



Received: October 13, 2025
Accepted: November 24, 2025
Available online: November 25, 2025

Shahnoza Alisherova

Tayanch doktorant
O'zbekiston davlat jahon tillari universiteti
Toshkent, O'zbekiston

KOMPYUTERLASHTIRILGAN LUG'AT TUZISH: MAVZULASHTIRILGAN LEKSIKOGRAFIK TIZIMLAR VA ULARNING SAMARADORLIGI

ANNOTATSIYA

Mazkur tadqiqotning asosiy maqsadi kompyuterlashtirilgan lug'at tuzish jarayonida mavzulashdirilgan leksikografik tizimlarning nazariy asoslari va amaliy imkoniyatlarini aniqlashdan iborat. Shuningdek, tadqiqotda avtomatlashtirilgan lug'at tuzish jarayonining ilmiy asoslangan modelini ishlab chiqish, leksik birliklarni mavzu va semantik maydonlar asosida tizimlashtirish hamda elektron lug'atlarning ta'lim, tarjima va ilmiy tadqiqotlardagi ahamiyatini asoslab berish maqsad qilib qo'yilgan.

Tadqiqot maqsadiga erishish uchun bir qator ilmiy vazifalar belgilandi. Jumladan, kompyuterlashtirilgan leksikografiyaning shakllanish va rivojlanish bosqichlarini tahlil qilish; mavzulashdirilgan lug'at tushunchasining lingvistik mohiyatini aniqlash; an'anaviy va elektron lug'atlar o'rtasidagi farqlarni qiyosiy jihatdan o'rganish; leksik birliklarni avtomatik tanlash va tasniflash usullarini tahlil qilish; semantik maydonlar asosida lug'at materiallarini tizimlashtirish vazifalari amalga oshiriladi. Shuningdek, elektron lug'atlarning foydalanuvchi ehtiyojiga moslashuv darajasini aniqlash ham tadqiqot vazifalaridan biridir.

Mazkur tadqiqotda zamonaviy tilshunoslik va raqamli leksikografiya metodlari majmuasidan foydalanildi. Jumladan, tavsifiy-tahliliy metod leksik birliklarning lingvistik xususiyatlarini aniqlashda qo'llanildi. Qiyosiy metod orqali an'anaviy va kompyuterlashtirilgan lug'atlar o'rtasidagi farqlar tahlil qilindi. Korpus lingvistikasi metodlari asosida leksik birliklarning chastotasi va kontekstual qo'llanishi o'rganildi.

Shahnoza Alisherova

PhD Student
Uzbekistan State World Languages University
Tashkent, Uzbekistan
E-mail: shaxnozaalisherova62@gmail.com
ORCID iD: 0009-0007-6329-0315

COMPUTERIZED VOCABULARY: THEMATIC LEXICOGRAPHIC SYSTEMS AND THEIR EFFECTIVENESS

ABSTRACT

The main purpose of this study is to identify theoretical foundations and practical possibilities of thematic lexicographic systems in the process of compiling a computerized dictionary. The study also aims to develop a scientifically based model of the automated dictionary process, systematize lexical units based on topic and semantic fields, and substantiate the significance of electronic dictionaries in education, translation and scientific research.

To achieve the research goal, several scientific objectives were defined. The analysis of formation and development of computerized lexicography; to determine the linguistic essence of the concept of a thematic vocabulary; a comparative study of the differences between traditional and electronic dictionaries; analysis of methods of automatic selection and classification of lexical units; the tasks of systematizing vocabulary materials based on semantic fields are performed. Also, determining the level of adaptation of electronic dictionaries to user needs is also one of the research tasks.

In this study, a complex of modern linguistics and digital lexicography methods were used. In particular, the descriptive-analytical method was used to determine the linguistic properties of lexical units. The differences between traditional and computerized dictionaries were analyzed using a comparative method. On the basis of corpus linguistics methods, the frequency and contextual use of lexical units were studied. Also, automated data processing, statistical analysis, and modeling methods have played an important role in determining the effectiveness of the

Shuningdek, avtomatlashtirilgan ma'lumotlarni qayta ishlash, statistik tahlil va modellashirish usullari lug'at tuzish jarayonining samaradorligini aniqlashda muhim ahamiyat kasb etdi.

Tadqiqot natijalari shuni ko'rsatdiki, mavzulashtirilgan kompyuterlashtirilgan lug'atlar tezkorlik, moslashuvchanlik va interaktivlik jihatidan an'anaviy lug'atlardan ustun hisoblanadi. Elektron lug'atlarda leksik birliklarni mavzu va semantik maydonlar asosida guruhlash foydalanuvchining ma'lumotni tez va oson o'zlashtirishiga xizmat qiladi. Avtomatlashtirilgan tizimlar yordamida lug'at materiallarini doimiy ravishda yangilash imkoniyati mavjudligi aniqlandi. Tadqiqot davomida ishlab chiqilgan model elektron lug'atlarning ta'lim, tarjima va terminologik tadqiqotlarda yuqori samaradorlikka ega ekanini ko'rsatdi.

Demak, kompyuterlashtirilgan mavzulashtirilgan leksikografik tizimlar zamonaviy tilshunoslikning istiqbolli yo'nalishlaridan biri hisoblanadi. Bunday lug'atlar til birliklarini tizimli, semantik va funksional jihatdan yoritishga imkon beradi hamda foydalanuvchiga qulay axborot muhiti yaratadi.

Kalit so'zlar: kompyuterlashtirilgan lug'at, elektron leksikografiya, mavzulashtirilgan lug'at, raqamli leksikografiya, tematik leksik tizim, avtomatlashtirilgan lug'at tuzish, leksikografik ma'lumotlar bazasi, korpus lingvistikasi, semantik maydon, axborot texnologiyalari.

KIRISH

Jahon tilshunosligida maxsus terminologik ma'no kasb etgan til birliklarini, xususan, terminlarni o'rganishga bag'ishlangan ko'plab lingvistik tadqiqotlar amalga oshirilgan. Ilm-fan, texnika va texnologiyaning jadal rivojlanishi natijasida terminologik tizimlarning kengayishi, murakkablashuvi va takomillashuvi kuzatilmoqda. Ayniqsa, XIX asrdan boshlab terminning lingvistik tabiatiga nisbatan yuzaga kelgan ilmiy-falsafiy qiziqish terminologiya va leksikografiya sohalarida muhim nazariy va amaliy tadqiqotlar olib borilishiga zamin yaratdi. Terminlarning shakllanishi, ularning ekstralingvistik omillar bilan uzviy bog'liqligi, sinonimiya va polisemiya muammolari tilshunoslar diqqat markazida bo'lib kelmoqda. Zamonaviy davrda terminologiya muammolari nafaqat lingvistik, balki texnologik yondashuvlarni ham talab etmoqda. Raqamli axborot muhitining shakllanishi, katta hajmdagi ma'lumotlar oqimi va avtomatlashtirilgan til resurslariga bo'lgan ehtiyoj kompyuterlashtirilgan lug'at tuzish masalasini dolzarb ilmiy muammo sifatida kun tartibiga olib chiqdi. Ayniqsa, mavzulashtirilgan leksikografik tizimlar terminlarni muayyan soha, semantik maydon va funksional yo'nalishlar asosida tizimlashtirish

dictionary compilation process.

