

VII SHO‘BA KOMPYUTER LEKSIKOGRAFIYASI

UDK: 811.511.141.004

O‘ZBEK TILI DIALEKTLARI IZOHLI LUG‘ATINI YARATISH MODEL VA ALGORITMI

Eshmurodova Madina Dilshod qizi,

Magistrant

madinadilshodovna313@gmail.com

Muhammad al-Xorazmiy nomidagi Toshkent Axborot Texnologiyalari Universiteti

Xudoykulova Nargiza Ravshanovna,

Magistrant

n.khudaykulova@newuu.uz

Muhammad Al-Xorazmiy nomidagi Toshkent Axborot Texnologiyalari Universiteti

Annotatsiya. Ushbu maqolada o‘zbek tilining turli dialektlaridagi so‘zlarni tizimli ravishda tahlil qilish va izohli lug‘at yaratish uchun zamonaviy algoritmlar va metodlarni ishlab chiqish masalasi ko‘riladi. Bunda dialektal farqlarni aniqlash va tushunish uchun lingvistik (fonetik, morfologik, sintaktik) va texnologik yondashuvlar (semantik tahlillar) qo‘llanilish ahamiyati ko‘rib o‘tiladi. Bunda o‘zbek tili dialektlarining o‘ziga xos xususiyatlarini saqlagan holda ularni adabiy tilga moslashtirish usullari, shuningdek, dialektlararo farqlarni avtomatik tarzda aniqlash va tasniflash uchun mashina o‘rganish va tabiiy tilni qayta ishlash (NLP) texnologiyalarini samarali qo‘llash masalalari o‘rganiladi.

Natijada, yaratilgan izohli lug‘at nafaqat ilmiy tahlil va tadqiqotlar uchun, balki til o‘rganish va dialektal farqlarni o‘rganishga qiziquvchilar uchun amaliy resurs sifatida ham foydalidir.

Абстракт. В данной статье рассматривается разработка современных алгоритмов и методов для систематического анализа слов из различных диалектов узбекского языка и создания толкового словаря. При этом подчёркивается важность применения как лингвистических (фонетических, морфологических, синтаксических), так и технологических подходов (семантический анализ) для выявления и понимания диалектных различий. Также изучаются методы адаптации диалектов к литературному языку с сохранением их особенностей, а также эффективное применение технологий машинного обучения и обработки естественного языка (NLP) для автоматического выявления и классификации диалектных различий.

В результате созданный толковый словарь будет полезен не только для научного анализа и исследований, но и как практический ресурс для изучающих язык и интересующихся диалектными различиями.

Abstract. This article discusses the development of modern algorithms and methods for the systematic analysis of words from different dialects of the Uzbek language and the creation of an explanatory dictionary. The importance of using both linguistic (phonetic, morphological, syntactic) and technological (semantic analysis) approaches to identify and understand dialectal differences is emphasised. In addition, methods of adapting dialects to the literary language while preserving their characteristics are analysed, as well as the effective use of machine learning and natural language processing (NLP) technologies for the automatic detection and classification of dialectal differences.

The resulting explanatory dictionary will be of use to the scientific community as well as to language learners and those interested in dialectal differences.

Kalit soʻzlar: *Oʻzbek tili, dialektlar, sheva soʻzlar, izohli lugʻat, dialektologiya, tabiiy tilni qayta ishlash (NLP), mashina oʻrganish (ML), semantik tahlil, fonetik tahlil, morfologik tahlil, adabiy til, kompyuter lingvistikasi, raqamli lugʻat, algoritmlar, tizimli tasniflash, madaniy meros, til boyligi.*

