ISNN 3030-3907

2026
MAY
ISSUE - 5
VOLUME - 3

DEVELOPMENT OF SCIENCE

scientific journal



Development of science

Ilmiy jurnal

MAY 2026
Issue - 5, VOLUME - 3

ISSN 3030 -3907

Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo‘yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro‘yxatiga kiritilgan.

BOSH MUHARRIR:
Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.
TAHRIRIYAT KENGASH RAISI:
Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor
TAHRIRIYAT KENGASH A’ZOLARI

TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ

Bozorov G‘ayrat Rashidovich
Buxoro Davlat Texnika Universiteti, professor
Abdurahmonov Olim Rustamovich
Buxoro Davlat Texnika Universiteti, professor
Adizov Bobur Zamirovich
O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor
Hayitov Ruslan Rustamjonovich
Buxoro Davlat Texnika Universiteti, professor
Jumayev Qayum Karimovich
Buxoro Davlat Texnika Universiteti, professor.
Berdiyev Dorob Murodovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor
Maxmudov Rafiq Amonovich
Buxoro Davlat Texnika Universiteti, professor.
Tilloyev Lochin Ismatilloyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Raxmonov Qahramon Sanakulovich
Buxoro Davlat Texnika Universiteti, dotsent.
Barakayev Nusratullo Rajabovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.
Jo‘rayev Xayrullo Fayziyevich
Buxoro Davlat Texnika Universiteti, professor.
Panoyev Erali Rajabboyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Ataullayev Sherzod Nabiullyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Malik Negmurodovich
Buxoro Davlat Texnika Universiteti, dotsent.
Axmedova Ozoda Bahronovna
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.
Safarov Bahri Jumayevich
Buxoro Davlat Texnika Universiteti, dotsent.
Xaydarov Axtam Amonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Safarov Jasur Alijon o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.

KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ

Axmedov Voxid Nizomovich
Buxoro Davlat Texnika Universiteti, professor
Umarov Baqo Bafoyevich
Buxoro Davlat Universiteti, professor
Kodirov Abduaxad Abduraximovich
Qarshi davlat Universiteti, professor
Qurbonov Mingniqul Jumagulovich
Qarshi davlat Universiteti, dotsent
Turemuratov Sharibay Naurizbaevich
O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor
Tursunov Kaxor Shonazarovich
Qarshi davlat Universiteti, professor
Do‘stov Hamro
Buxoro Davlat Texnika Universiteti, professor
Panoyev Nodir Shavkatovich
Buxoro Davlat Texnika Universiteti, dotsent.
Olimov Bobur Bahodir o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Kasimov Sherzod Abduazimovich
Termiz Davlat Universiteti, professor.
Aliqulov Rustam Valiyevich
Termiz Davlat Universiteti, professor
Tropov Xamza Tursunovich
Samarqand Davlat Universiteti, professor

IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ

Nizamov Aslitdin Badreddinovich
Buxoro Davlat Texnika Universiteti, professor.
Boboyev Akmal Choriyevich
Buxoro Davlat Texnika Universiteti, professor.
Ergashxojayeva Shaxnoza Djasurovna
Toshkent davlat iqtisodiyot universiteti, professor
Saidaxmedova Nodira Ilxomovna
Toshkent davlat iqtisodiyot universiteti, professor
Muxsinov Bekzod Toxirovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shakarov Qulmat Ashirovich
Renessans ta'lim universiteti,dotsent.
Jumaeva Zulfiya Qayumovna
Buxoro Davlat Texnika Universiteti, dotsent.
Norova Salomat Yusupovna
Buxoro Davlat Texnika Universiteti, dotsent.
Avezova Shaxnoza Maxmudjonovna
Buxoro Davlat Texnika Universiteti, dotsent.
Jalilov Jamshid G‘anijonovich
Toshkent davlat iqtisodiyot universiteti, dotsent
Xasanova Gulruh Djumanazarovna
Buxoro Davlat Texnika Universiteti, dotsent.

PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ

Siddiqova Sadoqat G‘afforovna
Buxoro Davlat Texnika Universiteti, dotsent.
Zoyirov Erkin Xalilovich
Buxoro Davlat Texnika Universiteti, DSc, professor
Gaffarov Feruz Hasanovich
Puchon Universiteti, *professor*
Ruziyev Davron Yuldashevich
Buxoro Davlat Universiteti, professor
Muradova Firuza Rashidovna
Buxoro Davlat Texnika Universiteti, professor.
Yunusova Gulandon Samiyevna
Buxoro Davlat Texnika Universiteti, professor.
To‘xtayeva Zebo Sharifovna
Buxoro Davlat Texnika Universiteti, professor.
Navro‘zova Gulchixra Nigmatovna
Buxoro Davlat Texnika Universiteti, dotsent.
Uzoqova Laylo Polvonovna
Buxoro muhandislik – texnologiya instituti, dotsent.
To‘raqulova Marjona Qiyom qizi
Buxoro Davlat Texnika Universiteti, dotsent.
Muradova Zarina Rashidovna
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Sanjar Aslonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shodiyev Jahongir Jo‘raqulovich
Buxoro Davlat Texnika Universiteti, dotsent
Gaffarov Laziz Hasanovich
Buxoro muhandislik – texnologiya instituti, dotsent.

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.

ISSN 3030-3907

Development of science ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.

ISSN 3030-3907

O‘ZBEK TILI UCHUN SEMANTIK QIDIRUV TIZIMINI ISHLAB CHIQISHNING ALGORITMIK VA DASTURIY ASOSLARI

Jo‘rakulov Tolib Toxirovich
Navoiy davlat universiteti dotsenti
Email: jamshidbektoxirov2021@gmail.com

Annotatsiya. Mazkur maqolada o‘zbek tili uchun semantik qidiruv tizimini ishlab chiqishning algoritmik va dasturiy asoslari tadqiq etiladi. O‘zbek tili morfologik jihatdan boy, leksik-semantik munosabatlari murakkab va raqamli lingvistik resurslari nisbatan cheklangan tillardan biri hisoblanadi. Tadqiqotda o‘zbek tili milliy korpusi, elektron korpus, internet axborot matnlari korpusi, mualliflik korpusi, sinonimlar bazasi, idiomalar bazasi, semantik teglar va veb-korpusga oid ilmiy ishlar tahlil qilindi.

Kalit so‘zlar: semantik qidiruv, o‘zbek tili, kam resursli til, NLP, korpus lingvistikasi, embedding, FAISS, Python, semantik moslik.

Аннотация. В данной статье исследуются алгоритмические и программные основы разработки семантической поисковой системы для узбекского языка. Узбекский язык относится к морфологически богатым языкам со сложными лексико-семантическими отношениями и относительно ограниченными цифровыми лингвистическими ресурсами. В рамках исследования были проанализированы научные работы, посвящённые национальному корпусу узбекского языка, электронному корпусу, корпусу интернет-текстов, авторскому корпусу, базе синонимов, базе идиом, семантическим тегам и веб-корпусу.

Ключевые слова: семантический поиск, узбекский язык, язык с малым объёмом ресурсов, NLP, корпусная лингвистика, векторное представление, FAISS, Python, семантическое сходство.

Abstract. This article investigates the algorithmic and software foundations for developing a semantic search system for the Uzbek language. The Uzbek language is considered to be morphologically rich, with complex lexical-semantic relationships and relatively limited digital linguistic resources. The study analyzes scholarly works related to the national corpus of the Uzbek language, an electronic corpus, a corpus of internet information texts, an author's corpus, a synonym database, an idiom database, semantic tags, and a web corpus.

Keywords: semantic search, Uzbek language, low-resource language, NLP, corpus linguistics, embedding, FAISS, Python, semantic similarity.

KIRISH.

Raqamli axborot hajmining keskin ortib borishi tabiiy tilni qayta ishlash, matnlarni avtomatik tahlil qilish va qidiruv tizimlarining semantik imkoniyatlarini kengaytirishni dolzarb vazifaga aylantirmoqda. Oddiy kalit so‘zga asoslangan

qidiruv tizimlarida foydalanuvchi kiritgan so‘zlar hujjat tarkibida aynan uchrashi asosiy mezon sifatida qaraladi. Bunday yondashuv morfologik shakllari xilma-xil, sinonimik qatlamlari keng va kontekstga bog‘liq ma’nolari kuchli bo‘lgan tillarda, xususan o‘zbek tilida, yetarli natija bermaydi.