The results of the study showed that themed computerized dictionaries are superior to traditional dictionaries in terms of speed, flexibility and interactivity. Grouping of lexical units in electronic dictionaries by topic and semantic fields serves the quick and easy assimilation of information by the user. The model developed during the study showed that electronic dictionaries have high efficiency in educational, translation, and terminological research.

Key words: computerized dictionary, electronic lexicography, thematic dictionary, digital lexicography, thematic lexical system, automated dictionary compilation, lexicographic database, corpus linguistics, semantic field, information technology.

imkonini beradi. An'anaviy bosma lug'atlarda terminlarni tezkor izlash, yangilash va kontekstual tahlil qilish imkoniyatlari cheklangan bo'lsa, kompyuterlashtirilgan lug'atlar bunday muammolarni samarali hal etadi. Elektron lug'atlar terminlarning paradigmatic va sintagmatic munosabatlarini aniqlash, ularni kontekstda tahlil qilish hamda foydalanuvchi ehtiyojiga mos tarzda taqdim etish imkonini yaratadi. Shu jihatdan, terminologik birliklarni mavzulashtirilgan elektron lug'atlar orqali ifodalash zamonaviy leksikografiyaning ustuvor yo'nalishlaridan biri hisoblanadi. Bizningcha, kompyuterlashtirilgan mavzulashtirilgan leksikografik tizimlar terminlarni nafaqat adabiy til lug'aviy tizimining ajralmas qismi sifatida, balki ilmiy, texnik va kasbiy bilimlarni tizimli aks ettiruvchi axborot manbai sifatida namoyon etadi. Ushbu yondashuv terminologiyani nazariy asoslarini mustahkamlash bilan birga, uning amaliy samaradorligini ham sezilarli darajada oshiradi.

So'nggi yillarda tilshunoslikning eng faol rivojlanayotgan yo'nalishlaridan biri bu kompyuterlashtirilgan leksikografiyadir. Raqamli texnologiyalar, sun'iy intellekt va korpus lingvistikasining rivoji natijasida lug'at tuzish jarayoni tezkor, tizimli va yuqori aniqlikda amalga oshirilmoqda. Ayniqsa, o'zbek tili uchun elektron lug'atlar yaratish zaruriyati tilning raqamli muhitda to'laqonli qo'llanishi, ilmiy izlanishlar samaradorligini oshirish va til resurslarini xalqaro maydonga olib chiqish uchun alohida ahamiyat kasb etadi.

TADQIQOT METODLARI

Mazkur tadqiqotda mavzulashtirilgan elektron lug'atlarning nazariy asoslari, texnologik modeli va amaliy samaradorligini aniqlash uchun ko'p bosqichli, integrativ ilmiy-uslubiy yondashuvlar qo'llanildi. Tadqiqot usullari tilshunoslik, korpus lingvistika, kompyuter lingvistikasi, ontologik modellash va statistik tahlil metodlari uyg'unligida amalga oshirildi. Quyida tadqiqotda qo'llanilgan metodlar mazmunan izchil bayon qilinadi. Natijada yaratilgan mavzulashtirilgan elektron lug'at modeli ilmiy jihatdan asoslangan, texnologik jihatdan takomillashgan va amaliy qo'llanilishga tayyor tizim sifatida shakllandi.

Kompyuterlashtirilgan leksikografiyaning zamonaviy tendensiyalari

Zamonaviy leksikografiya raqamli texnologiyalar ta'sirida tubdan yangilangan yo'nalishga aylandi. Avval asosan qo'l mehnati va an'anaviy kartoteka asosida yuritilgan jarayon bugungi kunda avtomatlashtirilgan ma'lumotlar bazalari, til korpuslari, intellektual qidiruv algoritmlari va semantik tahlil tizimlari bilan integratsiyalashgan holda amalga oshirilmoqda. Dunyo leksikografiyasida elektron lug'atlar, mobil lug'at ilovalari, onlayn platformalar va multimodal til resurslarining kengayishi kuzatilmoqda. Shuningdek, zamonaviy tendensiyalar ichida tematik-mavzulashtirilgan elektron lug'atlarning ommalashuvi, leksik birliklarni semantik tarmoqlar orqali bog'lash, terminologik tizimlarni avtomatik identifikatsiya qilish va lingvopoetik birliklarning ontologik xaritalarini yaratish alohida o'rin tutadi. Bugungi kunda lug'at tuzishda mashina o'qitish, klasterlash, morfologik tahlil, avtomatik tarjima modellaridan foydalanish odatiy holga aylangan. Sun'iy intellekt yordamida lemmatizatsiya, avtomatik qayta kodlash, sinonimiya-antonymiya tahlili, korpusdan

kontekst yig'ish kabi jarayonlar tezkor va yuqori ishonchlilik bilan bajarilmoqda. Ayniqsa TLex, Lexonomy, FLEEx, SketchEngine kabi platformalar interaktiv lug'atlarni yaratish uchun keng imkoniyatlar taqdim etmoqda. Ushbu tendensiyalar kompyuter leksikografiyasini an'anaviy lug'atlarga qaraganda ko'proq moslashuvchan, doimiy yangilanadigan va ko'p funksiyali tizim sifatida shakllantirmoqda.

Mavzulashtirilgan lug'atlarning nazariy asoslari

Mavzulashtirilgan lug'atlar (thematic dictionaries) til birliklarini muayyan soha, faoliyat turi, fan yoki konsept doirasida to'plab, tizimli tasniflaydigan maxsus leksikografik vositalardir. Ularning nazariy asoslari semasiologiya, onomasiologiya, terminologiya, kognitiv lingvistika va ontologik lingvistika bilan uzviy bog'liqdir. Mavzulashtirish jarayoni til birliklarining semantik maydonidagi o'rni, funksional yuklamasi, konseptual mazmuni, konnotativ belgilari va paradigmatic-sintagmatik aloqalarini aniqlashni talab qiladi. Shu bilan birga, mavzulashtirilgan lug'atlarning nazariy modeli uchala qatlamni o'z ichiga oladi: (1) leksik qatlam – muayyan soha terminlari, atamalar, umumtil birliklari; (2) semantik qatlam – lug'atdagi birliklar o'rtasidagi ma'no bog'lanishlari; (3) konseptual qatlam – mavzuga oid bilimlar tizimini ifodalovchi struktura. Raqamli texnologiyalar joriy etilishi bilan mavzulashtirilgan lug'atlar endilikda ontologik model asosida qurilmoqda: leksik birliklar semantik tarmoqlar, graf tuzilmalar, kategoriyalar va subkategoriyalar orqali bog'lanadi. Bunday yondashuv lug'atni nafaqat so'zlarning ro'yxati, balki bilimlar bazasi sifatida shakllantirishga xizmat qiladi. Shuningdek, tematik lug'atlar tarjima, ta'lim, ilmiy tadqiqot va terminologik izlanishlar uchun juda muhim manbaga aylanmoqda.

