Ushbu onlayn konferensiya barcha fan tarmoqlaridan ilmiy tezislarni qabul qiladi.

MAVZU: KOMPYUTER LINGVISTIKASI: SUN'IY INTELLEKT VA TABIIY TILNI QAYTA ISHLASH (NLP)

Nasirdinova Orzuxon Abdulxay qizi
Andijon davlat chet tillari instituti
Xorijiy filologiya va turizm fakulteti 1-bosqich talabasi

Annotatsiya

Mazkur maqolada kompyuter lingvistikasining zamonaviy rivojlanish tendensiyalari, sun'iy intellekt texnologiyalari va tabiiy tilni qayta ishlash (NLP) yo'nalishining nazariy hamda amaliy jihatlari yoritilgan. Shuningdek, o'zbek tilini kompyuter yordamida qayta ishlash, elektron lug'atlar va lingvistik korpuslar yaratish, avtomatik tarjima tizimlarini takomillashtirish hamda matnlarni avtomatik tahlil qilish masalalari ko'rib chiqilgan. O'zbek olimlarining kompyuter lingvistikasi sohasidagi ilmiy izlanishlari asosida milliy til resurslarini rivojlantirish va sun'iy intellekt texnologiyalarini o'zbek tiliga tatbiq etish istiqbollari tahlil qilingan.

Аннотация

В данной статье рассматриваются современные тенденции развития компьютерной лингвистики, а также теоретические и практические аспекты применения технологий искусственного интеллекта и обработки естественного языка (NLP). Особое внимание уделяется компьютерной обработке узбекского языка, созданию электронных словарей и лингвистических корпусов, совершенствованию систем машинного перевода и автоматическому анализу текстов. На основе научных исследований узбекских учёных в области компьютерной лингвистики анализируются перспективы развития национальных языковых ресурсов и внедрения технологий искусственного интеллекта в обработку узбекского языка.

Abstract

This article examines current trends in computational linguistics and the theoretical and practical aspects of applying artificial intelligence technologies and natural language processing (NLP). Particular attention is given to the computational processing of the Uzbek language, the development of electronic dictionaries and linguistic corpora, the improvement of machine translation systems, and automatic text analysis. Based on the research of Uzbek scholars in computational linguistics, the article analyzes the prospects for developing national language resources and implementing artificial intelligence technologies for Uzbek language processing.

Kalit so'zlar: kompyuter lingvistikasi, sun'iy intellekt, tabiiy tilni qayta ishlash, NLP, o'zbek tili, lingvistik korpus, avtomatik tarjima, elektron lug'at.

Ключевые слова: компьютерная лингвистика, искусственный интеллект, обработка естественного языка, NLP, узбекский язык, лингвистический корпус, машинный перевод, электронный словарь.

Keywords: computational linguistics, artificial intelligence, natural language processing, NLP, Uzbek language, linguistic corpus, machine translation, electronic dictionary.

KIRISH

XXI asrda axborot-kommunikatsiya texnologiyalarining jadal rivojlanishi inson faoliyatining barcha sohalarida tub o'zgarishlarni yuzaga keltirmoqda. Ayniqsa, sun'iy intellekt texnologiyalarining

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.

takomillashuvi inson va kompyuter o'rtasidagi muloqot imkoniyatlarini kengaytirib, tabiiy tillarni avtomatik qayta ishlash masalasini zamonaviy ilm-fanning dolzarb yo'nalishlaridan biriga aylantirdi.

Bugungi kunda matnlarni avtomatik tahlil qilish, nutqni tanib olish, mashina tarjimasi, elektron lug'atlar yaratish hamda matn mazmunini anglash kabi jarayonlar kompyuter lingvistikasi va tabiiy tilni qayta ishlash (Natural Language Processing — NLP) texnologiyalari asosida amalga oshirilmoqda.

Kompyuter lingvistikasi tilshunoslik, informatika, matematik modellashtirish va sun'iy intellekt kesishgan nuqtada shakllangan fanlararo yo'nalish bo'lib, tabiiy tillarning tuzilishi va qonuniyatlarini kompyuter yordamida o'rganish hamda qayta ishlash usullarini ishlab chiqishga xizmat qiladi. Ushbu sohaning rivojlanishi nafaqat axborotni qayta ishlash tezligi va samaradorligini oshiradi, balki turli tillarda so'zlashuvchi insonlar o'rtasidagi muloqotni yengillashtirish, raqamli ta'lim resurslarini takomillashtirish va milliy tillarning raqamli makondagi imkoniyatlarini kengaytirishga ham yordam beradi.

