2	Harbiy-taktik tizim	chovush, turg‘oq	Jangovar shiddat, harbiy intizom va xavfsizlik ruhi	Funksional nisbiy lakuna
3	Etnografik va maishiy	sarpo, abzal, ayil, yugan	Milliy etiket, mehmondo‘stlik, shaxsiy qadriyat bo‘yog‘i	Etnografik mutlaq lakuna
4	Diniy-marosim va etiket	fotiha, duoyixayr	Ilohiy himoya, ruhiy sadoqat, taqdir in’ikosi	Konseptual va uslubiy lakuna

Xulosa va takliflar (conclusion/recommendations). Muhammad Alining “Ulug‘ saltanat” epopeyasi matnida lingvomadaniy lakunallik hodisasini til nazariyasi aspektida tadqiq etish natijasida quyidagi fundamental ilmiy xulosalarga kelindi:

–tarixiy proza matnidagi lakunalar shunchaki iste’moldan chiqqan arxaizm yoki istorizmlar emas, balki muayyan bir davr (Temuriylar davri) va milliy etnosning lisoniy dunyo manzarasini, ijtimoiy-madaniy kodlarini o‘zida saqlovchi va matn tuzilishini shakllantiruvchi kognitiv-tushunchaviy birliklar hisoblanadi;

–asar diskursidan aniqlangan lakunalar tartibsiz leksik to‘plam bo‘lmay, balki til tizimida mikrotizimlarni (davlatchilik, harbiy-taktik, etnografik va diniy-marosim maydonlarni) hosil qilishi asoslandi. Bu maydonlar badiiy matn ichida o‘zaro valentlik aloqalariga kirishib, asarning umumiy lingvomadaniy imkoniyatlarini yuzaga chiqaradi.

Matn tarkibi bilan bog‘liq va konnotativ tahlillar shuni ko‘rsatdiki, yasoq, turg‘oq, chovush, sarpo kabi birlamchi lakunalar badiiy jumlalar ichida shunchaki axborot tashuvchi vazifasini bajarmaydi. Ular personajlar muloqotida mutloq hokimiyat, ijtimoiy hamjamiyat, jangovar shiddat va milliy etiket kabi yuqori emotsional-ekspressiv va estetik uslubiy bo‘yoqlarni ta’minlaydi. Tarixiy asarlar tarkibidagi lisoniy bo‘shliqlarni tadqiq etish o‘zbek tili so‘z xazinasining tarixiy taraqqiyot dinamikasini hamda til va milliy mentalitet munosabatlarining asosiy tamoyillarini yoritishda muhim o‘rin tutadi.

Kelgusida o‘zbek tarixiy prozasidagi lingvomadaniy lakunalarning to‘liq elektron tezaurus-lug‘atlarini yaratish hamda ularni kompyuter lingvistikasi qoliplariga solib, milliy korpus bazasiga integratsiya qilish amaliy tilshunoslik uchun istiqbolli yo‘nalishlardan hisoblanadi.

Foydalanilgan adabiyotlar ro‘yxati (References):

(1). Muhammad Ali. Ulug‘ saltanat. Roman-epopeya. Toshkent, “Sharq”, 2013, 640-bet.

(2). Мунен Ж., Винай Ж., Дарбельне Ж. Вопросы теории перевода и сопоставительной лингвистики. Москва, “Прогресс”, 1980, с. 45–48.

(3). Сорокин Ю.А., Марковина И.Ю. Текст и культура: Проблемы межкультурного общения. Москва, “Наука”, 1989, с. 87–91.

(4). Nurmonov A. Tanlangan asarlar. 3-jild, Toshkent, “Akademnashr”, 2012, 44–46-betlar.

(5). Раҳматуллаев Ш. Ўзбек тилининг изоҳли фразеологик луғати. Тошкент, “Ўқитувчи”, 1978, 18–22-betlar.

(6). Mahmudov N. Tilning mukammal tadqiqi yo‘llarini izlab... “O‘zbek tili va adabiyoti”, Toshkent, 2012, 5-son, 3–16-betlar.

(7). Худойберганова Д. Матннинг антропоцентрик тадқиқи. Монография, Тошкент, “Фан”, 2013, 112–115-betlar.