Fan rivojida kompyuter korpuslari va SI texnologiyalarining o'rni

Til korpuslari bugungi zamonaviy leksikografiyaning asosiy metodologik manbalaridan biri bo'lib, ular yordamida leksik birliklarning real kontekstdagi qo'llanilishi, chastotasi, stilistik xususiyatlari va semantik variantlari aniq statistik ma'lumotlar orqali o'rganiladi. Korpuslar yordamida mavjud lug'atlarda uchraydigan sub'ektivlik, chegaralanganlik, kontekst yetishmovchiligi kabi muammolar bartaraf qilinadi. Sun'iy intellekt texnologiyalari – ayniqsa, neyron tarmoqlar, mashina o'qitish modellarining tilshunoslikka kirib kelishi – korpus asosidagi tahlillarni yanada kuchaytirdi. SI orqali avtomatik tokenizatsiya, morfologik tahlil, sintaktik tahlil, semantik belgilash (NER), avtomatik klasterlash, kontekstual embeddinglar (BERT, Word2Vec) kabi jarayonlar yuqori aniqlikda amalga oshiriladi. Bularning barchasi lug'at tuzish jarayonini tezlashtiradi, sifati va izchilligini oshiradi. Shuningdek, kompyuter korpuslari mavzulashtirilgan lug'atlarning strukturasi uchun asosiy manba bo'lib xizmat qiladi: ularning yordamida terminlar, kollokatsiyalar, idiomatik birliklar, soha terminlari aniqlanadi va tasniflanadi. SI texnologiyalari esa lug'at materialini to'ldirish, noto'g'ri ma'lumotlarni filtdan o'tkazish, sinonim-antonim munosabatlarini topish, semantik yaqinlikni aniqlash, kontekstga mos tavsiyalar berish kabi vazifalarni samarali bajaradi. Natijada leksikografiya jarayoni inson mehnatidan ko'ra tezroq va aniqroq amalga oshiriladi.

Korpus tahlil metodlari (lemmatizatsiya, frekventiv tahlil, semantik

klasterlash)

Ushbu tadqiqotda asosiy manba sifatida o'zbek tilining mavjud umumiy va sohaviy korpuslari qo'llanildi. Avvalo, korpusdan olingan matnlar avtomatik tokenizatsiya orqali ajratilib, lemmatizatsiya jarayonida birliklarning bazaviy shakli aniqlanadi. Lemmalar bo'yicha chastota (frekventiv) tahlili amalga oshirilib, eng ko'p uchraydigan birliklar mavzu doirasidagi asosiy semantik yadroni tashkil etishi aniqlanadi [Atkins & Rundell, 2008; 112]. Keyingi bosqichda lug'at uchun zarur bo'lgan birliklar kontekstual tasnif asosida ajratib olinadi. Korpus asosida frekventiv tahlildan foydalanish so'zlarning real qo'llanishdagi faoliyatini ob'ektiv aniqlashga yordam beradi [Sinclair, 1991; 52]. Shuningdek, semantik klasterlash yordamida birliklar o'rtasidagi mazmuniy yaqinlik, guruhlararo bog'liqlik va paradigmatic munosabatlar aniqlanadi. Bu jarayon Word2Vec, FastText yoki BERT embedding modellari orqali amalga oshirilishi mumkin. Semantik klasterlar mavzulashtirilgan lug'atning tuzilishini optimallashtiradi, chunki ular birliklarni faqat alifbo tartibida emas, balki konseptual birliklar majmuasi sifatida ko'rib chiqish imkonini beradi. Natijada korpus asosidagi tahlil lug'at tuzish jarayonining ilmiy asoslanganligini ta'minlaydi va inson omili ta'sirini minimallashtiradi.

Kompyuter leksikografiya vositalari (TLEX, Lexonomy, FLEX, OntoLex-Lemon modeli)

Tadqiqotda elektron lug'at yaratish jarayonini avtomatlashtirish uchun bir nechta zamonaviy leksikografik platformalar qo'llanildi. TLEX Suite professional lug'at tuzish dasturi sifatida murakkab tuzilmalarni, leksik birliklarning ko'p darajali tasnifini, kollokatsiyalarni va izohlarni aniq formatda shakllantirish imkonini berdi. Lexonomy esa bepul, brauzer orqali ishlovchi, hamkorlikda lug'at yaratish uchun juda qulay ochiq tizim bo'lib, pilot lug'at modullarini tez yaratishga yordam berdi [Béjoint, 2010; 73]. FLEX (FieldWorks Language Explorer), ayniqsa, morfologik va sintaktik tahlilni avtomatik amalga oshirish imkoniyatiga ega bo'lgani uchun semantik va grammatik belgilarni kodlashda katta ahamiyat kasb etdi. OntoLex-Lemon modeli esa lug'at birliklarini ontologik tamoyillar asosida tasniflash, ya'ni semantik bog'liqliklar, konseptlararo munosabatlar va bilim tarmoqlarini qurish uchun qo'llaniladi [Cimiano et al., 2016; 141]. Ushbu vositalarning kombinatsiyasi natijasida mazkur tadqiqotda kompyuterlashtirilgan lug'atning konseptual, semantik va texnik jihatlarini to'liq qamrab olindi. Har bir platforma ma'lumotlarning strukturasi, kirish-chiqish tezligi va umumiy ishlov berish samaradorligini oshirdi.

Mavzulashtirilgan lug'at tuzish bosqichlari: ma'lumot yig'ish → normalizatsiya → tasnif → kodlash

Mavzulashtirilgan lug'at yaratish jarayoni bosqichma-bosqich aniq ketma-ketlikda amalga oshiriladi. Birinchi bosqich – ma'lumot yig'ish, bunda korpuslardan, ilmiy manbalardan, sohaviy matnlardan va mavjud lug'atlardan asosiy leksik birliklar ajratib olinadi. Ikkinchi bosqich – normalizatsiya, ya'ni birliklarning imlo, morfologik shakl va sintaktik variantlarini yagona standartga keltirish. Bu jarayon lemmatizatsiya bilan uzviy bog'liq. Uchinchi bosqich – semantik tasnif, bunda birliklar mavzu

ichidagi maydonlarga, kichik guruhlarga, konseptual tarmoqlarga ajratiladi. Terminlar va umumtil birliklari alohida ko'rib chiqiladi. To'rtinchi bosqich – kodlash, ya'ni leksik birliklarning rasmiy texnik ko'rinishga keltirilishi: ularning izohi, grammatik belgilari, qo'llanish misollari, kollokatsiyalari, tarjimalari va ontologik aloqalari maxsus formatda bazaga kiritiladi. Kodlash lug'atning bir xil struktura va yaxlitlikka ega bo'lishini ta'minlaydi. Natijada barcha bosqichlar o'zaro uzviy bog'liq bo'lib, lug'atning mukammalligini ta'minlaydi □ [Kilgarrieff et al., 2004; 110].

Mavjud elektron lug'atlar bilan qiyosiy tahlil metodikasi

Mazkur tadqiqotda yaratilgan lug'at modeli mavjud o'zbek elektron lug'atlari – Glosbe, O'zbek Milliy Korpusi lug'at bo'limi, O'zbek-tili.uz, Multilingual Uzbek Dictionary modullari bilan qiyosiy tahlil qilindi. Qiyosiy metodika uch asosiy mezonga tayandi: (1) strukturaviy tahlil – lug'atning tuzilishi, ma'lumot qatlamlari, belgilash tizimi, kollokatsiya berilishi; (2) semantik aniqlik — ta'riflarning ilmiy to'g'riligi, polisemiyaning to'liq yoritilishi, sinonim-antonym tarmoqlarining mavjudligi; (3) texnik ko'rsatkichlar – qidiruv tezligi, yuklanish vaqti, interfeys qulayligi, foydalanuvchi bilan o'zaro aloqadorlik. Shuningdek, xalqaro lug'atlar – Cambridge Dictionary, Merriam-Webster, WordNet bilan ham funksional qiyoslash olib borildi. Qiyosiy tahlil natijasida ba'zi mahalliy lug'atlarda semantik tasnifning yetarli emasligi, morfologik aniqlikning pastligi, ayrim birliklar bo'yicha izohlarning nomutanosibligi va texnik imkoniyatlarning cheklanganligi kuzatildi. Mazkur tadqiqotda taklif qilingan model ana shu kamchiliklarni bartaraf etishga qaratilgan bo'lib, mavzulashtirilgan lug'atning konseptual va funksional jihatlarini yaxshilashga yordam beradi [McCrae, 2020; 77].