Oʻzbek tili, ayniqsa uning turli hududlar doirasidagi rang-barang dialektlari xalqimizning madaniy va tarixiy merosining ajralmas qismi hisoblanadi. Ularni izohli lugʻatda jamlash, saqlash va tarqatish orqali tilning turli shakllarini va lugʻat boyligini aniqlash, mazmunini tushunish imkonini beradi va oʻzbek tilini oʻrganayotganlar uchun dialektal farqlarni tushunishda qoʻllanma boʻlib xizmat qiladi. Dialektlarni tahlil qilish va izohli lugʻat yaratishda zamonaviy algoritmlar va modellarni qoʻllash oʻzbek tilini raqamlashtirish va kompyuterlingvistikani rivojlantirishga imkon yaratadi. Bu, oʻz navbatida, matnni qayta ishlash, nutqni tanish, avtomatik tarjima kabi sohalarda samarali tizimlarni yaratishga yordam beradi. Oʻzbek dialektal (sheva) soʻzlarni tahlil qilish, ularni izohlash va oʻxshashliklar va farqlarni aniqlashda sunʼiy intellekt, mashinali oʻqitish (ML) va tabiiy tilni qayta ishlash (NLP) texnologiyalarini qoʻllash oʻzbek tili izohli lugʻatini dialektlari izohli lugʻati uchun model va algoritm ishlab chiqish orqali bu jarayonni yanada samarali qilish mumkin. Bunda, dialektlarni izohli lugʻatga kiritish uchun avtomatik alifbo moslashuvi, semantik tahlil va morfologik strukturani aniqlashga qaratilgan algoritmlar ishlab chiqilishi kerak.

Oʻzbek tili dialektlarini oʻrganish va ularning farqlarini aniqlash lingvistika sohasida muhim ahamiyat kasb etadi, dialektologiya, leksikologiya va morfologiya kabi tilshunoslik yoʻnalishlarida yangi dialektal birliklarni aniqlash, ularning

semantik xususiyatlarini tahlil qilish uchun izohli lug'atlar va algoritmik modellar asosiy vosita bo'lib xizmat qiladi.

Dialektal farqlar tahlili nafaqat o'quvchilarga va balki o'zbek tilini o'rganayotgan xorijliklar uchun mo'ljallangan ta'lim materiallari yaratishda katta ahamiyatga ega. Bu materiallar dialektlar o'rtasidagi leksik, fonetik va morfologik tafovutlarni yoritib, til o'rgatish jarayonini yanada samarali bo'lishiga yordam beradi.

Dialektal tizimlar zamonaviy texnologiyalar bilan integratsiyalashganda u kompyuter lingvistikasi doirasida juda keng qo'llanilishi mumkin, tabiiy tilni qayta ishlash (NLP) va mashina o'rganish (Machine Learning) texnologiyalaridan foydalangan holda, turli dialektlar orasidagi farqlarni avtomatik aniqlash, tahlil qilish va tarjima qilish imkoniyati mavjud. Bu jarayon, ayniqsa, avtomatik tarjima va nutqni tanish tizimlarida muhim ahamiyat kasb etadi. Bir dialektida yozilgan matnni boshqa inson uchun tushunarli bo'lgan dialektga yoki rasmiy adabiy tilda ifodalash uchun maxsus algoritmlar ishlab chiqilishi mumkin.

Nutqni tanish tizimlari uchun o'zbek tilining turli hududiy shakllarini aniqlay oladigan modellar yaratish, tilning tabiatini chuqurroq o'rganish va uni texnologik qurilmalarda to'g'ri tanish imkonini beradi.

Shuningdek, dialektlar o'zbek xalqining boy madaniy merosini o'zida mujassam etadi. Turli hududlardagi dialektal birliklar va ularning izohli hujjatli shakllari tarixiy-lingvistik tadqiqotlar uchun muhim manba bo'lishi mumkin. Shu jihatdan, bu model madaniy merosni saqlash, hujjatlashtirish va kelajak avlodlarga yetkazish vositasi sifatida ham xizmat qiladi.

O'zbek tili juda ko'p dialektal shakllarga (shevalarga) ega. Har bir dialekt quyida ko'rsatilgan o'ziga xos fonetik, leksik, grammatik va semantik xususiyatlarga ega:

Dialekt /Sheva	Hudud	Xususiyati
Shimoliy (Samarqand)	Samarqand atrofida rivojlangan	O'ziga xos fonetik va leksik o'zgarishlar
Janubiy (Farg'ona)	Farg'ona vodiysi va Toshkent v.da mavjud	Ba'zi turkcha va forscha so'zlar ko'proq ishlatiladi
G'arbiy (Buxoro)	Buxoro viloyatida tarqalgan	O'ziga xos grammatik va fonetik xususiyatlar
Shimoliy-g'arbiy (Xorazm)	Xorazm viloyatida rivojlangan	O'ziga xos fonetik xususiyatlar, ba'zi so'zlar boshqacha talaffuz qilinadi
Janubi-sharqiy (Surxondaryo)	Surxondaryo viloyatida mavjud	Janubiy turkiy til qoidalariga ta'sirida fonetik va leksik o'zgarishlar bor
Qoraqalpog'	Qoraqalpog'iston hududida mavjud	O'ziga xos grammatik va fonetik farqlar
Toshkent (Markaziy)	Toshkent atrofida rivojlangan	O'zbek tilining markaziy varianti