Abduqaxxorov Sardorbek Akramjon o‘g‘li (Andijon davlat Chet tillari instituti dotsenti;
e-mail: sardorbek.akramovich@gmail.com)

O‘ZBEK TILI ENSIKLOPEDIK KORPUSINING LINGVISTIK ARXITEKTURASI VA METHONTOLOGY ASOSIDA KONSEPTUAL ONTOLOGIYANI SHAKLLANTIRISH BOSQICHLARI

Annotatsiya. *Maqolada o‘zbek tili ensiklopedik korpusining klassik lingvistik korpusdan farqlanuvchi ko‘p qatlamli lingvistik arxitekturasi nazariy jihatdan asoslangan. Korpusning markaziy birligi token yoki gap emas, balki bilim bloki, maqola va soha-moduli ekanligi ko‘rsatilgan. Methontology metodologiyasi asosida konseptual ontologiyani shakllantirishning ketma-ket bosqichlari – spetsifikatsiya, konseptuallashtirish, formalizatsiya, implementatsiya, baholash va hujjatlashtirish – ishlab chiqilgan. Bilim sohalarining UO‘K asosida taqsimlanishi va kross-soha aloqalari yoritilgan.*

Kalit so‘zlar: *ensiklopedik korpus, lingvistik arxitektura, methontology, konseptual ontologiya, annotatsiya, UO‘K, OWL/RDF.*

ЭТАПЫ ФОРМИРОВАНИЯ КОНЦЕПТУАЛЬНОЙ ОНТОЛОГИИ НА ОСНОВЕ ЛИНГВИСТИЧЕСКОЙ АРХИТЕКТУРЫ ЭНЦИКЛОПЕДИЧЕСКОГО КОРПУСА УЗБЕКСКОГО ЯЗЫКА

Аннотация. *В статье обосновывается многоуровневая лингвистическая архитектура энциклопедического корпуса узбекского языка, отличающая его от классического лингвистического корпуса. Показано, что центральной единицей такого корпуса является не токен или предложение, а блок знаний, статья или предметный модуль. На основе методологии Methontology разработаны последовательные этапы формирования концептуальной онтологии: спецификация, концептуализация, формализация, реализация, оценка и документирование. Также в статье представлена классификация предметных областей, основанная на Универсальной десятичной классификации (УДК), и описаны межпредметные связи.*

Ключевые слова: *энциклопедический корпус, лингвистическая архитектура, мефонтология, коцептуальная онтология, аннотация, UO'K, OWL/RDF.*

STAGES OF CONSTRUCTING A CONCEPTUAL ONTOLOGY BASED ON THE LINGUISTIC ARCHITECTURE OF THE ENCYCLOPEDIC CORPUS OF THE UZBEK LANGUAGE

Annotation. *The article theoretically substantiates the multi-layered linguistic architecture of the Uzbek encyclopedic corpus, distinct from the classical linguistic corpus. It is shown that the central unit of the corpus is not the token or sentence but the knowledge block, entry and domain module. Based on the Methontology methodology, the sequential stages of forming the conceptual ontology – specification, conceptualisation, formalisation, implementation, evaluation and documentation – are developed. The distribution of knowledge domains by UDC and cross-domain links are examined.*

Keywords: *encyclopedic corpus, linguistic architecture, Methontology, conceptual ontology, annotation, UDC, OWL/RDF.*

Kirish (Introduction). O'zbek tili ensiklopedik korpusini yaratish muammosi matnlarni yig'ish vazifasidan ancha keng bo'lib, til birliklari, bilim tizimlari va raqamli axborot o'rtasidagi aloqalarga bog'liqdir. Bunday korpusda maqola matnlari, atamalar, atoqli nomlar, tarixiy va madaniy ma'lumotlar, faktlar, iqtiboslar hamda manbalar yagona axborot maydonida jamlanadi.

Biroq bu majmua faqat hajmning kattaligi bilan ilmiy qiymatga ega bo'la olmaydi. Uning ahamiyati har bir birlikning til shakli, lug'aviy ma'nosi, grammatik xususiyati, matndagi vazifasi va bilimlar tizimidagi o'rni o'zaro bog'langan holda izohlanishida namoyon bo'ladi. Ana shunday lingvistik arxitektura ensiklopedik korpusni oddiy matnlar to'plamidan ilmiy tahlil va raqamli bilim almashinuvi uchun xizmat qiladigan mukammal manbaga aylantiradi.

Ensiklopedik korpus – korpus tushunchasining alohida turi bo'lib, u klassik lingvistik korpuslardan bir qator parametrlar bilan farqlanadi: lingvistik korpusning birlamchi maqsadi til hodisalarini empirik o'rganish, ensiklopedik korpusniki esa tushunchalar tizimi haqidagi bilimni tilshunoslik xususiyatlari bilan birga taqdim etishdir. D.Biber, S.Conrad va R.Reppen korpus tipologiyasida "maxsuslashgan korpus" toifasini ajratadi; ensiklopedik korpus shu toifaga kiruvchi, ammo ikki yoqlamali lingvistik va bilim-tashkiliy funksiyaga ega kategoriyadir.