O'zbek tilining raqamli muhitdagi mavqeini mustahkamlash, uning boy leksik-semantik imkoniyatlarini zamonaviy axborot texnologiyalari vositasida namoyon etish bugungi o'zbek tilshunosligining muhim vazifalaridan biridir. O'zbek tilining agglutinativ xususiyati, so'zlarning qo'shimchalar yordamida turli grammatik shakllarga kirishi, erkin so'z birikmalari hamda kontekstga bog'liq ma'no xususiyatlari tilni avtomatik qayta ishlash jarayonida o'ziga xos ilmiy va texnologik muammolarni yuzaga keltiradi. Ushbu muammolarni hal etish uchun o'zbek tilining elektron korpuslarini shakllantirish, morfologik tahlil algoritmlarini ishlab chiqish, so'z va iboralarning semantik xususiyatlarini modellashtirish hamda mashina tarjimasi tizimlarini takomillashtirish zarur.

O'zbekistonda kompyuter lingvistikasi, elektron lug'atshunoslik, avtomatik tarjima va matnlarni lingvistik tahlil qilish yo'nalishlarida olib borilayotgan ilmiy izlanishlar milliy til resurslarini raqamlashtirish uchun muhim asos bo'lib xizmat qilmoqda. O'zbek olimlarining ushbu sohadagi tadqiqotlari til hodisalarini formal tavsiflash, lingvistik ma'lumotlarni tizimlashtirish va ularni kompyuter dasturlari yordamida qayta ishlash masalalarini o'rganishga qaratilgan. Mazkur izlanishlarni sun'iy intellektning zamonaviy yondashuvlari bilan uyg'unlashtirish o'zbek tilini raqamli texnologiyalar bilan integratsiyalashning yangi imkoniyatlarini ochadi.

ADABIYOTLAR TAHLILI VA METODOLOGIYASI

Kompyuter lingvistikasi zamonaviy tilshunoslikning jadal rivojlanayotgan yo'nalishlaridan biri bo'lib, tabiiy tillarni matematik modellashtirish, kompyuter yordamida tahlil qilish va avtomatlashtirilgan til tizimlarini yaratish masalalarini qamrab oladi. Ushbu sohaning nazariy asoslarini o'rganishda o'zbek tilshunosligida yaratilgan ilmiy adabiyotlar, o'quv qo'llanmalar va monografiyalar muhim ahamiyat kasb etadi.

O'zbek kompyuter lingvistikasining nazariy asoslarini o'rganishda A. Rahimovning "Kompyuter lingvistikasi asoslari" nomli o'quv qo'llanmasi muhim manbalardan biri hisoblanadi. Mazkur asarda kompyuter lingvistikasining asosiy tushunchalari, nazariy masalalari va til hodisalarini kompyuter yordamida o'rganish imkoniyatlari yoritilgan. Muallifning ishlari tilshunoslik va axborot texnologiyalari o'rtasidagi o'zaro aloqani anglash hamda lingvistik birliklarni formal tavsiflashning ahamiyatini tushunishda nazariy asos bo'lib xizmat qiladi.[1]

A. Po'latov va S. Muhamedovaning "Kompyuter lingvistikasi" nomli o'quv qo'llanmasi ham o'zbek tilshunosligida ushbu yo'nalishning shakllanishi va rivojlanishini o'rganishda ahamiyatlidir. Unda kompyuter lingvistikasining asosiy tushunchalari, matematik lingvistika bilan aloqasi hamda til hodisalarini kompyuter vositasida tahlil qilish masalalari ko'rib chiqilgan. Ushbu ilmiy yondashuvlar

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.

tabiiy tilni qayta ishlash texnologiyalarining lingvistik asoslarini anglash uchun muhim nazariy zamin yaratadi.[2]

Kompyuter lingvistikasining zamonaviy rivojlanish bosqichida N.Abdurahmonovaning ilmiy ishlari alohida e'tiborga loyiq. Olimaning "Kompyuter lingvistikasi" darsligida tabiiy tillarni kompyuter texnologiyalari yordamida qayta ishlashning lingvistik asoslari hamda o'zbek kompyuter lingvistikasining nazariy va amaliy masalalari yoritilgan. Shuningdek, muallifning o'zbek tili elektron korpusining kompyuter modellariga bag'ishlangan monografiyasi til birliklarini raqamli muhitda tizimlashtirish, lingvistik ma'lumotlarni qayta ishlash va milliy til resurslarini rivojlantirish masalalarini o'rganishda muhim manba hisoblanadi.[3]