Shuning uchun, dialektlar o'rtasidagi farqlarni aniqlash va tasniflash birinchi asosiy muammo bo'lib, bu jarayonni ilmiy va texnik jihatdan to'g'ri tashkil etish masalasi mavjud. Buning uchun, o'zbek tili dialektlarining tarkibini aniq tasniflash va ularni turli xususiyatlar (fonetik, leksik, sintaktik, semantik) bo'yicha tahlil qilish va so'zlarning sinonimlari, antonimlari va semantik farqlarini ajratib ko'rsatish talab etiladi. Dialektal so'zlarni yagona izohli lug'atda to'plash va ularga tushuntirish berish, ularning ma'nolarini va ishlatilish kontekstini ta'riflash - bu yana bir katta muammo. Bunday kata hajmdagi ma'lumotlarni kitob holatida to'plash hozirgi kun foydalanuvchilari uchun qiyinchilik tug'dirishi mumkin, chunki bunda kerakli manbaani samarali topa olish va undan ma'lumot izlab topish ko'p vaqt talab qiladi. Shuning uchun, lug'at yaratish jarayonini avtomatlashtirish muhim ahamiyat kasb etadi, buning uchun mashinali o'qitish (ML) va tabiiy tilni qayta ishlash (NLP) texnologiyalarining integratsiyasi zarur.

Yaratilish texnologiyalari:

Ma'lumot to'plash va tahlil qilish:

- O'zbek tilining dialektlarini aniqlash va to'plash uchun tahlil qilish manbaalarini tanlash kerak. Bu manbaalar og'zaki va yozma matnlar,

lugʻatlar, adabiy va ilmiy asarlar, xalqaro va mahalliy lugʻatlar boʻlishi mumkin.

- Turli hududlardan (shaharlar, qishloqlar) olingan matnlar yoki yozma asarlarni tahlil qilish. Bu matnlar orqali oʻziga xos soʻzlar, iboralar va grammatik shakllarni yigʻish mumkin.
- Dialektal soʻzlar va ularning izohlarini jamlash uchun maxsus maʼlumotlar bazasi yaratish zarur. Bu bazaga dialektal soʻzlarning turli variantlarini va ularning semantik maʼnolarini kiritish mumkin.

Dialektlarni tasniflash va tahlil qilish:

- Har bir dialektning fonetik, leksik, sintaktik, semantik xususiyatlarini tahlil qilish. Masalan, bir dialektda “koʻrmoq” soʻzi boshqasida “koʻrish” deb aytiladi, bu dialektal farqlarni toʻgʻri tasniflash kerak boʻladi.
- Soʻzlar va iboralarining maʼnolarini ajratish, sinonimlar va antonimlarni aniqlash.

Tabiiy tilni qayta ishlash (NLP) va mashina oʻrganish (ML):

- Dialektal soʻzlarni avtomatik tanib olish, ularning semantik xususiyatlarini tahlil qilish va izohlarini aniqlash uchun AI modellarini yaratish. Bu modellar dialektal farqlarni oʻrganish va tasniflashda yordam beradi.
- Oʻzbek tili uchun maxsus NLP modellarini yaratish, masalan, soʻzlarning morfologiyasini tahlil qilish, semantik va sintaktik xususiyatlarini aniqlash.

Lugʻat tizimini yaratish va vizualizatsiya:

- Yaratilgan lugʻatni foydalanuvchi uchun qulay tarzda taqdim etish, interaktiv tizimlar (web ilovalar, mobil ilovalar) yordamida lugʻatni izlash va foydalanishni osonlashtirish.
- Foydalanuvchilar uchun izohli lugʻatning tarjima qilish yoki dialektal soʻzlarni boshqa dialektlarga oʻzgartirish imkoniyatini yaratish.