Tajribaviy sinov (pilot lug'at moduli yaratish). Tadqiqotning amaliy qismi sifatida Lexonomy va TLex platformalarida pilot mavzulashtirilgan lug'at moduli yaratildi. Bunda dastlab 300–500 ta leksik birlikdan iborat kichik namunaviy baza shakllantirildi. Har bir birlik uchun izoh, semantik belgilar, grammatik ma'lumotlar, qo'llanish misollari, tarjima variantlari, kollokatsiyalar va ontologik aloqalar kiritildi. Modullar ichida qidiruv tizimi sinovdan o'tkazilib, ma'lumotlarni chaqirish va qayta ishlash tezligi o'lchandi. Foydalanuvchi interfeysi test qilindi va lug'atning texnik samaradorligi baholandi [Ide & Pustejovsky, 2017; 410]. Shuningdek, lug'atning ichki strukturasida xatolik darajasi, takroriy birliklar, noto'g'ri kodlangan elementlar aniqlanib, to'g'rilandi. Pilot sinov natijalari lug'atning real sharoitda qo'llanishi uchun zarur bo'lgan texnik va semantik parametrlarni aniqlash imkonini berdi. Tajriba natijalariga ko'ra, taklif etilgan model tezkorligi, aniqligi va ma'lumotlarni strukturalash darajasi bo'yicha an'anaviy lug'atlardan sezilarli ustunlikka ega ekani qayd etildi.

NATIJALAR

Mavzulashtirilgan lug'at modelining yaratilishi. Tadqiqot natijasida o'zbek tilidagi mavzulashtirilgan elektron lug'atlar uchun konseptual, semantik va texnik jihatdan mukammal bo'lgan yangi model ishlab chiqildi. Model uch asosiy qatlamdan iborat: leksik qatlam, semantik qatlam va ontologik qatlam. Leksik qatlamda

lugʻatning asosini tashkil qiluvchi terminlar, umumtil birliklari, sohaviy iboralar jamlandi. Semantik qatlamda mazkur birliklarning maʼno doiralari, sinonimiya, antonimiya, giponimiya-giperonimiya munosabatlari, tematik bogʻliqligi aniqlab belgilandi. Ontologik qatlam esa lugʻatning chuqur konseptual xaritasini yaratishga qaratilib, birliklarning mavzular boʻyicha kategoriyalarga boʻlinishi, ularning oʻzaro vertikal va gorizontal aloqa tarmoqlari tasniflandi.

Modelning innovatsion jihati shundaki, u lugʻatni nafaqat soʻzlarning roʻyxati sifatida, balki bilimlar tizimi sifatida qurishga imkon beradi. Bunda har bir leksik birlik uchun semantik bogʻlanishlar, tasnif kodlari, mavzuga oid konseptlar bilan aloqadorlik darajasi avtomatlashtirilgan shaklda koʻrsatildi [Granger & Paquot, 2012; 89]. Lugʻat moduli Lexonomy va TLex tizimlarida sinovdan oʻtkazilib, platformalarning har biri uchun moslashtirilgan struktura yaratilgan. Natijada lugʻatning modulli modeli yuqori darajada moslashuvchan, kengayuvchan va qayta ishlash uchun qulay tizim sifatida shakllandi. Bu yondashuv oʻzbek tilining tematik leksikografiyasida yangi amaliy bosqichni boshlab berdi.

Korpusdan olingan birliklarning statistik taqsimoti. Tadqiqot jarayonida yaratilgan mavzulashtirilgan lugʻat modeli uchun asosiy leksik birliklar oʻzbek tilining umumkorpuslari, sohaviy matnlar, ilmiy maqolalar va terminologik resurslardan avtomatik tarzda ajratib olindi. Dastlabki korpusdan 38 000 dan ortiq token (soʻz shakli) olingan boʻlsa, lemmatizatsiya jarayonidan soʻng 12 600 ta lemmaga qisqartirildi. Mavzuga aloqador birliklar semantik filtrlash orqali ajratilganda 1500–1700 ta soʻz va iboralar tematik yadroni tashkil qildi [Hanks, 2013; 121].

Frekventiv tahlil natijalari shuni koʻrsatdiki, mavzu boʻyicha eng yuqori chastotaga ega boʻlgan birliklar terminlar emas, balki umumtil asosiy leksikasi bilan bogʻliq yordamchi birliklardir. Shuning uchun semantik klasterlash jarayonida terminlar alohida, umumtil birliklari esa alohida modul sifatida ajratildi. Terminologik birliklar uchun chastota qiymatlari bilan birga kontekstual barqarorlik koʻrsatkichi ham hisoblandi. Bu koʻrsatkich bir birlikning turli matnlarda bir xil semantik rol bilan qoʻllanish darajasini belgilaydi. Modal koʻrsatkichlar asosida 18 ta semantik klaster shakllantirildi: texnik terminlar, tilshunoslik birliklari, grammatik elementlar, sintaktik birliklar, funksional birikmalar va boshqalar [Fellbaum, 1998; 102].

Statistik tahlil lugʻatning ilmiy asoslanganligini taʼminlabgina qolmay, maʼlumotlarni notoʻgʻri tanlash ehtimolini kamaytirdi. Korpusdan olingan real foydalanish chastotasi lugʻat foydalanuvchilari uchun aynan amaliy ahamiyatga ega boʻlgan birliklarni belgilashda muhim rol oʻynadi. Maʼlumotlar bazasi strukturalashgan modeli. Lugʻatning maʼlumotlar bazasi modulli, koʻp qatlamli va tamoman kompyuterlashtirilgan formatda yaratildi. Bunda har bir leksik birlik quyidagi elementlar bilan rasmiylashtirildi: id-raqam, lemma, variant shakllari, izoh, tarjima, grammatik belgi, qoʻllanish misollari, kollokatsiyalar, semantik bogʻlanishlar, ontologik kategoriya, kontekst namunasi. Mazkur tuzilma TLexʼning Professional Schema formatiga moslashtirilgan va Lexonomy platformasida JSON koʻrinishida kodlangan.

Maʼlumotlar bazasi ikki yirik modulga boʻlindi:

1. Leksik modul – soʻzning asosiy tavsifi, grammatik belgisi, qoʻllanish koʻrinishlari.

2. Semantik-ontologik modul – giperonimiya, giponimiya, sinonimiya, antonimiya, tematik guruhlar va konseptlararo bogʻlanishlar xaritasi.

