss, Erik (2024). Artificial intelligence in language assessment. In A. Kunnan (Ed.) *The Concise Companion to Language Assessment*, Hoboken, NJ: *Wiley-Blackwell*. (pp. 112–125);
- Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), 1-191.

REFERENCES

- Attali, Y., & Burstein, J. (2006). Automated essay scoring with e-rater. *The Journal of Technology, Learning and Assessment*, 4(3), p. 127;
- Chomsky, N. (1957). Syntactic Structures. *The Hague: Mouton*. p. 117;
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Human language technologies, Volume 1*. (pp. 4171-4186);
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179-211.
- Graves, A., Schmidhuber, J., & Mohamed, A. (2006). Off-line handwriting recognition with multidimensional recurrent neural networks. *Advances in neural information processing systems*, 18, 545-552.
- Graves, A., Schmidhuber, J., & Mohamed, A. (2006). Off-line handwriting recognition with multidimensional recurrent neural networks. *Advances in neural information processing systems*, 18, 545-552;
- Gilles, Peter & Ziegler, Evelyn (2021). Exploring corpus linguistics approaches in Linguistic Landscape research with automatic text recognition software. *In Evelyn Ziegler & Heiko*. Berlin: *Peter Lan*. (pp. 65–86);
- Lyons, Kate (2019). Let’s get phygital: Seeing through the ltered’ landscapes of Instagram. *Linguistic Landscape*, 5(2), 179–197;
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *In Advances in neural information processing systems* (pp. 3111-3119);

- Malinowski, David (2018). Linguistic Landscape. In Aek Phakiti, Paul De Costa, Luke Plonsky & Sue Stareld (Eds.) *The Palgrave handbook of applied linguistics research methodology*. London: *Palgrave Macmillan*, (pp. 869–885);
- OpenAI (2024b). How ChatGPT and our language models are developed, *OpenAI*;
- Jurafsky, D., & Martin, J. H. (2019). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed.). *Pearson*. 361-365;
- Purschke, Christoph (2021). Crowdscape. Participatory research and the collaborative (re)construction of Linguistic Landscapes with Lingscape. *Linguistics vanguard*, 7(s1), 20190032 . p.9;
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135;
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *In Advances in neural information processing systems* (pp. 3111-3119).
- Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532 -1543);
- Rosendal, Tove, Nielsen, Helle Lykke, Järlehed, Johan, Milani, Tomasso M. & Löfdahl, Maria (2023). Language, translocality and urban change: Online and offline signage in four Gothenburg neighbourhoods. *Linguistic Landscape*, 9 (2), 181–210;
- Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. *Psychology Press*. p.47-50;
- Voss, Erik (2024). Artificial intelligence in language assessment. In A. Kunnan (Ed.) *The Concise Companion to Language Assessment*, Hoboken, NJ: *Wiley-Blackwell*. (pp. 112–125);
- Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), 1-191.