Oʻzbek tili dialektlari va ularning izohli lugʻatini yaratish mavzusida xorijda ham, Oʻzbekiston hududida ham bir qancha ilmiy tadqiqotlar olib borilgan. Ushbu tadqiqotlar asosan tilshunoslik, dialektologiya, leksikologiya, va kompyuter lingvistikasiga oid boʻlib, har ikki mintaqada tilni chuqur oʻrganish va dialektal farqlarni aniqlashga eʼtibor qaratilgan.

Ushbu yoʻnaltirishda ilk tadqiqot namunalarini soʻz mulkining sultoni Alisher Navoiyning asarlarida ham kuzatishimiz mumkin. “**Mahbub ul-qulub**” asarida

Navoiy o'zining tilshunoslik va grammatikaga doir qarashlarini bayon etadi. “Mahbub ul-qulub” asarida Navoiy o'zbek tilining o'ziga xos xususiyatlarini, uning go'zalligini va boyligini ta'kidlaydi. Bu asar tilning strukturasini tahlil qilishda muhim ahamiyatga ega. 20-asrning o'rtalaridan boshlab O'zbekistonning turli hududlarida (Farg'ona vodiysi, Qashqadaryo, Surxondaryo, Buxoro, Samarqand va boshqalar) tilshunoslar o'zbek tilining turli dialektlarini tahlil qilib, ularning fonetik, grammatik va leksik xususiyatlarini o'rgandilar.

- **Qori Niyozov va Abdullaev Shukrullo** kabi tilshunoslar o'zbek tilining turli dialektlarini tahlil qilgan va ushbu dialektlar orasidagi farqlarni o'rgangan.

- **Erkinov Mirzaehsan** (O'zbekistonda dialektologiya sohasida yetakchi tadqiqotchilardan biri) o'z tadqiqotlarida o'zbek tilining har bir dialekti uchun alohida lug'atlar yaratish, so'zlarning fonetik va morfologik o'zgarishlarini aniqlashni maqsad qilgan. U dialektal so'zlarning tarixi, semantik xususiyatlari, va ishlatilish kontekstlarini o'rgangan.

- **G'afur G'ulom va Erkin Vohidov** kabi o'zbek shoir va yozuvchilari dialektal leksikani o'z asarlarida aks ettirib, tilning turli shakllarini ommaga tanishtirganlar. Shuningdek, o'zbek tilining turli dialektlaridagi leksik o'zgarishlarni o'rganish bo'yicha ilmiy izlanishlar mavjud.

O'zbek tilining dialektlarini o'rganishda tashqi omillarning roli ham muhim. Masalan, **Toshkent va Samarqand** kabi shaharlar o'zgarishlarga moyil bo'lgan, chunki bu hududlarda turli etnik guruhlar va ko'chmanchilar birgalikda yashagan. Bu hududlarda o'zbek tilining turli shakllari, jumladan, so'zlarning fonetik xususiyatlari o'zgarib ketgan. Bu mavzu ham ko'plab ilmiy tadqiqotlarda yoritilgan.

- Sovet hukumati davrida ham, muayyan tadqiqotlar olib borilgan bo'lib ulardan, **Yekaterina G. Prokofyeva** (SSSR davrida) o'zbek tilining dialektlarini va turk til oilasiga mansub boshqa tillar o'rtasidagi aloqalarni o'rganib chiqqan. U dialektal so'zlarni tahlil qilish, ularning tarixi va semantik xususiyatlarini izohlashda zamonaviy metodlarni qo'llagan.
- **Maksim Il'ich Kormanov va Anatoliy S. Pocheptsov** kabi tadqiqotchilar kompyuter lingvistikasiga asoslangan dialektal izlanishlarni olib borgan. Ular o'zbek tilining turli dialektlarining morfologiyasini va sintaksisini tahlil qilish uchun matematik modellar va algoritmlardan foydalanganlar.
- **Sun'iy intellekt va tabiiy tilni qayta ishlash (NLP)** sohasida yirik chet el tadqiqotchilari, masalan, **Berkeley Universiteti** va **Max Planck Instituti** lingvistik tadqiqotlar olib boradi. Bu tadqiqotlar doirasida o'zbek tilining dialektal xususiyatlarini tahlil qilish va so'zlarni avtomatik ravishda tanib olish uchun mashina o'rganish va algoritmlardan foydalanish ko'zda tutilgan.
- **Oxford Universiteti** va **Cambridge Universiteti** kabi nufuzli ilmiy muassasalarda, o'zbek tili va uning dialektlari haqida xalqaro tadqiqotlar olib

borilgan. Bular asosan o'zbek tilining boshqa turkiy tillar bilan aloqalarini va dialektlararo o'zgarishlarni o'rganishga qaratilgan.