Maqolaning maqsadi o'zbek tili ensiklopedik korpusining lingvistik arxitekturasini klassik korpusdan farqlovchi belgilar asosida asoslash va Methontology metodologiyasi asosida konseptual ontologiyani shakllantirish bosqichlarini ishlab chiqishdan iborat.

Metodlar (Methods). Tadqiqotda korpus, ontologik modellashtirish va qiyosiy tahlil metodlari qo'llanildi. Korpus dizayni D.Biber, S.Conrad, R.Reppen tipologiyasi va T.McEnery, A.Hardie funksional tasnifi asosida olib borildi. Ontologiya yaratish jarayoni Noy-McGuinness amaliy tartibi va Methontology metodologiyasi asosida modellashtirildi.

Klassik lingvistik korpus va ensiklopedik korpus yetti parametr (maqsad, matn tabiati, reprezentativlik, annotatsiya yo'nalishi, foydalanuvchi profili, strukturaviy birlik, yangilanish tipi) bo'yicha qiyoslandi. Bilim sohalari UO'K (Universal o'nlik klassifikatsiyasi) asosida taqsimlandi; annotatsiya darajalari R. Garside, G.Leech, T.McEnery andazasida tizimlashtirildi.

Natijalar (Results). Qiyosiy tahlil ensiklopedik korpusning o'ziga xos arxitekturaviy yechimlarni talab qilishini ko'rsatdi, uning strukturaviy birligi alohida token yoki gap emas, balki bilim bloki bo'lib, u o'z navbatida maqola va soha-modul darajalariga birlashadi. E.Atwell ko'rsatganidek, korpus arxitekturasi ko'p qatlamli tizim bo'lib, ensiklopedik korpus uchun u olti qatlamdan iborat: birlamchi matn, lingvistik annotatsiya, tarkibiy-strukturaviy, ontologik-konseptual, bibliografik metama'lumot va foydalanuvchi interfeysi qatlamlari.

Annotatsiya korpus arxitekturasining markaziy tushunchasi sifatida aniqlandi; R.Garside, G.Leech va T.McEnery uni korpusning haqiqiy ilmiy qiymatini belgilovchi asosiy parametr deb baholaydi. Ensiklopedik korpus annotatsiyasi sakkiz darajadan iborat: tokenlashtirish, lemmatizatsiya, morfologik, sintaktik (Universal Dependencies), semantik (NER, FrameNet), ontologik (OWL/RDF/SKOS), bibliografik (Dublin Core) va diskursiv (bilim blok turi) darajalar.

Bilim sohalari UO'K asosida muvozanatli taqsimlandi: ba'zi sohalar (8 – tilshunoslik va adabiyot; 9 – tarix; 7 – san'at) milliy materiallar bilan boy ta'minlangan, boshqalari (5 – tabiiy fanlar; 6 – texnika) ko'p hollarda o'zlashma atamalar bilan to'ldirilgan. Soha modullari o'rtasida kross-soha aloqalari belgilanishi zarurligi ko'rsatildi, chunki real dunyoda bilim sohalari bir-biriga zich bog'langan.

Methontology metodologiyasi asosida konseptual ontologiyani shakllantirishning oltita bosqichi ishlab chiqildi. Birinchi bosqich – spetsifikatsiya: ontologiyaning maqsadi, qamrovi, foydalanuvchi profili va kompetensiya savollari aniqlanadi (masalan, "XV asr o'zbek shoirlari kim?"). Ikkinchi bosqich – konseptuallashtirish: tushunchalar ro'yxati, ierarxiyasi, atributlar va munosabatlar tuziladi.

Uchinchi bosqich – formalizatsiya: konseptual model yarim-formal mantiqiy ifodaga o'tkaziladi (masalan, "shoir – kamida bitta she'riy asar yaratadigan inson"). To'rtinchi bosqich – implementatsiya: ontologiya OWL/RDF tilida, Protégé vositasida kodlanadi. Beshinchi bosqich – baholash: to'liqlik, izchillik va qo'llanuvchanlik mezonlari bo'yicha tekshiriladi. Oltinchi bosqich – hujjatlashtirish: har bir tushuncha va munosabatning ma'nosi, kelib chiqishi qayd etiladi.