O'zbek tilini avtomatik qayta ishlashda morfologik tahlil ham muhim o'rin tutadi. Sh. Hamroyevaning o'zbek tili morfologik analizatorining lingvistik ta'minotiga bag'ishlangan tadqiqoti so'zlarning grammatik xususiyatlarini avtomatik aniqlash, morfologik shakllarni farqlash va til birliklarini kompyuter dasturlari yordamida tahlil qilish masalalarini o'rganish uchun ahamiyatlidir.[4]

Bunday tadqiqotlar o'zbek tilining grammatik tuzilishini hisobga oluvchi avtomatlashtirilgan tizimlarni ishlab chiqish zaruratini ko'rsatadi.

Tahlil qilingan adabiyotlar shuni ko'rsatadiki, o'zbek kompyuter lingvistikasida nazariy tushunchalarni shakllantirish, elektron lug'atlar yaratish, lingvistik korpuslarni ishlab chiqish va morfologik tahlil vositalarini takomillashtirish bo'yicha muhim ilmiy izlanishlar olib borilgan. Shu bilan birga, sun'iy intellektning zamonaviy usullarini, jumladan, mashinaviy o'qitish va chuqur o'qitish yondashuvlarini o'zbek tilining grammatik hamda semantik xususiyatlari bilan uyg'unlashtirish masalalari keyingi tadqiqotlar uchun dolzarb yo'nalish bo'lib qolmoqda.

Shu nuqtayi nazardan, mazkur maqolada kompyuter lingvistikasining nazariy asoslari va tabiiy tilni qayta ishlash texnologiyalari o'zaro bog'liqlikda o'rganilib, ularning o'zbek tiliga tatbiq etish imkoniyatlari tahlil qilinadi.

Mazkur tadqiqot kompyuter lingvistikasi, sun'iy intellekt va tabiiy tilni qayta ishlash texnologiyalarining nazariy hamda amaliy xususiyatlarini o'rganishga qaratilgan. Tadqiqot metodologiyasi o'zbek tilshunosligida yaratilgan ilmiy adabiyotlarni tahlil qilish, mavjud nazariy qarashlarni umumlashtirish va til birliklarini avtomatik qayta ishlash jarayonlarining asosiy xususiyatlarini aniqlash tamoyillariga asoslanadi.

Tadqiqotning nazariy-metodologik asosini A. Rahimov, A. Po'latov, S. Muhamedova va N. Abdurahmonovaning kompyuter lingvistikasi hamda tabiiy tillarni kompyuter yordamida qayta ishlashga oid ilmiy ishlari tashkil etadi. Ushbu manbalar yordamida kompyuter lingvistikasining asosiy tushunchalari, lingvistik modellashtirish tamoyillari va o'zbek tilining raqamli resurslarini rivojlantirish masalalari o'rganildi.[5]

Tadqiqot davomida quyidagi metodlardan foydalanish nazarda tutildi:
1. Tavsifiy tahlil metodi. Ushbu metod yordamida kompyuter lingvistikasi va tabiiy tilni qayta ishlash tushunchalarining mazmuni, asosiy vazifalari hamda qo'llanilish yo'nalishlari tavsiflanadi. Shuningdek, til birliklarini kompyuter yordamida qayta ishlashning nazariy xususiyatlari yoritiladi.

2. Qiyosiy tahlil metodi. Mazkur metod orqali an'anaviy lingvistik tahlil usullari bilan avtomatlashtirilgan tahlil yondashuvlari o'rtasidagi farqlar o'rganiladi. O'zbek tilining morfologik va grammatik xususiyatlarini hisobga olishda turli yondashuvlarning imkoniyatlari hamda cheklovlari taqqoslanadi.

3. Tizimli tahlil metodi. Kompyuter lingvistikasi, sun'iy intellekt va tabiiy tilni qayta ishlash o'zaro bog'liq ilmiy-texnologik tizim sifatida ko'rib chiqiladi. Ushbu yondashuv tilshunoslik bilimlari,

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.

lingvistik ma'lumotlar bazalari, algoritmlar va dasturiy vositalarning o'zaro aloqasini aniqlashga yordam beradi.