- **Amerika va Yevropa ilmiy markazlarida** o'zbek tilining turli dialektlari o'rtasidagi farqlarni o'rganish, xalqaro kommunikatsiyada dialektal o'zgarishlarni hisobga olish uchun metodlar ishlab chiqish ishlari olib borilmoqda. Bu dialektlararo tafovutlarni aniqlash va tarqatishda madaniy almashinuvlar, masalan, **tillararo tarjima tizimlari** va **interkultural muloqot** uchun ahamiyatli bo'lishi mumkin.

O'zbekiston hududida va chet elda turli dialektologik tadqiqotlar olib borilgan va hozirda ham bu soha rivojlanmoqda. Chet elda kompyuter lingvistika va sun'iy intellekt yordamida dialektal farqlarni tahlil qilishda ilg'or texnologiyalar qo'llanilsa, O'zbekistonda esa hozirgi kunda asosan an'anaviy tilshunoslik metodlari va zamonaviy tadqiqotlar birgalikda olib borilmoqda. Bu tadqiqotlar tilning boyligini, uning dialektal farqlarini o'rganishda katta ahamiyatga kasb etadi.

O'zbek tili dialektlari izohli lug'atini yaratish va dialektologiya sohasida olib borilgan tadqiqotlar yutuqlari va kamchiliklari mavjud. Quyida ular tahlil qilinadi:

Yutuqlar:

O'zbekistonda amalga oshirilgan dialektologik tadqiqotlar, o'zbek tilining turli hududlaridagi dialektal farqlarni aniqlash va tasniflashda katta yutuqlarga erishdi. Masalan, **Alisher Navoiyning asarlari** orqali tilning tarixiy rivojlanishini o'rganish, **Erkinov Mirzaehsan** kabi mutaxassislar tomonidan dialektlarning fonetik va grammatik xususiyatlari tahlil qilindi. **Sovet davrida** va undan keyin olib borilgan ishlarda, o'zbek tilining turli dialektlari, ularning fonetik, leksik va morfologik xususiyatlari to'liq o'rganildi. O'zbekistonlik tadqiqotchilar tomonidan **dialektal lug'atlar** yaratish va **so'zlarning semantik tahlili** amalga oshirilgan. Bu ilmiy ishlar tilni yaxshiroq tushunishga va o'rganishga yordam beradi, ayniqsa, tilni o'rganayotganlar yoki tarjimonlar uchun foydali. **Kompyuterlingvistika** va **tabiiy tilni qayta ishlash (NLP)** texnologiyalaridan foydalanish, masalan, dialektlarni avtomatik tanib olish va tahlil qilish imkoniyatini yaratgan. Bu, lug'at va dialektal tahlillarni avtomatlashtirishda katta yutuqdir.

Chet eldagi ilmiy markazlar, masalan, **Oxford Universiteti** va **Max Planck Institutida** olib borilgan tadqiqotlar dialektlararo o'zgarishlarni va turk tillari o'rtasidagi aloqalarni o'rganish orqali O'zbekistondagi dialektologik tadqiqotlar bilan taqqoslash imkonini yaratdi. Bu tadqiqotlar xalqaro ilmiy hamkorlikni va o'zbek tilini global miqyosda o'rganishni rivojlantirdi.

Mashina o'rganish (machine learning) va **sun'iy intellekt** texnologiyalaridan foydalanish dialektal so'zlarni avtomatik tanib olish va ularni tahlil qilishda katta yutuq keltirdi. Bu sohada chet elda amalga oshirilgan tadqiqotlar so'zlarning morfologiyasini va semantik aloqalarini o'rganish uchun yangi imkoniyatlarni yaratdi.