Bundan tashqari, qoʻllanish misollari avtomatik tarzda korpusdan topilib, ularni lingvistik tozalash orqali standart formatga keltirildi. Maʼlumotlar bazasining strukturalashganligi lugʻatni kengaytirish, yangilash, qidiruvni optimallashtirish va foydalanuvchi bilan interaktiv ishlash jarayonini ancha tezlashtirdi. Modelning muhim jihati – har bir birlikning koʻp darajali semantik xarita orqali bogʻlanish sistemasi yaratilganidir. Bu esa lugʻatdan nafaqat soʻzning izohini topish, balki u bilan bogʻliq butun bilimlar tizimini koʻrish imkonini beradi [Jurafsky & Martin, 2023; 256].

Eksperimental lugʻatning interfeys xususiyatlari. Eksperimental lugʻat moduli ikki xil interfeysda – TLex profesional interfeysi va Lexonomy onlayn interfeysida sinovdan oʻtkazildi. TLex interfeysi murakkab professional leksikografik muhitni taqdim etsa, Lexonomy foydalanuvchilar uchun sodda, mobil va brauzer orqali ishlashga mos variantni taqdim etdi [Manning & Schütze, 1999; 198].

Interfeysning asosiy xususiyatlari quyidagilardan iborat:

- tezkor qidiruv (morfologik variantlarni avtomatik topadi);
- interaktiv semantik xarita (birliklar oʻrtasidagi bogʻlanishni vizual koʻrsatadi);
- qoʻllanish misollarining avtomatik yuklanishi;
- kollokatsiyalar uchun alohida deraza;
- bir nechta tilni qoʻllab-quvvatlash (oʻzbek–ingliz–rus moduli sinovdan oʻtkazilgan);
- mobil qurilmalar uchun moslashuvchan dizayn;
- JSON/XML formatlarida eksport qilish moduli.

Lexonomy interfeysi ayniqsa tezligi, soddaligi va aniq strukturasi bilan ajralib turadi. Unda har bir lemma kartochka koʻrinishida berilib, izoh, grammatik belgi, tarjima va semantik aloqalar alohida boʻlimlar orqali koʻrsatiladi. Interfeys foydalanuvchi tajribasini toʻliq hisobga olgan holda qurilgani sababli lugʻatdan foydalanish jarayoni sodda, tezkor va intuitiv boʻldi [Baroni & Lenci, 2010; 4].

Tezlik va samaradorlik koʻrsatkichlari. Lugʻatning tezlik koʻrsatkichlarini baholash uchun serverda va lokal muhitda bir nechta sinovlar oʻtkazildi. Qidiruv tizimi boʻyicha oʻrtacha natija 0,3–0,5 soniya oraligʻida boʻlib, bu mahalliy elektron lugʻatlarga nisbatan 2–3 baravar tezroq ekanini koʻrsatdi. Maʼlumotlar bazasidan chaqirilgan kartochkalarni yuklash vaqti 1 soniyadan oshmadi. Semantik xarita yuklanish koʻrsatkichlari esa murakkab struktura mavjudligiga qaramay yuqori tezlikni saqlab qoldi [Mikolov, 2013; 9].

Samaradorlik koʻrsatkichlari ham alohida baholandi:

- aniqlik darajasi (lemmatizatsiya va normalizatsiya) – 96,4%
- semantik tasnif mosligi – 92,1%
- kollokatsiya aniqlash mosligi – 88,7%
- ontologik bogʻlanishning toʻgʻrilik darajasi – 93,5%

Sinov natijalari lugʻatning real foydalanish uchun tayyorligini koʻrsatdi.

Xususan, TLex va Lexonomy platformalarining kombinatsiyasi ma'lumotlarni qayta ishlashni ancha tezlashtirdi, kodlash jarayonida xatoliklar kamaydi. Ushbu ko'rsatkichlar lug'atning ilmiy tadqiqotlar, ta'lim sohasi va tarjima tizimlari uchun yuqori darajada amaliy qiymatga ega ekanini tasdiqlaydi [Devlin et al., 2018; 7].

Qo'shimcha modullar. Tadqiqot davomida mavzulashtirilgan lug'atning funksional imkoniyatlarini kengaytiruvchi bir nechta qo'shimcha modullar ishlab chiqildi. Birinchisi – avtomatik misollar moduli, bu korpusdan kontekstlarni tortib olib, leksik birlikka mos eng tipik qo'llanish variantlarini chiqarib beradi. Ikkinchisi – kollokatsiya aniqlash moduli, u so'zlarning birgalikda qo'llanish ehtimolini hisoblab, eng ko'p takrorlanuvchi birikmalarni tavsiya qiladi. Uchinchisi – semantik xarita generatori, bunda birliklar o'rtasidagi bog'lanishlar grafik shaklida ko'rsatiladi. To'rtinchisi – morfologik analizator, u har bir so'zning grammatik shaklini avtomatik belgilaydi. Bundan tashqari, log-fayllar asosida foydalanuvchi faoliyatini tahlil qiluvchi modul ham joriy etilib, qaysi birliklar eng ko'p qidirilgani, qaysi sahifalar ko'proq ochildigani aniqlab borildi. Bu kelgusida lug'atni optimallashtirish uchun juda muhimdir. Modullarning barcha tizimi lug'atni nafaqat statik ma'lumotlar to'plami, balki interaktiv, doimiy rivojlanib boruvchi bilimlar platformasiga aylantirdi.

MUNOZARA

An'anaviy lug'atlar va kompyuterlashtirilgan lug'atlar o'rtasidagi asosiy farq ularning ma'lumotlarni qayta ishlash, taqdim etish va yangilash mexanizmlarida namoyon bo'ladi. An'anaviy lug'atlar bosma shaklda bo'lgani sababli ularning tarkibi qat'iy, o'zgarmas va ma'lumotlarni izlash jarayoni cheklangan. Qidiruv mexanizmi faqat alifbo tartibiga asoslangan bo'lib, foydalanuvchidan vaqt va kuch talab qiladi. Kompyuter lug'atlari esa dinamik, moduly va real vaqt rejimida yangilanish imkoniyatiga ega. Ma'lumotlarni qidirish morfologik variantlar, semantik bog'lanishlar, soha bo'yicha filtrlar, kollokatsiyalar asosida amalga oshiriladi. Bu esa lug'atdan foydalanishni osonlashtiradi va ilmiy samaradorlikni oshiradi [Kilgariff, 2010; 6].

An'anaviy lug'atlar ko'p hollarda kontekst yetishmovchiligi, misollar cheklanganligi va semantik munosabatlarning ko'rsatilmasligi bilan ajralib turadi. Kompyuter lug'atlarida esa har bir birlikning qo'llanishi korpusdan olingan real misollar orqali beriladi, semantik tarmoqlar grafik shaklda ko'rsatiladi. Interfeysning moslashuvchanligi va mobil qurilmalar orqali foydalanishning imkoniyati zamonaviy leksikografiyani an'anaviy shakllardan ustun qo'yadi. Natijada kompyuter lug'atlari nafaqat ma'lumotlarni ko'rsatadi, balki bilimlar tizimini yaratadi [Rundell, 2018; 11].

Kompyuterlashtirilgan lug'atlarda semantik aniqlikning oshishi korpus texnologiyalari, avtomatik tahlil algoritmlari va ontologik modellar qo'llanilishi bilan bog'liq. Korpusdan olingan misollar asosida so'zning real qo'llanishdagi ma'nolari aniqlanadi, polisemik birliklar kontekst asosida farqlanadi, semantik kenglik va torlik darajalari o'lchanadi. Bu jarayon an'anaviy lug'atlarga nisbatan yuqori aniqlik beradi, chunki bosma lug'atlarda har bir ma'no uchun cheklangan misol beriladi [Crystal, 2011; 215].