4. Umumlashtirish va sintez metodi. O'rganilgan ilmiy adabiyotlar hamda nazariy qarashlar asosida umumiy xulosalar shakllantiriladi. Natijada o'zbek tilini avtomatik qayta ishlashda mavjud imkoniyatlar, muammolar va istiqbolli yo'nalishlar aniqlanadi.

5. Lingvistik modellashtirish yondashuvi. Ushbu yondashuv til birliklarini kompyuter tizimlarida formal ifodalash, so'zlarning grammatik xususiyatlarini aniqlash va matnlarni avtomatik tahlil qilish tamoyillarini o'rganishda qo'llanadi. Xususan, o'zbek tilining qo'shimchalar orqali grammatik shakllanishi va so'zlarning kontekstga bog'liq ma'nolarini hisobga olish masalalariga e'tibor qaratiladi.

Tadqiqotning asosiy obyekti kompyuter lingvistikasi va tabiiy tilni qayta ishlash texnologiyalari bo'lsa, uning predmeti sun'iy intellekt vositalarining o'zbek tilini avtomatik tahlil qilish, lingvistik ma'lumotlarni qayta ishlash va raqamli til resurslarini rivojlantirishdagi imkoniyatlaridan iborat.

Tadqiqot jarayonida ilmiy adabiyotlar tahlili asosida olingan nazariy xulosalar umumlashtirilib, o'zbek tilini kompyuter yordamida qayta ishlashning dolzarb masalalari tizimlashtiriladi. Metodologik yondashuvning asosiy maqsadi mavjud ilmiy qarashlarni izchil tahlil qilish, muammolarni aniqlash va sun'iy intellekt texnologiyalarini o'zbek tiliga tatbiq etish istiqbollarini ilmiy jihatdan asoslashdan iborat.

NATIJA VA MUHOKAMA

Tadqiqot davomida kompyuter lingvistikasi, sun'iy intellekt va tabiiy tilni qayta ishlash texnologiyalarining o'zaro aloqasi hamda ularni o'zbek tiliga tatbiq etish imkoniyatlari tahlil qilindi. O'zbek olimlarining ilmiy ishlari va sohaga oid nazariy qarashlarni o'rganish natijasida milliy til resurslarini rivojlantirish, lingvistik ma'lumotlarni tizimlashtirish va avtomatik tahlil vositalarini takomillashtirish ushbu yo'nalishning muhim vazifalari ekanligi aniqlandi.[6]

Birinchidan, o'zbek tilining elektron resurslarini kengaytirish tabiiy tilni qayta ishlash tizimlari samaradorligini oshirishning asosiy shartlaridan biri sifatida namoyon bo'ldi. Elektron korpuslar, raqamli lug'atlar va lingvistik ma'lumotlar bazalari til birliklarining qo'llanish xususiyatlarini o'rganish, so'zlar o'rtasidagi semantik munosabatlarni aniqlash hamda til hodisalarini kompyuter yordamida tahlil qilish imkonini beradi. N. Abdurahmonovaning o'zbek tili elektron korpusi va kompyuter lingvistikasiga oid tadqiqotlari ushbu resurslarni yaratishning nazariy hamda amaliy ahamiyatini ko'rsatadi.[7

Ikkinchidan, o'zbek tilining morfologik xususiyatlarini avtomatik aniqlash tabiiy tilni qayta ishlashdagi muhim masalalardan biri ekani belgilandi. O'zbek tilida so'zlarga turli grammatik va so'z yasovchi qo'shimchalarning qo'shilishi natijasida ko'plab so'zshakllar hosil bo'ladi. Shu sababli avtomatik tahlil tizimi so'zning o'zak va qo'shimchalarini farqlashi, grammatik ma'nolarini aniqlashi hamda so'zning gapdagi vazifasini hisobga olishi zarur. Sh. Hamroyeva va B. Mengliyevning o'zbek tili morfologik analizatoriga bag'ishlangan tadqiqotida morfoanalizator tarkibiy birliklari va ularni lingvistik jihatdan ta'minlash masalalari yoritilgan.[8]

Uchinchidan, mashina tarjimasi tizimlarini takomillashtirishda lingvistik bilimlarning ahamiyati aniqlandi. Avtomatik tarjima faqat so'zlarning lug'aviy muqobillarini topish bilan cheklanmaydi. Unda grammatik shakllar, gap tuzilishi, so'zlarning kontekstdagi ma'nosi va uslubiy xususiyatlarini ham hisobga olish talab etiladi. N. Abdurahmonovaning mashina tarjimasining lingvistik ta'minotiga oid tadqiqotlari tarjima tizimlarini yaratishda tilshunoslik bilimlarining muhim o'rin tutishini ko'rsatadi.[9]

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.