Semantik qidiruvning asosiy maqsadi foydalanuvchi so‘rovi va hujjatlar o‘rtasidagi yuzaki leksik moslikni emas, balki ularning mazmuniy yaqinligini aniqlashdan iborat. Sentence Transformers hujjatlar va so‘rovlarni yagona vektor fazoga joylashtirib, eng yaqin embeddinglarni topish orqali semantik qidiruvni amalga oshirishini ko‘rsatadi; bunday qidiruv sinonimlar, qisqartmalar va ayrim imlo xatolarida ham kalit so‘zli qidiruvga nisbatan mosroq natija berishi mumkin.

O‘zbek tili uchun semantik qidiruv tizimini yaratish masalasi nafaqat dasturiy muammo, balki lingvistik ta’minot, korpus bazasi, morfologik tahlil, leksik-semantik teglash, sinonimik munosabatlar va sohaviy terminlar bilan bevosita bog‘liq ilmiy-amaliy masaladir. Shuning uchun ushbu maqolada semantik qidiruv tizimi o‘zbek tili korpusshunosligi bo‘yicha mavjud ilmiy tadqiqotlar bilan bog‘langan holda yoritiladi.

ADABIYOTLAR TAHLILI.

O‘zbek tili uchun semantik qidiruv tizimini ishlab chiqish mavjud korpusshunoslik tadqiqotlari bilan uzviy bog‘liq. Chunki semantik qidiruvda matnlarning shunchaki elektron shakli emas, balki ularning lingvistik, morfologik, semantik va kontekstual belgilari ham muhim ahamiyat kasb etadi. O‘zbek tili uchun semantik korpuslar yaratish va ulardan foydalanishni takomillashtirish masalalariga oid tadqiqotlar mamlakatimizda va xorijda bir qator olimlar tomonidan olib borilgan. Jumladan, G.I.Toirova, N.Z. Abduraxmonova, O.X.Abdullayeva, Sh.Xamroyeva [1-4], O‘zbek tili ta’limiy korpusini tuzishning lingvistik asoslari bo‘yicha O‘. Xoliyorov, D. Ahmedova, N.S. G‘ulomova [5-7] amaliy ishlar, N Reimers., I.Gurevych, E.Kuriyozov, D.Vilares, C.Gómez-Rodríguez [13-14] kabi olimlar tomonidan olib borilgan.

Tahlillardan ko‘rinadiki, o‘zbek tili bo‘yicha mavjud tadqiqotlar korpus yaratish, annotatsiyalash, parsing, sinonimlar, idiomalar, neologizmlar va semantik teglar bazasini shakllantirish masalalarini chuqur yoritgan. Biroq ushbu resurslarni yagona semantik qidiruv tizimi doirasida birlashtirish, foydalanuvchi so‘rovi va hujjat mazmuni o‘rtasida embedding asosida semantik moslikni hisoblash hamda Python asosida ishlaydigan dasturiy prototip yaratish alohida ilmiy-amaliy muammo sifatida namoyon bo‘ladi.

TADQIQOT METODOLOGIYASI.

Tadqiqot materiali sifatida o‘zbek tilidagi matnlar korpusi nazarda tutiladi. Korpus quyidagi manbalardan shakllantirilishi mumkin:

- ilmiy maqolalar;

ISSN 3030-3907

•yangilik va internet axborot matnlari;
•ta’limiy matnlar;
•badiiy va mualliflik matnlari;
•sohaviy terminlar va neologizmlar;
•sinonimlar, idiomalar va semantik teglar bazasi.

Korpus matnlari JSON, CSV yoki ma’lumotlar bazasi shaklida saqlanishi mumkin. Har bir hujjat quyidagi atributlarga ega bo‘lishi maqsadga muvofiq:

```
{
  "id": "doc_001",
  "title": "O‘zbek tili korpusida semantik teglash",
  "text": "Matn mazmuni...",
  "source": "maqola",
  "domain": "kompyuter lingvistikasi",
  "tags": ["korpus", "semantika", "NLP"]
}
```

Semantik qidiruv algoritmi

Taklif etilayotgan tizim quyidagi bosqichlarda ishlaydi:

1. **Matnni yig‘ish** - korpusga hujjatlar kiritiladi.
2. **Oldindan qayta ishlash** - apostrof, imlo, katta-kichik harf, tinish belgilari, stop-so‘zlar va morfologik shakllar normallashtiriladi.
3. **Lingvistik boyitish** - sinonimlar, idiomalar, neologizmlar va semantik teglar orqali hujjat tavsifi kengaytiriladi.
4. **Embedding yaratish** - matnlar vektor ko‘rinishiga o‘tkaziladi.
5. **FAISS indekslash** - vektorlar tezkor qidiruv indeksiga joylashtiriladi.
6. **So‘rovni qayta ishlash** - foydalanuvchi so‘rovi ham embeddingga aylantiriladi.
7. **Semantik moslikni hisoblash** - kosinus yaqinlik yoki ichki ko‘paytma asosida eng yaqin hujjatlar aniqlanadi.
8. **Gibrid reytinglash** - embedding natijasi, kalit so‘z mosligi, sinonimik moslik va semantik teglar asosida yakuniy reyting chiqariladi.

SBERT modeli BERT arxitekturasini semantik mazmunli gap embeddinglarini hosil qilishga moslashtiradi va kosinus o‘xshashlik orqali gaplar o‘rtasidagi semantik yaqinlikni hisoblash imkonini beradi. Reimers va Gurevych tadqiqotida SBERT oddiy BERT/RoBERTa bilan eng o‘xshash juftlarni izlashdagi hisoblash xarajatini keskin kamaytirishi ko‘rsatilgan.

O‘zbek tili uchun transformer modellaridan foydalanishda BERTbek kabi tilga xos modellar muhim ahamiyatga ega. BERTbek tadqiqotida o‘zbek tili kam resursli til sifatida ko‘riladi va BERT asosidagi oldindan o‘qitilgan modellar

ISSN 3030-3907

sentiment tahlili, ko‘p yorliqli mavzu tasnifi va nomlangan obyektlarni aniqlash vazifalarida baholangan.

MATEMATIK MODEL

Har bir hujjat matni quyidagi ko‘rinishda vektorlashtiriladi:

$$D_i \to \vec{d}_i$$

Foydalanuvchi so‘rovi esa:

$$Q \to \vec{q}$$

Semantik o‘xshashlik quyidagi formula orqali aniqlanadi:

$$sim(Q, D_i) = \frac{\vec{q} \cdot \vec{d}_i}{\|\vec{q}\| \|\vec{d}_i\|}$$

Yakuniy reyting quyidagicha aniqlanishi mumkin:

$$Score(Q, D_i) = \alpha S_{emb} + \beta S_{lex} + \gamma S_{tag}$$

Bu yerda:

- S_{emb} - embedding asosidagi semantik yaqinlik;
- S_{lex} - kalit so‘z va sinonimlar asosidagi moslik;
- S_{tag} -semantik teglar asosidagi moslik;
- α, β, γ - vazn koeffitsiyentlari.

Dasturiy ta’minot modeli

Tizim arxitekturasi

Taklif etilayotgan dasturiy ta’minot quyidagi modullardan tashkil topadi [1-jadval]:

1-jadval.

Dasturiy ta’minot quyidagi modular va vazifalari

Modul	**Vazifasi**
CorpusLoader	korpus fayllarini yuklash
TextPreprocessor	matnni tozalash va normallashtirish
SemanticAdapter	sinonim, idioma, neologizm va semantik teglar bilan boyitish
EmbeddingModel	matnlarni vektorlashtirish
FaissIndexer	vektorlarni indekslash va tezkor qidirish
Ranker	gibrid reytinglash
API	foydalanuvchi so‘rovini qabul qilish va natijani qaytarish

FAISS kutubxonasi zich vektorlar ichida samarali similarity search va clustering bajarish uchun ishlab chiqilgan bo‘lib, Python wrapperlariga ham ega. FAISS indeks tuzilmasi qurilgach, yangi vektor berilganda eng yaqin vektorlarni topish vazifasini bajaradi.