Kamchiliklar:

- Hozirgi tadqiqotlar, ayniqsa, O'zbekistondagi hududiy dialektlar to'liq o'rganilmagan. O'zbekistonda ko'plab kichik qishloq va viloyatlarda yashovchi aholi orasidagi dialektal farqlar hali to'liq aniqlanmagan.

- **Ba'zi dialektal so'zlar** yoki **yangi so'zlar** izohli lug'atlarga kiritilmagan, chunki bu dialektal shakllar hali ilmiy jurnallarda yoki tadqiqotlarda yetarlicha yoritilmagan.

- **Mashina o'rganish** va **tabiiy tilni qayta ishlash (NLP)** texnologiyalarining tillar o'rtasidagi dialektal farqlarni to'liq aniqlashdagi cheklovlari mavjud. O'zbek tilidagi dialektal farqlar murakkab va kontekstual bo'lib, avtomatik tizimlar hali ham ularni to'liq tushunish va izohlashda qiyinchiliklarga duch keladi.

- **Mashina o'rganish modellari** dialektal farqlarni aniqlashda yuqori aniqlikni ta'minlashda qiyinchiliklarga duch kelishi mumkin, chunki dialektlarning o'ziga xos grammatik va fonetik xususiyatlari mavjud.

- **O'zbek tilining dialektlari** haqidagi mavjud tadqiqotlar ko'pincha an'anaviy metodlarga asoslangan. Ba'zi holatlarda ilmiy ishlar va lug'atlar ko'proq klassik lingvistik metodlarga asoslangan, ammo zamonaviy texnologiyalar, masalan, **yuzaga keladigan yangi dialektal so'zlar** va **onlayn lug'atlarni** yaratish jarayonida hali ham yetarlicha foydalanilmayapti.

- **Dialektal til materiallarini raqamlashtirish** masalasi O'zbekistonda to'liq amalga oshirilmagan. Ko'plab dialektal matnlar va yozuvlar hozirda raqamlashtirilmagan va tizimli tarzda arxivlanmagan.

- **Xalqaro hamkorlikda**, ayniqsa, kompyuter lingvistikasi va sun'iy intellekt sohasida O'zbekiston va chet el ilmiy markazlari o'rtasidagi o'zaro aloqalar hali ham yetarli darajada rivojlanmagan. Bu esa, masalan, dialektal so'zlarni aniqlash va tahlil qilishda yangi yondashuvlar va metodlarni ishlab chiqish imkoniyatlarini cheklaydi.

- O'zbek tili dialektlari vaqt o'tishi bilan o'zgarib bormoqda, shuning uchun eski dialektlarni o'rganish va hujjatlashtirish qiyinlashmoqda. Ba'zi dialektlar yo'qolib bormoqda, bu esa ilmiy tadqiqotlar va lug'atlar yaratishda qiyinchiliklarga olib keladi.

O'zbek tili dialektlari izohli lug'atini yaratish va dialektologiya sohasida olib borilgan tadqiqotlar o'zining muhim yutuqlariga ega, lekin ular hali ham to'liq va mukammal emas. Kompyuterlingvistika, sun'iy intellekt va zamonaviy ilmiy metodlarni qo'llash, dialektal farqlarni aniqlash va tahlil qilishda katta imkoniyatlar yaratgan. Biroq, dialektal farqlarni to'liq qamrab olish, raqamlashtirish, va ilmiy hamkorlikni kuchaytirish bo'yicha hali ko'plab ishlar qilish kerak.

“O'zbek tili dialektlari izohli lug'atini yaratish modeli va algoritmi” mavzusining metodologiyasi ilmiy tadqiqotlar, texnologiyalar va tilshunoslik metodlariga asoslanadi. Bu metodologiya bir nechta bosqichlardan tashkil topgan bo'lib, dialektal farqlarni tahlil qilish, lug'at yaratish va ma'lumotlarni tizimli

ravishda taqdim etishni o'z ichiga oladi. Quyida ushbu metodologiyani asosiy bosqichlari keltirilgan:

Dialektlarni o'rganish va tahlil qilish: Dialektlar o'rtasidagi farqlarni aniqlash uchun tilshunoslik metodlari, masalan, fonetik, morfologik, syntaktik va semantik tahlillar ishlatiladi. Har bir dialektning o'ziga xos xususiyatlarini aniqlashda quyidagi metodlar qo'llaniladi:

- Fonetik tahlil: Dialektal farqlarni aniqlash uchun so'zlarning talaffuzi, intonatsiyasi va aksentlari o'rganiladi. Bu dialektlarning fonetik xususiyatlarini tahlil qilishda yordam beradi.
- Dialektal so'zlarning shakllanishi va o'zgarishini o'rganish. Masalan, bir dialektda so'zlarning tugallanma shakllari boshqacha bo'lishi mumkin.
- Dialektlar o'rtasida sintaktik tuzilishdagi farqlarni aniqlash, masalan, jumalarni qurish va ularda ishlatiladigan predikatlar va qo'shimchalarning farqlanishi.
- Dialektal so'zlarning ma'nolarini o'rganish. Ba'zi so'zlar bir dialektda boshqa ma'nolarga ega bo'lishi mumkin. Bu semantik tahlilning muhim qismi bo'ladi.

Dialektal leksikalarni yig'ish: Dialektal lug'atlarni yaratish uchun, har bir dialektga oid so'zlar, iboralar, sinonimlar, antonimlar va so'z birikmalarini yig'ish zarur. Bu jarayon quyidagilarni o'z ichiga oladi:

- Og'zaki manbalarni yig'ish: Mahalliy aholi, so'zlashuvdagi dialektal farqlarni to'plash. O'zbek tilining turli hududlarida yashovchi odamlar bilan suhbatlar o'tkazish va ulardan og'zaki so'zlarni yig'ish.
- Yozma manbalarni yig'ish: Mahalliy matnlar, adabiy asarlar, ilmiy ishlanmalar va boshqa yozma manbalar orqali dialektal leksikani to'plash.
- Ma'lumotlarni tuzish va tizimlashtirish: Dialektal so'zlarni tizimli ravishda tuzish uchun ma'lumotlar bazasi yaratiladi.

Bu bazaga har bir so'zning:

- Fonetik, grammatik va semantik tavsifi;
- Izohlar va tarixiy konteksti kiritiladi;
- Sinonimlar, antonimlar va so'z birikmalarining ro'yxati kiradi.

Raqamli resurslar va raqamlashtirish: Dialektal so'zlarni yig'ish va tasniflashda raqamli texnologiyalarni qo'llash zarur. Raqamlashtirish orqali dialektal so'zlarni izlash va tahlil qilish imkoniyati oshadi. Ma'lumotlar bazasini yaratishda quyidagi metodlar qo'llaniladi:

- XML, JSON va SQL kabi ma'lumotlar formatlari.
- Web ilovalar va mobil ilovalar orqali dialektal lug'atga onlayn kirish imkoniyatini yaratish.

Mashina o'rganish va tabiiy tilni qayta ishlash (NLP) texnologiyalaridan foydalanish

So'zlarni tanib olish va tahlil qilish: Dialektal so'zlarni avtomatik ravishda aniqlash uchun NLP algoritmlarini qo'llash zarur. Bu metod quyidagilardan iborat bo'lishi mumkin:

- So'zlarni morfologik tahlil qilish: Mashina o'rganish yordamida dialektal so'zlarni morfologik tahlil qilish, so'zning ildizini va qo'shimchalarini aniqlash.
- So'zlarning vektorial ifodasi (word embeddings): Dialektal so'zlarning ma'nolarini va semantik aloqalarini aniqlash uchun sun'iy neyron tarmoqlari yoki word2vec kabi modellarni ishlatish.
- Semantik tahlil: So'zlarning semantik ma'nolarini aniqlash va dialektal so'zlarni izohli tarzda tasvirlash. Bu jarayon uchun, BERT yoki GPT kabi til modellaridan foydalanish mumkin.

Dialektlarni tasniflash va ajratish: Mashina o'rganish modellaridan (masalan, k-means clustering, random forests, support vector machines) foydalangan holda, dialektlar o'rtasidagi farqlarni aniqlash va avtomatik ravishda tasniflash mumkin. Bu dialektal farqlarni o'rganish va ularni tizimli tarzda guruhlash imkonini beradi.

Leksik va sintaktik o'zgarishlarni aniqlash: Dialektlarning leksik va sintaktik o'zgarishlarini aniqlashda, NLP texnologiyalaridan foydalanish, jumladan, dependency parsing, named entity recognition (NER) kabi usullarni qo'llash mumkin.