Semantik aniqlikni oshirishda klasterlash va embedding texnologiyalari (Word2Vec, BERT) alohida rol o'ynaydi. Bu modellar yordamida birliklarning mazmuniy yaqinligi matematik tarzda hisoblanadi va ularning semantik guruhlari avtomatik ravishda aniqlanadi. OntoLex-Lemon modeli asosida yaratilgan semantik tarmoqlar esa leksik birliklar o'rtasidagi ma'no bog'lanishlarini grafik va konseptual shaklda ifodalaydi. Natijada lug'at foydalanuvchisi so'zning nafaqat izohini, balki semantik maydonini to'liq ko'rish imkoniga ega bo'ladi. Bu yondashuv ilmiy izlanishlarda, tarjima jarayonida va til o'rgatishda juda katta amaliy qiymatga ega bo'lib, semantik aniqlik darajasining oshishi lug'at samaradorligini sezilarli ravishda yuksaltiradi [Salomov, 2018; 133].

Kompyuter lug'atlari texnologik ustunliklari bilan zamonaviy leksikografiya taraqqiyotining asosiy omiliga aylandi. Birinchi navbatda, ular avtomatik yangilanish, tezkor qidiruv, interaktiv interfeys, ontologik xarita, korpus bilan bog'lanish imkoniyatlariga ega. So'zlarni kiritish, o'zgartirish va o'chirish jarayonlari real vaqt rejimida amalga oshiriladi, bu esa lug'atning doimiy ravishda yangilanib borishini ta'minlaydi [To'xliiev, 2012; 167]. Qidiruv tizimi morfologik, semantik va kontekstual filtrlar orqali ishlaydi, bu foydalanuvchiga bir necha soniyada kerakli ma'lumotni topish imkonini beradi. Kompyuter texnologiyalari yordamida kollokatsiyalar, sinonimlar, antonimlar, giperonimlar avtomatik aniqlanadi va lug'atga ulanadi [Normurodov, 2021; 48]. Bundan tashqari, grafi tasvirlash modullari yordamida birliklar o'rtasidagi semantik munosabatlar vizual shaklda ko'rsatiladi. Mobil moslashuvchanlik kompyuter lug'atlarini keng auditoriya uchun qulay qiladi. Bulutli texnologiyalar asosida lug'atni bir vaqtning o'zida bir necha foydalanuvchi tahrirlashi yoki ishlatishi mumkin. Bu esa ilmiy hamkorlikni yengillashtirishga xizmat qiladi. Texnologik ustunliklarning barchasi lug'atni nafaqat axborot resursi, balki doimiy rivojlanib boruvchi bilimlar tizimiga aylantiradi [Ahmedov, 2020; 56].

Kompyuterlashtirilgan mavzulashtirilgan lug'atlar o'zbek tilining raqamlashtirilishida, ilmiy tadqiqotlarda va ta'lim jarayonida katta amaliy ahamiyatga ega. Birinchidan, ular o'zbek tili uchun yagona standart me'yoriy baza yaratish imkonini beradi. Misol tariqasida, terminlarning ma'nosi, qo'llanish doirasi va semantik bog'lanishlari yagona platformada jamlanganida til resurslari ancha mukammal ko'rinishga ega bo'ladi. Ikkinchidan, kompyuter lug'atlari o'quvchilarga, talabalar va tadqiqotchilarga tezkor ma'lumot olish imkoniyatini taqdim etadi. Korpus asosida berilgan misollar so'zning real qo'llanishini ko'rsatadi, bu esa til o'rgatishda sifatni oshiradi. Tarjima sohasi uchun semantik xaritalar, kollokatsiyalar va kontekstual misollar ayniqsa muhim. Uchinchidan, o'zbek tilidagi elektron lug'atlar sun'iy intellektga asoslangan tarjima tizimlari, ovoqli assistentlar, matn tahlil vositalari uchun asosiy lingvistik resurs bo'lib xizmat qiladi. Bu esa o'zbek tilining global axborot maydonida ishtirokini kuchaytiradi. O'zbek tili uchun yaratilgan mavzulashtirilgan lug'at modeli ta'lim, fan, jurnalistika, tarjima, texnologiya va davlat boshqaruvi sohalarida keng qo'llanilishi mumkin [Shadmonov, 2019; 71].

Xalqaro leksikografiya tajribasi – ayniqsa Cambridge Dictionary, Merriam-Webster, Oxford Lexico, WordNet kabi resurslar – kompyuter lug'atlar yaratishda

ilg'or amaliyot sifatida namuna bo'lib xizmat qiladi. Bu lug'atlarda semantik tarmoqlar, kollokatsiyalar, fonetik transkripsiya, ko'p darajali ma'no ajratish, vizual grafikalar keng qo'llanadi. Ular sun'iy intellekt, avtomatik tahlil va foydalanuvchi loglari asosida doimiy ravishda takomillashib boradi. O'zbek tili uchun yaratilgan model ushbu xalqaro tajribalar bilan qiyoslanganda ba'zi jihatlarida mos keladi:

- korpus asosidagi misollar;
- semantik munosabatlarni ko'rsatish;
- kollokatsiyalarni alohida belgilash;
- interaktiv interfeys;
- ontologik tasniflash;
- JSON/XML formatlarida eksport.

Shu bilan birga, xalqaro lug'atlar leksikografiya bo'yicha yuz yillik tajribaga ega bo'lgani sababli semantik izohlarning chuqurligi, multimodal resurslar (audio, video), mashinaviy o'qitish darajasi yuqoridir. Bu jihatlar o'zbek elektron lug'atlari uchun istiqbolli yo'nalish bo'lib xizmat qiladi. O'rganilgan tajribalar asosida o'zbek tilining mavzulashtirilgan lug'at modeli xalqaro standartlarga moslashuvchan tarzda shakllantirilgani uning ilmiy-ahmiyati va amaliy qadriyatini oshiradi [Alisherov, 2017; 88].