O'zbek tili ensiklopedik korpusi uchun amaliy bosqichlar belgilandi: "O'zbek tilining izohli lug'ati" asosida bazaviy leksik-konseptual ro'yxat tuzish, "O'zbekiston milliy ensiklopediyasi" asosida sohalar bo'yicha modullashtirilgan konseptual model yaratish, OWL/RDF formatida implementatsiya, DBpedia va Wikidata bilan integratsiya hamda uzluksiz texnik xizmat ko'rsatish.

Muhokama (Discussion). Ensiklopedik korpusda bir tushuncha turli sohalarda turli ma'nolarga ega bo'lishi (P.Bouquet kontekstual ontologiyalari) muhim metodologik masaladir. Masalan, "ildiz" tushunchasi botanikada o'simlik organi, matematikada funksiya qiymati, tilshunoslikda so'z asosini ifodalovchi morfema sifatida talqin qilinadi. korpusning ontologik tashkilida polisemiya holatlari aniq belgilanishi lozim. M.Hepp ta'kidlaganidek, ontologiyani har qanday sohada cheksiz batafsillashtirib bo'lmaydi – amaliy maqsadlardan kelib chiqib chegara qo'yiladi.

O'zbek tili agglutinativ til bo'lganligi sababli morfologik annotatsiya alohida ahamiyatga ega: bir so'z shaklida bir nechta morfema-grammatik ma'no birlashadi (kitob+lar+imiz+da). Avtomatik morfologik analizator yaratish korpus lingvistikasining markaziy texnik vazifasidir. B.Bahriddinova ta'kidlaganidek, sifatli morfologik analizator yaratish lokal ensiklopedik korpus arxitekturasining birinchi va asosiy texnik shartidir.

G.Sampson ta'kidlaganidek, annotatsiya nafaqat texnik vazifa, balki muayyan tilshunoslik nazariyasining timsolidir. Shu sababli o'zbek ensiklopedik korpusi annotatsiya sxemasini tanlashda o'zbek tilshunosligining mahalliy maktablari hisobga olinishi lozim, aks holda korpus chet til konseptual sxemalarining nusxasi bo'lib qoladi. Sh.Safarov ta'kidlaganidek, milliy konseptual sxema lokal til–madaniyat an'analari va xalqaro standartlarni muvofiqlashtirilgan tarzda taqdim etishi kerak.

Ontologiyaning tashqi resurslar bilan integratsiyasi alohida muhim bosqichdir. Linked Data tamoyillari asosida o'zbek ensiklopedik korpusining ontologiyasi DBpedia va Wikidata kabi xalqaro bilim bazalariga ulanishi mumkin. L.Ehrlinger va W.Wöß ta'kidlaganidek, zamonaviy ontologiya o'zining alohida mavjudligi bilan emas, balki ko'p qatlamli bilim infrastrukturasidagi muvofiqlashgan vazifasi bilan qiymatga ega.

Konseptual ontologiyani shakllantirish bosqichlari shu tariqa nafaqat texnik, balki metodologik va madaniy ahamiyatga ega bo'lgan murakkab loyihaviy faoliyatdir. Uning natijasi sifatida o'zbek tili ensiklopedik korpusi nafaqat lingvistik tadqiqotlar uchun, balki keng ko'lamli ilmiy va amaliy maqsadlar uchun ham tayanch resurs bo'lib xizmat qila oladi.

Xulosa (Conclusion). O'zbek tili ensiklopedik korpusining lingvistik arxitekturasi klassik korpusdan matn tabiati, annotatsiya yo'nalishi, foydalanuvchi profili, strukturaviy birligi va yangilanish usuli bilan farqlanuvchi maxsus tip sifatida asoslandi; uning markaziy birligi token yoki gap bilan chegaralanmay, bilim bloki, maqola va soha-moduliga tayanadi.

Methontology yondashuvi asosida konseptual ontologiyani shakllantirishning ketma-ket bosqichlari — spetsifikatsiya, konseptuallashtirish, formalizatsiya, implementatsiya, baholash va hujjatlashtirish — ilmiy asoslanib, o'zbek tili leksikografiyasi sharoitidagi amaliy yo'riqnomasi yaratildi. Ontologiyaning

D.Bpedia va Wikidata bilan integratsiyasi milliy bilim resurslarini xalqaro miqyosga olib chiqish imkonini beradi.

Foydalanilgan adabiyotlar ro'yxati:

(1). Biber D., Conrad S., Reppen R. Corpus Linguistics: Investigating Language Structure and Use. Cambridge, CUP, 1998.

(2). McEnery T., Hardie A. Corpus Linguistics: Method, Theory and Practice. Cambridge, CUP, 2012.

(3). Noy N., McGuinness D. Ontology Development 101. Stanford, 2001.