To‘rtinchidan, sun’iy intellekt va tabiiy tilni qayta ishlash texnologiyalarini o‘zaro integratsiyalash o‘zbek tilining raqamli imkoniyatlarini kengaytirish uchun istiqbolli yo‘nalish ekanligi asoslandi. Mashinaviy o‘qitish va chuqur o‘qitish algoritmlari matnlardagi takrorlanuvchi qonuniyatlarni aniqlash, matnlarni tasniflash, nutqni tanib olish va savollarga avtomatik javob berish kabi vazifalarda qo‘llanilishi mumkin. Biroq bunday tizimlarning o‘zbek tilida samarali ishlashi uchun yetarli hajmdagi, sifatli va to‘g‘ri belgilangan til ma’lumotlari talab etiladi.

Beshinchidan, tadqiqot natijalari o‘zbek tiliga xos lingvistik xususiyatlarni hisobga olish zarurligini ko‘rsatdi. Jumladan, qo‘shimchalar orqali grammatik ma’no ifodalanishi, omonim shakllarning mavjudligi, so‘zlarning kontekstga qarab turli ma’nolar anglatishi va gap bo‘laklari o‘rtasidagi munosabatlar avtomatik tahlil jarayonida alohida e’tibor talab qiladi. Ushbu xususiyatlarni e’tibordan chetda qoldirish matn mazmunini noto‘g‘ri talqin qilish yoki tarjimada xatolarga olib kelishi mumkin.[10]

Umuman olganda, tahlil natijalari kompyuter lingvistikasini rivojlantirishda nazariy tilshunoslik, raqamli til resurslari va zamonaviy sun’iy intellekt usullarini birgalikda qo‘llash muhimligini ko‘rsatadi. O‘zbek tilining grammatik va semantik xususiyatlarini hisobga oluvchi modellarni ishlab chiqish milliy tilning raqamli muhitdagi imkoniyatlarini kengaytirish uchun zarur ilmiy-amaliy shartlardan biridir.

Olingan tahliliy natijalar kompyuter lingvistikasini faqat texnologik vositalar majmui sifatida emas, balki tilshunoslik bilimlari va hisoblash usullarini birlashtiruvchi fanlararo yo‘nalish sifatida baholash lozimligini ko‘rsatadi.

Kompyuter tabiiy til bilan ishlashi uchun til birliklarining shakli, grammatik vazifasi, ma’nosi va o‘zaro munosabatlarini muayyan qoidalar yoki hisoblash modellari orqali ifodalash zarur. Shu jihatdan, A. Rahimovning kompyuter lingvistikasi asoslariga oid qo‘llanmasi sohaning nazariy tushunchalarini anglash uchun muhim manba hisoblanadi.[11]

O‘zbek tiliga sun’iy intellekt texnologiyalarini tatbiq etishda, eng avvalo, tilning ichki qonuniyatlarini chuqur o‘rganish talab qilinadi. Masalan, bitta o‘zakka turli qo‘shimchalar qo‘shilishi natijasida grammatik ma’nosi farqlanuvchi ko‘plab shakllar hosil bo‘lishi mumkin. Oddiy matn qidiruvi bunday shakllarni har doim ham o‘zaro bog‘lay olmaydi. Morfologik tahlil vositalari esa so‘zning tuzilishini va grammatik xususiyatlarini aniqlash orqali matnni yanada chuqurroq qayta ishlash imkonini beradi. Sh. Hamroyeva va B. Mengliyevning tadqiqotlari o‘zbek tili morfologik analizatorini yaratishda uning tarkibiy birliklari va lingvistik ta’minotiga e’tibor qaratish zarurligini asoslaydi.[12]