ISSN 3030-3907

Namunaviy korpus fayli


```
[
  {
    "id": "doc_001",
    "title": "O‘zbek tili milliy korpusini yaratish masalalari",
    "text": "Mazkur hujjatda o‘zbek tili milliy korpusini yaratishning nazariy va amaliy jihatlari yoritiladi.",
    "source": "ilmiy maqola",
    "domain": "korpus lingvistikasi",
    "tags": ["korpus", "o‘zbek tili", "lingvistik baza"]
  },
  {
    "id": "doc_002",
    "title": "Semantik qidiruv tizimlarida embedding modellar",
    "text": "Semantik qidiruv tizimlarida matnlar vektor fazoga o‘tkazilib, mazmuniy yaqinlik asosida qidiriladi.",
    "source": "texnik maqola",
    "domain": "NLP",
    "tags": ["semantik qidiruv", "embedding", "sun'iy intellekt"]
  },
  {
    "id": "doc_003",
    "title": "O‘zbek tilida sinonimlar bazasidan foydalanish",
    "text": "Sinonimlar bazasi foydalanuvchi so‘rovini kengaytirish va qidiruv natijalarini yaxshilashga xizmat qiladi.",
    "source": "ilmiy maqola",
    "domain": "semantika",
    "tags": ["sinonim", "semantik moslik", "qidiruv"]
  }
]
```


NATIJALAR.

Tadqiqot natijasida o‘zbek tili uchun semantik qidiruv tizimining algoritmik va dasturiy modeli ishlab chiqildi. Taklif etilgan model oddiy kalit so‘zli qidiruvdan farqli ravishda so‘rov va hujjatlar o‘rtasidagi mazmuniy yaqinlikni aniqlashga yo‘naltirilgan. Modelda matnni oldindan qayta ishlash, sinonimik kengaytirish, semantik embedding yaratish, FAISS orqali indekslash va gibrid reytinglash jarayonlari bir butun tizim sifatida qaraldi.

Taklif etilgan tizim quyidagi imkoniyatlarga ega [2-jadval]:

ISSN 3030-3907

2-jadval.

Tizim imkoniyatlari va mazmuni

Imkoniyat	Mazmuni
Semantik moslikni aniqlash	So‘rov va hujjat mazmuni embedding orqali solishtiriladi
Sinonimlar bilan ishlash	So‘rov kengaytirilib, yaqin ma’noli hujjatlar topiladi
Tezkor qidiruv	FAISS orqali katta hajmdagi vektorlar ichidan tez natija olinadi
O‘zbek tili xususiyatlarini hisobga olish	Apostrof, stop-so‘z, morfologik shakl va semantik teglar e’tiborga olinadi
Kengaytirish imkoniyati	Idiomalar, neologizmlar, parsing va semantik teglar bazasi qo‘shilishi mumkin

Natijaviy modelda foydalanuvchi “o‘zbek tilida mazmuniy izlash tizimi” deb so‘rov kiritsa, tizim faqat shu so‘zlar aynan uchragan hujjatlarni emas, balki “semantik qidiruv”, “axborot topish”, “matnlararo mazmuniy moslik”, “embedding asosidagi izlash” kabi yaqin mazmundagi hujjatlarni ham chiqarishi mumkin.

MUHOKAMA VA NATIJALAR.

O‘zbek tili uchun semantik qidiruv tizimini ishlab chiqishda asosiy murakkabliklar tilning agglutinativ tuzilishi, so‘z shakllarining ko‘pligi, apostrofli harflarning turlicha yozilishi, sinonimik qatlamning kengligi, idiomatik birliklarning ko‘pligi va sohaviy neologizmlarning tez yangilanib borishi bilan bog‘liq.

Adabiyotlar tahlili shuni ko‘rsatadiki, o‘zbek tili bo‘yicha korpus yaratish, annotatsiyalash, sinonimlar bazasi, idiomalar bazasi, parsing, semantik teglash va veb-korpusga oid tadqiqotlar yetarli darajada shakllanib bormoqda. Biroq mazkur resurslarni semantik qidiruv tizimining yagona dasturiy arxitekturasiga integratsiya qilish hali alohida ilmiy-amaliy vazifa bo‘lib qolmoqda.

Taklif etilgan tizimning ilmiy yangiligi quyidagilarda namoyon bo‘ladi:

1. o‘zbek tili korpusshunosligi bo‘yicha mavjud ilmiy ishlar semantik qidiruv tizimi arxitekturasi bilan bog‘landi;
2. qidiruv jarayoni leksik moslikdan embedding asosidagi mazmuniy moslikka o‘tkazildi;
3. sinonimlar, idiomalar, neologizmlar va semantik teglar qidiruvni boyituvchi alohida lingvistik qatlam sifatida qaraldi;
4. Python, Sentence Transformers va FAISS asosida amaliy dasturiy prototip yaratish imkoni taklif qilindi;

ISSN 3030-3907

5. o‘zbek tili uchun gibrid reytinglash formulasi ishlab chiqildi.