Izohli lug'at yaratish: Dialektal so'zlarning izohlarini yaratish uchun, har bir so'zning ishlatilish kontekstini va tarixini taqdim etish kerak. Bu jarayonda quyidagilar amalga oshiriladi:

- Dialektal so'zlar va iboralar uchun izohlar: Har bir so'zning ma'nosi, ishlatilish konteksti va sinonimlari kiritiladi.
- Hikoyalar va namunalar: So'zlar ishlatilgan matnlar yoki xalq og'zaki ijodi namunalarini kiritish.

Interaktiv tizimlar va vizualizatsiya: Izohli lug‘atni foydalanuvchilar uchun qulay tarzda taqdim etish uchun interaktiv tizimlar yaratish. Bu quyidagilarni o‘z ichiga oladi:

- Web ilovalar: Foydalanuvchilarga dialektal so‘zlarni izlash va o‘rganish imkonini beradi.
- Mobil ilovalar: Dialektal so‘zlarni mobil qurilmada o‘rganish va tahlil qilish.
- Ma’lumotlar vizualizatsiyasi: So‘zlar va dialektlar o‘rtasidagi aloqalarni ko‘rsatish uchun grafikalar va diagrammalar yaratish.

Dialektlararo aloqalarni tahlil qilish: Dialektlar o‘rtasidagi aloqalarni aniqlash va tahlil qilishda, semantic networks (semantik tarmoqlar) va ontology-based approaches (ontologiyalar asosida yondashuvlar) qo‘llanilishi mumkin. Bu so‘zlar o‘rtasidagi semantik bog‘lanishlarni va dialektal farqlarni yanada chuqurroq tushunishga yordam beradi.

Ushbu mavzu doirasidagi metodologiya tilshunoslik, kompyuter lingvistika, mashina o‘rganish va tabiiy tilni qayta ishlash texnologiyalaridan foydalangan holda, dialektal so‘zlarni tizimli ravishda yig‘ish, tahlil qilish va izohli lug‘at yaratish jarayonlarini o‘z ichiga oladi. Yaratilgan lug‘at interaktiv tizimlar yordamida foydalanuvchilarga taqdim etiladi, shu bilan birga dialektal farqlarni o‘rganish va tahlil qilish uchun zamonaviy texnologiyalarni qo‘llash ilmiy tadqiqotlarda yangi imkoniyatlarni yaratadi.

Foydalanilgan adabiyotlar:

1. Iskandarov, N. (2001). O‘zbek tilining dialektologiyasi. – Tashkent: O‘zbekiston Milliy Ensiklopediyasi.
2. Raximov, B. (1997). O‘zbek tilining dialektik rivojlanishi va uning muammolari. – Tashkent: Fan.
3. Sultonov, X. (2006). O‘zbek tili va uning dialektik turlari. – Tashkent: O‘zbekiston davlat nashriyoti.
4. Usmonov, N. (2003). O‘zbek tili grammatikasi va dialektologiyasi. – Tashkent: Akademnashr.
5. Xolmatov, Sh. (2009). O‘zbek tilining dialektik va leksik xususiyatlari. – Tashkent: O‘qituvchi.
6. Abduqodirov, S. (2012). O‘zbek tili dialektlarining fonetik xususiyatlari. – Tashkent: Fan va texnologiya.
7. “O‘zbek tili va adabiyoti” jurnali
8. “Filologiya masalalari” jurnali
9. “O‘zbekiston tilshunosligi” jurnali



10. O'zbek tilshunoslarining ilmiy jurnali - (<https://www.tilshunoslik.uz>)
11. S. Datta Investigating English-Language Dialect-Adjusted Models. A Thesis Submitted to the Faculty in partial fulfillment of the requirements for the degree of Bachelor of Arts in Computer Science by Samiha Datta Advised by Professor Soroush Vosough. Dartmouth College Dartmouth Digital Commons. June 2023. p95
12. A.Melikyan Technology for creating digital explanatory dictionaries. E3S Web of Conferences 210, 02003 (2020)Don State Technical University, 344003, Gagarin Sq., 1, Rostov-on-Don, Russia 2020. p11
13. www.language.uz
14. www.tilshunoslik.uz