Muhokama davomida elektron korpuslarning ahamiyati ham alohida namoyon bo‘ladi. Korpuslar turli janr va uslublarga mansub matnlarni jamlash, so‘zlarning real qo‘llanishini kuzatish, grammatik shakllar chastotasini aniqlash hamda til birliklarining kontekstual xususiyatlarini o‘rganish imkonini beradi. Bunday resurslar nafaqat nazariy tilshunoslik, balki avtomatik tarjima, matn tasnifi, imlo tekshiruvi va axborot qidiruvi tizimlari uchun ham muhim ma’lumot manbai bo‘lib xizmat qiladi. N. Abdurahmonovaning o‘zbek tili elektron korpusiga bag‘ishlangan monografiyasi ushbu yo‘nalishning ilmiy va amaliy ahamiyatini yoritadi.[13]

Shu bilan birga, elektron korpusning mavjudligi o‘z-o‘zidan yuqori sifatli sun’iy intellekt tizimini yaratishni kafolatlamaydi. Korpusdagi matnlar imlo jihatidan to‘g‘ri, turli soha va uslublarni qamrab olgan hamda zarur hollarda morfologik yoki sintaktik belgilar bilan ta’minlangan bo‘lishi kerak. Ma’lumotlarning bir xil uslubga haddan tashqari bog‘lanib qolishi yoki ayrim so‘z shakllarining kam uchrashi model natijalarining aniqligiga ta’sir qilishi mumkin. Shuning uchun milliy til

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.

resurslarini yaratishda matnlarni tanlash, tekshirish, belgilash va yangilab borish jarayonlariga tizimli yondashish zarur.

Mashina tarjimasi masalasida ham shunga o'xshash holat kuzatiladi. So'zma-so'z tarjima ayrim sodda jumlalarda qoniqarli natija berishi mumkin, biroq murakkab grammatik tuzilmalar, iboralar, ko'pma'noli so'zlar va kontekstga bog'liq mazmunni ifodalashda cheklovlarga ega. N. Abdurahmonovaning mashina tarjimasining lingvistik ta'minotiga oid tadqiqoti tarjima tizimlarida tilga xos bilimlarni hisobga olish masalasining dolzarbligini ko'rsatadi.[14]

Demak, tarjima sifatini yaxshilash uchun statistik yoki neyron modellarni lingvistik tahlil usullari va sifatli parallel matnlar bilan uyg'unlashtirish istiqbolli yondashuv hisoblanadi.

Yana bir muhim jihat — o'zbek tilidagi sun'iy intellekt tizimlarining natijalarini baholash masalasidir. Tizimning samaradorligini faqat matnni tez qayta ishlashi bilan o'lchash yetarli emas. Morfologik tahlilda grammatik belgilarni to'g'ri aniqlash, tarjimada mazmunni saqlash, matn tasnifida xatoliklarni kamaytirish va nutqni tanib olishda so'zlarni aniq qayd etish kabi mezonlar ham hisobga olinishi lozim. Baholash mezonlarining aniq belgilanishi turli algoritmlarning imkoniyatlarini xolis taqqoslashga yordam beradi.

Shuningdek, sun'iy intellekt tizimlari o'zbek tilining barcha uslublari va hududiy variantlarini bir xil darajada qayta ishlay olmasligi mumkin. Shevaga xos birliklar, yangi so'zlar, qisqartmalar, norasmiy yozishmalar va imlo xatolari model natijalariga ta'sir ko'rsatadi. Bu esa tizimlarni ishlab chiqishda adabiy til bilan bir qatorda real hayotda uchraydigan turli til ko'rinishlarini ham hisobga olish zarurligini ko'rsatadi.

XULOSA

Xulosa o'rnida qayd etish joizki, kompyuter lingvistikasi zamonaviy tilshunoslikning istiqbolli yo'nalishlaridan biri bo'lib, u tabiiy tillarni kompyuter yordamida tahlil qilish, qayta ishlash va raqamli muhitda qo'llash imkoniyatlarini kengaytiradi. Sun'iy intellekt texnologiyalarining jadal rivojlanishi natijasida tabiiy tilni qayta ishlash (NLP) tizimlari matnlarni avtomatik tahlil qilish, mashina tarjimasi, nutqni tanib olish, matn yaratish hamda inson va kompyuter o'rtasidagi muloqotni takomillashtirishda muhim ahamiyat kasb etmoqda.