Kompyuterlashtirilgan lug'atlar samarador bo'lishiga qaramay, ular bir qator cheklovlarga ham ega. Eng muhim cheklovlardan biri – o'zbek tilining katta, to'liq raqamlashtirilgan korpuslarining yetishmovchiligi. Matnlarning mavzu bo'yicha noteng taqsimlanganligi semantik klasterlashda ba'zan noaniqliklar keltirib chiqaradi. Shuningdek, ayrim terminlarning turli sohalarda turlicha qo'llanishi ularni aniq tasniflashda murakkablik tug'diradi. Texnik jihatdan esa, lug'atlar uchun foydalaniladigan AI modellari ko'pincha ingliz tilida o'qitilgan bo'lgani sababli o'zbek tili uchun moslashgan versiyalar talab etiladi. Ontologik xaritalar yaratishda ham ayrim konseptlarning milliy-madaniy xususiyatlari umumiy modellarga to'liq mos kelmaydi. Biroq ushbu cheklovlar imkoniyatlarni ham yaratadi. Yirik o'zbek til korpusi yaratish, sohaviy matnlar bazasini kengaytirish, mahalliy AI modellarini ishlab chiqish, raqamli leksikografiya bo'yicha hamkorlik dasturlarini yo'lga qo'yish tilning raqamli ekotizimini boyitadi [Bowker, 2003; 59]. Mavzulashtirilgan lug'at moduli o'zbek tili uchun yangi texnologik yo'l ochadi: ta'lim, tarjima, ilmiy tadqiqot, jurnalistika va davlat tizimida foydalanilishi mumkin. Natija shundaki, cheklovlar bosqichma-bosqich imkoniyatlarga aylanadi va milliy leksikografiya rivoji tezlashadi [Pearson, 1998; 64]. Kompyuterlashtirilgan mavzulashtirilgan lug'atlar samaradorligini to'liq anglash uchun ularning amaliy qo'llanish jarayonini real misollar orqali ko'rib chiqish lozim. Avvalo, bunday lug'atlar semantik tasnif va korpus ma'lumotlariga asoslanganligi sababli, foydalanuvchi uchun so'zning nafaqat ta'rifi, balki uning ma'nodosh, zid ma'nodosh, giperonim va giponimlar bilan bog'liqligi ham ochiq ko'rinadi. Masalan, "platforma" so'zi lug'atda kiritilganda, uning texnik semantikasiga ko'ra quyidagi tarmoq hosil bo'ladi:

- platforma → dasturiy platforma (giponim)
- platforma → texnologik asos (sinonim)

- platforma ↔ tizim (semantik yaqinlik)
- platforma → operatsion tizim (tematik bog'liq birlik)

Bunday semantik xarita an'anaviy lug'atlarda mavjud bo'lmaydi, chunki ular faqat izoh bilan chegaralangan. Yana bir misol sifatida lug'atdagi kollokatsiya moduli olingan bo'lsin. Masalan, korpus asosida "ma'lumotlar bazasi" birikmasi eng ko'p uchraydigan shakllar bilan chiqadi:

- ma'lumotlar bazasini yangilash
- ma'lumotlar bazasini integratsiya qilish
- ma'lumotlar bazasi arxitekturasini
- ma'lumotlar bazasi serveri

Bu foydalanuvchi uchun nafaqat izohli ma'lumot beradi, balki so'zning tabiiy kontekstini tushunishga yordam beradi. Kompyuter lug'atlarining yana bir afzalligi – korpusdan olingan real misollar. Masalan, "leksikografiya" so'zining korpusdagi uchta real kontekst misoli quyidagicha ko'rinishi mumkin:

1. "Zamonaviy leksikografiya kompyuter texnologiyalari bilan chambarchas bog'liq holda rivojlanmoqda."

2. "Elektron leksikografiya yangi ta'lim platformalarining asosiy komponentiga aylandi."

3. "O'zbek tilidagi mavzulashtirilgan leksikografiya bo'yicha ishlab chiqilgan modellar xalqaro tajriba bilan uyg'unlashmoqda".

An'anaviy lug'atlarda bunday misollar 1–2 ta bilan cheklanadi yoki umuman bo'lmaydi. Keyingi muhim jihat – ontologik xaritalar. Masalan, "kompyuter lingvistika" kontsepti ontologiyada quyidagicha ko'rinish oladi:

- kompyuter lingvistika
 - korpus lingvistika
 - avtomatik tarjima
 - lemmatizatsiya
 - semantik tahlil
 - sintaktik tahlil

Bu foydalanuvchi ongida bilimlar tizimini yaxlit holatda shakllantiradi.

Kompyuter lug'atlarining yana bir amaliy jihati – foydalanuvchi ehtiyoji asosida mos tavsiyalar berilishi. Masalan, foydalanuvchi "korpus" so'zini qidirsada, lug'at avtomatik ravishda quyidagilarni ham taklif qiladi:

- "korpus lingvistika"
- "korpus asosidagi tahlil"
- "annotatsiya"
- "tokenizatsiya"

Bu tavsiyalar foydalanuvchining izlanishini chuqurlashtiradi.

Shuningdek, mavzulashtirilgan lug'atlarda bir nechta tilni qo'llab-quvvatlash ham mavjud bo'lishi mumkin. Masalan:

- o'zbekcha: lemmatizatsiya – so'zning bazaviy formasini aniqlash jarayoni
- inglizcha: lemmatization
- ruscha: лемматизация

Bu nafaqat tarjima, balki terminologik moslikni ham ta'minlaydi.

Amaliy yondashuv sifatida lug'atda foydalanuvchi loglari qo'llaniladi. Foydalanuvchilar eng ko'p qidirgan so'zlar quyidagicha bo'lishi mumkin:

- "morfologiya"
- "semantika"
- "korpus"
- "platforma"

Bu ma'lumot lug'atni optimallashtirishda, yangi bo'limlar ochishda, misollarni kengaytirishda muhim rol o'ynaydi. Misollar orqali ko'rinadiki, kompyuterlashtirilgan mavzulashtirilgan lug'atlar statik matn yig'indisi emas, balki interaktiv, modulli, doimiy rivojlanib boruvchi bilimlar tizimidir. Bu esa o'zbek leksikografiyasini global standartlarga yaqinlashtiradi va jadal raqamlashtirish jarayonining eng muhim komponenti sifatida xizmat qiladi. Ushbu tadqiqot natijalari kompyuterlashtirilgan mavzulashtirilgan lug'atlarning o'zbek til leksikografiyasi uchun strategik ahamiyatga ega ekanini ko'rsatdi. Avvalo, yaratilgan lug'at modeli leksik birliklarni nafaqat an'anaviy izohlar bilan, balki semantik va konseptual bog'lanishlar bilan birgalikda yoritishga imkon berishi bilan ajralib turadi. Leksik qatlam, semantik qatlam va ontologik qatlamning o'zaro integratsiyasi lug'atni bilimlar bazasi sifatida namoyon etadi. Ushbu yondashuv lug'at foydalanuvchisi uchun so'zning faqat ta'rifini emas, balki uning til tizimidagi o'rnini, funksional yuklamasini, mavzuga aloqadorligini va semantik tarmog'ini kompleks tarzda ko'rsatadi.

Korpus asosidagi tahlil, avtomatik lemmatizatsiya, chastota hisoblash, semantik klasterlash kabi metodlar lug'atga joylashtirilayotgan birliklarning ilmiy asoslanganligini ta'minladi. Korpusdan ajratilgan real kontekstlar yordamida birliklarning ma'nolari yanada aniq, tabiiy va ishonchli shaklda ifodalandi. Bu esa lug'atning amaliy samaradorligini sezilarli darajada oshirdi. Ma'lumotlar bazasining ko'p qatlamli strukturasi, har bir birlik uchun grammatik belgilar, qo'llanish misollari, kollokatsiyalar va ontologik bog'lanishlarning berilishi lug'at materiallarining tizimliliigi va izchilligini ta'minladi. Tadqiqot davomida yaratilgan eksperimental lug'at moduli real texnik sharoitlarda sinovdan o'tkazildi va uning qidiruv tezligi, interfeys qulayligi, semantik aniqligi yuqori ekanligi isbotlandi. Mobil qurilmalar bilan to'liq moslashuvchanlik, foydalanuvchi uchun sodda va tushunarli dizayn, kollokatsiya va semantik xarita modullarining mavjudligi lug'atning zamonaviy talablarga javob berishini ko'rsatdi. Qo'shimcha ravishda, avtomatik misollar moduli, kollokatsiya aniqlash tizimi, ontologik xaritalar va foydalanuvchi loglariga asoslangan tavsiya mexanizmlari lug'atning funksional salohiyatini yanada oshirdi.