BERTbek kabi o‘zbek tiliga xos modellar kelgusida ushbu tizimga chuqurroq integratsiya qilinsa, qidiruv natijalarining lingvistik mosligi yanada oshishi mumkin. Shu bilan birga, TF-IDF va kosinus o‘xshashlik kabi klassik metodlar ham baseline sifatida muhimdir. Masalan, o‘zbek ta’limiy matnlarini avtomatik tasniflashga oid tadqiqotlarda TF-IDF va kosinus o‘xshashlik yondashuvlari baholash uchun foydali asos bo‘lishi ko‘rsatilgan.

XULOSA VA TAKLIFLAR.

Mazkur maqolada o‘zbek tili uchun semantik qidiruv tizimini ishlab chiqishning algoritmik va dasturiy asoslari tadqiq qilindi. Tadqiqotda o‘zbek tili milliy korpusi, elektron korpus, internet axborot matnlari korpusi, mualliflik korpusi, parsing, sinonimlar bazasi, idiomalar bazasi, leksik-semantik teglash, ta’limiy korpus va veb-korpusga oid ilmiy ishlar tahlil qilindi.

Tahlil natijalari shuni ko‘rsatdiki, o‘zbek tili uchun semantik qidiruv tizimini yaratishda faqat dasturiy algoritm yetarli emas. Bunday tizim samarali ishlashi uchun lingvistik resurslar, korpus bazasi, semantik teglar, sinonimlar, idiomalar va sohaviy terminlar integratsiyalashgan bo‘lishi kerak.

Python asosida taklif etilgan dasturiy prototip o‘zbek tilidagi matnlarni embeddingga aylantirish, FAISS orqali indekslash va foydalanuvchi so‘roviga mazmunan yaqin hujjatlarni topish imkonini beradi. Kelgusida ushbu tizimni BERTbek, morfologik analizator, parsing moduli, semantik teglar bazasi va veb-interfeys bilan boyitish orqali to‘liq funksional o‘zbek tili semantik qidiruv platformasiga aylantirish mumkin.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. Toirova G.I. O‘zbek tili milliy korpusini yaratishning nazariy va amaliy masalalari //Filologiya fanlari doktori (DSc) dissertatsiyasi avtoreferati. – Buxoro, 2021. – 54 b.
2. Abduraxmonova N.Z. O‘zbek tili elektron korpusining kompyuter modellari. //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2021. – 48 b.
3. Abdullayeva O.X. O‘zbek tilining internet axborot matnlari korpusini shakllantirishning nazariy va amaliy asoslari //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2022. – 50 b.
4. Xamroyeva Sh. O‘zbek tili mualliflik korpusini tuzishning lingvistik asoslari //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2018. – 46 b.
5. Raxmanova A. O‘zbek tili Milliy korpusini yaratishda kompyuter usullari //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2022. – 42 b.

ISSN 3030-3907

6. Xidirov O. Milliy korpus uchun parsing dasturi yaratishning lingvistik asoslari //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2021. – 44 b.
7. Begmatova G. O‘zbek milliy korpusida idiomalar bazasini yaratish //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2021. – 40 b.
8. Eshmo‘minov A.A. O‘zbek tili milliy korpusining sinonim so‘zlar bazasi. //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2019. – 44 b.
9. Xoliyorov O‘. O‘zbek tili ta’limiy korpusini tuzishning lingvistik asoslari //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2018. – 48 b.
10. Ahmedova D. Atov birliklarini o‘zbek tili korpuslari uchun leksik-semantik teglashning lingvistik asos va modellari //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2018. – 40 b.
11. G‘ulomova N.S. Alisher Navoiy mualliflik korpusi va uning semantik teglari bazasini yaratish //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2022. – 42 b.
12. Yulbarsov O.O. O‘zbek tili veb korpusi yordamida sohaviy neologizmlar bazasini yaratish //Filologiya fanlari bo‘yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati. – Toshkent, 2025. – 48 b. Reimers N.,
13. Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. EMNLP-IJCNLP, 2019.
14. Kuriyozov E., Vilares D., Gómez-Rodríguez C. BERTbek: A Pretrained Language Model for Uzbek. SIGUL, 2024.