O'zbek tilshunosligida kompyuter lingvistikasi bo'yicha olib borilayotgan ilmiy izlanishlar, jumladan, A. Rahimov, N. Abdurahmonova, Sh. Hamroyeva va boshqa olimlarning tadqiqotlari o'zbek tilining raqamli imkoniyatlarini kengaytirishga xizmat qilmoqda. Xususan, o'zbek tilining elektron korpuslarini shakllantirish, morfologik tahlil modellarini ishlab chiqish, so'zlarning grammatik xususiyatlarini avtomatik aniqlash va mashina tarjimasi uchun lingvistik ta'minot yaratish borasidagi ishlar milliy til texnologiyalarini rivojlantirishning muhim ilmiy asosini tashkil etadi.

Tadqiqot davomida ko'rib chiqilgan ilmiy manbalar tahlili shuni ko'rsatadiki, o'zbek tilini sun'iy intellekt tizimlarida samarali qo'llash uchun uning fonetik, leksik, morfologik va sintaktik xususiyatlarini chuqur o'rganish zarur. O'zbek tilining agglutinativ xususiyati, qo'shimchalarning ko'pligi, grammatik shakllarning xilma-xilligi hamda so'zlarning kontekstga bog'liq ma'nolari NLP tizimlarini ishlab chiqishda alohida yondashuvni talab qiladi. Shu sababli, o'zbek tiliga xos lingvistik qoidalarni hisobga oluvchi modellarni yaratish va ularni zamonaviy sun'iy intellekt usullari bilan uyg'unlashtirish dolzarb vazifalardan biri hisoblanadi.

Shuningdek, o'zbek tilidagi sifatli raqamli ma'lumotlar bazasini kengaytirish, belgilangan matn korpuslarini yaratish, ochiq lingvistik resurslarni ko'paytirish va ilmiy tadqiqotlar natijalarini amaliy dasturiy mahsulotlarga tatbiq etish muhim ahamiyatga ega. Bu yo'nalishdagi ishlar avtomatik tarjima sifatini oshirish, matnlarni mazmunan tahlil qilish, o'zbek tilidagi ovozli yordamchilarni

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.

takomillashtirish hamda ta'lim, ilm-fan va axborot texnologiyalari sohalarida tilga oid raqamli xizmatlarni rivojlantirishga yordam beradi.

Umuman olganda, kompyuter lingvistikasi va tabiiy tilni qayta ishlash texnologiyalarining rivojlanishi o'zbek tilining raqamli makondagi mavqeini mustahkamlash, milliy til resurslarini asrash va ulardan samarali foydalanish uchun keng imkoniyatlar yaratadi. Kelgusida o'zbek tiliga mos sun'iy intellekt modellarini takomillashtirish, turli sohalar uchun ixtisoslashtirilgan lingvistik dasturlar ishlab chiqish hamda tilshunoslar, dasturchilar va sun'iy intellekt mutaxassislari o'rtasidagi hamkorlikni kuchaytirish ushbu sohaning ustuvor yo'nalishlari bo'lib qoladi.

FOYDALANILGAN ADABIYOTLAR

1.Rahimov A. Kompyuter lingvistikasi asoslari. — Toshkent: Akademnashr, 2011. — 160 b.
2.Po'latov A., Muhamedova S. Kompyuter lingvistikasi. — Toshkent, 2007.
3.Po'latov A. Kompyuter lingvistikasi. — Toshkent, 2011.
4.Abdurahmonova N. Kompyuter lingvistikasi. — Toshkent: Nodirabegim, 2021. — 394 b.
5.Abdurahmonova N. Mashina tarjimasining lingvistik ta'minoti. — Toshkent: Muharrir, 2018. — 165 b.
6.Abdurahmonova N. O'zbek tili elektron korpusining kompyuter modellari. — Globeedit, 2021.
7.Abdurahmonova N. Mashina tarjimasining lingvistik asoslari. — Toshkent: Akademnashr, 2012. — 120 b.
8.Hamroyeva Sh. O'zbek tili mualliflik korpusini tuzishning lingvistik asoslari. — Monografiya, 2020.
9.Hamroyeva Sh. Korpus lingvistikasi atamalarining qisqacha izohli lug'ati. — Buxoro, 2018.
Muhamedov S.A., Piotrovskiy G.G. Инженерная лингвистика и опыт системно-статистического исследования узбекских текстов. — Toshkent: Fan, 1986.
10.Abjalova M.A. Tahrir va tahlil dasturlarining lingvistik modullari. — Toshkent: Nodirabegim, 2020.

Mualliflik huquqi egasi tomonidan ushbu material Creative Commons Attribution 4.0 (CC BY 4.0) litsenziyasi asosida tarqatiladi.