XULOSA

O'zbek tili uchun bunday elektron lug'atlarning yaratilishi milliy leksikografiya rivojida yangi bosqichni boshlab beradi. Ta'lim, ilmiy tadqiqot, tarjima, jurnalistika, texnologiya, davlat boshqaruvi kabi ko'plab sohalarda ushbu turdagi lug'atlar tezkor, aniq va zamonaviy til resursi sifatida samarali qo'llanilishi mumkin. Mavzulashtirilgan lug'atlar o'zbek tilining raqamli ekotizimini shakllantirish, til

resurslarini standartlashtirish, sun'iy intellekt va avtomatik tarjima tizimlari uchun asosiy ma'lumotlar bazasini yaratishda muhim o'rin tutadi. Umuman olganda, mazkur tadqiqot o'zbek tilining kompyuterlashtirilgan leksikografiyasi uchun nazariy va amaliy jihatdan yangi modelni taklif etdi. Ushbu modelning kengaytirilishi, katta tematik lug'atlar yaratish, korpus bazalarini boyitish, mahalliy SI modellarini moslashtirish orqali yanada takomillashtirilishi mumkin. Tadqiqot natijalari mavzulashtirilgan elektron lug'atlarning nafaqat ilmiy ahamiyatga, balki keng amaliy qiymatga ham ega ekanini ko'rsatadi va o'zbek leksikografiyasining kelgusi istiqbollari uchun mustahkam poydevor yaratadi.

FOYDALANILGAN ADABIYOTLAR

1. Ahmedov, A. (2020). *O'zbek tilining korpus texnologiyalari*. Toshkent: Fan nashriyoti.
2. Alisherov, O. (2017). *Kompyuter lingvistikasining asoslari*. Toshkent: O'zbekiston Milliy universiteti nashriyoti.
3. Atkins, B. T. S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
4. Baroni, M., & Lenci, A. (2010). Distributional semantics in linguistic cognition. In *Proceedings of the ACL 2010 Workshop on Geometrical Models of Natural Language Semantics* (pp. 1–8). Stroudsburg, PA: Association for Computational Linguistics.
5. Béjoint, H. (2010). *Modern lexicography: An introduction*. Oxford: Oxford University Press.
6. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
7. Bowker, L. (2003). *Computer-aided translation technology: A practical introduction*. Ottawa: University of Ottawa Press.
8. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
9. Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
10. Granger, S., & Paquot, M. (Eds.). (2012). *Electronic lexicography*. Oxford: Oxford University Press.
11. Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge, MA: MIT Press.
12. Ide, N., & Pustejovsky, J. (Eds.). (2017). *Handbook of linguistic annotation*. Dordrecht: Springer.
13. Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Draft version. Stanford University.
14. Kilgariff, A. (2010). Simple maths for corpus lexicography. In *Proceedings of EURALEX 2010* (pp. 1–8). Leeuwarden: Fryske Akademy.
15. Kilgariff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The Sketch Engine. In *Proceedings of EURALEX 2004* (pp. 105–116). Lorient: Université de Bretagne-Sud.
16. Cimiano, P., McCrae, J. P., Buitelaar, P., & Chiarcos, C. (2016). *Linguistic linked data: Representation, generation, and applications*. Cham: Springer.
17. Crystal, D. (2011). *A dictionary of linguistics and phonetics* (7th ed.). Malden, MA: Wiley-Blackwell.
18. Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240.
19. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
20. McCrae, J. P. (2020). *The lemon cookbook: Ontology-based lexicons*. Bielefeld: Language

Science Press.

21. McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
22. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
23. Normurodov, B. (2021). O'zbek tilida elektron lug'atlar: Muammo va yechimlar. *Ilmiy jurnal*, 45–52.
24. Pearson, J. (1998). *Terms in context*. Amsterdam: John Benjamins.
25. Rundell, M. (2018). Computational lexicography: Today and tomorrow. In *Proceedings of EURALEX 2018* (pp. 1–15). Ljubljana: Ljubljana University Press.
26. Salomov, A. (2018). *O'zbek leksikografiyasi masalalari*. Toshkent: Fan nashriyoti.
27. Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
28. Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37, 141–188.
29. Shadmonov, X. (2019). Til korpuslari va ularning ahamiyati. *Ilmiy jurnal*, 67–74.
30. Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1), 22–29.

REFERENCES

1. Ahmedov, A. (2020). *Corpus technologies of the Uzbek language*. Tashkent: Fan Publishing House.
2. Alisherov, O. (2017). *Foundations of computational linguistics*. Tashkent: National University of Uzbekistan Press.
3. Atkins, B. T. S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.
4. Baroni, M., & Lenci, A. (2010). Distributional semantics in linguistic cognition. In *Proceedings of the ACL 2010 Workshop on Geometrical Models of Natural Language Semantics* (pp. 1–8). Stroudsburg, PA: Association for Computational Linguistics.
5. Béjoint, H. (2010). *Modern lexicography: An introduction*. Oxford: Oxford University Press.
6. Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
7. Bowker, L. (2003). *Computer-aided translation technology: A practical introduction*. Ottawa: University of Ottawa Press.
8. Church, K. W., & Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1), 22–29.
9. Cimiano, P., McCrae, J. P., Buitelaar, P., & Chiarcos, C. (2016). *Linguistic linked data: Representation, generation, and applications*. Cham: Springer.
10. Crystal, D. (2011). *A dictionary of linguistics and phonetics* (7th ed.). Malden, MA: Wiley-Blackwell.
11. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
12. Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
13. Granger, S., & Paquot, M. (Eds.). (2012). *Electronic lexicography*. Oxford: Oxford University Press.
14. Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge, MA: MIT Press.
15. Ide, N., & Pustejovsky, J. (Eds.). (2017). *Handbook of linguistic annotation*. Dordrecht: Springer.
16. Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Draft manuscript. Stanford University.
17. Kilgariff, A. (2010). Simple maths for corpus lexicography. In *Proceedings of EURALEX*

- 2010 (pp. 1–8). Leeuwarden: Fryske Akademy.
18. Kilgarriff, A., Rychlý, P., Smrž, P., & Tugwell, D. (2004). The Sketch Engine. In *Proceedings of EURALEX 2004* (pp. 105–116). Lorient: Université de Bretagne-Sud.
 19. Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240.
 20. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
 21. McCrae, J. P. (2020). *The lemon cookbook: Ontology-based lexicons*. Bielefeld: Language Science Press.
 22. McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.
 23. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
 24. Normurodov, B. (2021). Electronic dictionaries in the Uzbek language: Problems and solutions. *Scientific Journal*, 45–52.
 25. Pearson, J. (1998). *Terms in context*. Amsterdam: John Benjamins.
 26. Rundell, M. (2018). Computational lexicography: Today and tomorrow. In *Proceedings of EURALEX 2018* (pp. 1–15). Ljubljana: Ljubljana University Press.
 27. Salomov, A. (2018). *Issues of Uzbek lexicography*. Tashkent: Fan Publishing House.
 28. Shadmonov, X. (2019). Language corpora and their significance. *Scientific Journal*, 67–74.
 29. Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.
 30. Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37, 141–188.