(4). Fernández-López M., Gómez-Pérez A. Methontology: From Ontological Art Towards Ontological Engineering. AAAI, 1997.

(5). Garside R., Leech G., McEnery T. Corpus Annotation. London, "Longman", 1997.

(6). Hepp M. Ontologies: State of the Art, Business Potential, and Grand Challenges. Ontology Management. Springer, 2008.

(7). Bird S., Klein E., Loper E. Natural Language Processing with Python (NLTK). O'Reilly, 2009.

(8). Bahriddinova B. O'zbek tili korpusining lingvistik ta'minoti. "O'zbek tili va adabiyoti", 201, 3-son.

(9). Safarov Sh. Kognitiv tilshunoslik. Jizzax, "Sangzor", 2006.

(10). Abdurakhmonova N. Kompyuter lingvistikasi. Darslik, "Globe Edit", 2020, 395-bet.

Hamroyeva Shahlo Mirdjonovna (Toshkent davlat O'zbek tili va adabiyoti universiteti prof.v.b., filologiya fanlari doktori (DSc); shahlo.xamrayeva@navoiy-uni.uz),

Umarova Nozimaxon Qobil qizi (Toshkent davlat O'zbek tili va adabiyoti universiteti 1-bosqich magistratura talabasi; nozimaumarova333@gmail.com)

O'ZBEK TILIDA SAVOL-JAVOB TIZIMINING LINGVISTIK VA TEXNOLOGIK MODELI

Annotatsiya. *Mazkur maqolada o'zbek tilidagi huquqiy diskurs xususiyatlari hamda huquqiy savol-javob tizimlarini yaratishning lingvistik va texnologik asoslari tahlil qilinadi. Tadqiqotda morfologik normalizatsiya, BM25 asosidagi qidiruv, transformer modellarga asoslangan semantik moslashtirish va qoida-asoslangan tekshiruvni birlashtiruvchi gibrid model taklif etiladi. Ushbu model o'zbek tilida huquqiy QA tizimlarining aniqligi, ishonchliligi va izohlanuvchanligini oshirishga xizmat qiladi.*

Kalit so'zlar: *huquqiy diskurs, savol-javob tizimi, huquqiy NLP, lingvistik model, texnologik model, BM25, transformer model, semantik moslik, agglutinativ til, o'zbek tili.*

Хамроева Шахло Мирджоновна (Ташкентский государственный университет узбекского языка и литературы, и.о. профессора, доктор филологических наук (DSc); shahlo.xamrayeva@navoiy-uni.uz),

Умарова Нозимахон Кобил кизи (Ташкентский государственный университет узбекского языка и литературы магистрант 1 курса; nozimaumarova333@gmail.com)

ЛИНГВИСТИЧЕСКАЯ И ТЕХНОЛОГИЧЕСКАЯ МОДЕЛЬ СИСТЕМЫ ВОПРОС-ОТВЕТ НА ОСНОВЕ ЮРИДИЧЕСКОГО ДИСКУРСА УЗБЕКСКОГО ЯЗЫКА

Аннотация. *В статье анализируются особенности юридического дискурса узбекского языка, а также лингвистические и технологические основы создания юридических систем вопрос-ответ. В исследовании предлагается гибридная модель, объединяющая морфологическую нормализацию, поиск на основе BM25, семантическое сопоставление на базе трансформерных моделей и проверку на основе правил. Данная модель способствует повышению точности, надёжности и интерпретируемости юридических QA-систем на узбекском языке.*

Ключевые слова: *юридический дискурс, система вопрос-ответ, юридический NLP, лингвистическая модель, технологическая модель, BM25, трансформерные модели, семантическое соответствие, агглютинативный язык, узбекский язык.*

Khamroyeva Shahlo Mirdjonovna (Tashkent State University of Uzbek Language and Literature Acting Professor, Doctor of Philological Sciences (DSc); shahlo.xamrayeva@navoiy-uni.uz),

Umarova Nozimakhon Qobil qizi (Tashkent State University of Uzbek Language and Literature 1st-year Master's student; nozimaumarova333@gmail.com)

LINGUISTIC AND TECHNOLOGICAL MODEL OF A LEGAL QUESTION-ANSWERING SYSTEM BASED ON UZBEK LEGAL DISCOURSE

Annotation. *This article analyzes the features of Uzbek legal discourse and the linguistic and technological foundations for developing legal question-answering systems. The study proposes a hybrid model that integrates morphological normalization, BM25-based retrieval, transformer-based semantic matching, and rule-based verification. This model contributes to improving the accuracy, reliability, and interpretability of legal QA systems in the Uzbek language.*