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.*

Keywords: *legal discourse, question-answering system, legal NLP, linguistic model, technological model, BM25, transformer models, semantic similarity, agglutinative language, Uzbek language.*

Kirish. Huquqiy savol-javob tizimlarining asosiy vazifasi foydalanuvchi savolini to‘g‘ri talqin qilish, tegishli huquqiy manbani aniqlash va ushbu manbaga asoslangan javobni shakllantirishdan iborat. Shu nuqtayi nazardan, huquqiy QA tizimlari oddiy axborot-qidiruv tizimlaridan farq qiladi. Ular nafaqat hujjatlarni topadi, balki savolning semantik mazmunini tushunadi va unga mos huquqiy normani aniqlaydi. Legal question-answering tizimlari bo‘yicha olib borilgan tadqiqotlarda ushbu yo‘nalish huquqiy axborot bilan ishlashning eng istiqbolli usullaridan biri sifatida baholanadi [2;3–5-b.]. Shuningdek, huquqiy NLP sohasidagi umumlashtiruvchi tadqiqotlarda savol-javob tizimlari huquqiy qidiruv, matn tasnifi va axborot ajratish kabi vazifalar bilan chambarchas bog‘liqligi ko‘rsatilgan [3;10–12-b.].

Mazkur tadqiqotning maqsadi–o‘zbek tilidagi huquqiy diskurs xususiyatlarini hisobga olgan holda savol-javob tizimlari uchun lingvistik va texnologik model ishlab chiqishdan iborat. Ushbu modelda lingvistik tahlil, axborot qidiruv, semantik moslashtirish va qoida-asoslangan tekshiruv bosqichlari integratsiyalashgan holda qaraladi. Taklif etilayotgan yondashuv o‘zbek tilida huquqiy QA tizimlarini yaratishda aniqlik, ishonchlilik va izohlanuvchanlikni oshirishga xizmat qiladi.

Mavzuga oid adabiyotlarning tahlili. Huquqiy savol-javob tizimlari (Legal Question Answering) so‘nggi yillarda tabiiy tilni qayta ishlash sohasining muhim yo‘nalishlaridan biriga aylangan. Ushbu yo‘nalishda olib borilgan tadqiqotlar savolni tushunish, tegishli huquqiy manbani aniqlash va manbaga asoslangan javobni shakllantirish kabi bosqichlarning muhimligini ko‘rsatadi. Martinez-Gil tomonidan olib borilgan tadqiqotda huquqiy QA tizimlari qidiruvga asoslangan va semantik yondashuvlarga ajratilib, ularning har biri alohida ustunlik va cheklovlarga ega ekanligi qayd etilgan [2;3–5-b.]. Huquqiy NLP sohasidagi keng qamrovli tahlillarda esa savol-javob tizimlari huquqiy qidiruv, matn tasnifi va axborot ajratish vazifalari bilan uzviy bog‘liq ekanligi ko‘rsatilgan. Ariai va hammualliflar huquqiy matnlarning murakkabligi, terminologik xususiyatlari va uzun strukturalari avtomatik tahlil jarayonini qiyinlashtirishini ta’kidlaydi [3; 10–12-b.].

Axborot qidiruv tizimlarida BM25 modeli hujjatlar ichidan tegishli segmentlarni aniqlashda samarali yondashuv sifatida keng qo‘llaniladi. Ushbu model statistik asosga ega bo‘lib, hujjatdagi so‘z chastotasi va muhimligini hisobga oladi [7;336–340-b.]. Biroq faqat qidiruvga asoslangan tizimlar semantik jihatdan o‘xshash, lekin turli shaklda ifodalangan matnlarni aniqlashda yetarli natija bermaydi. Shu sababli so‘nggi tadqiqotlarda transformer modellar asosidagi yondashuvlar keng qo‘llanilmoqda. BERT modeli matnning kontekstual ma’nosini aniqlash imkonini berib, savol-javob tizimlarida yuqori natijalar ko‘rsatmoqda [8;172–175-b.]. O‘zbek tilida esa UzBERT va BERTbek modellarining yaratilishi ushbu yondashuvlarni milliy tilga moslashtirish imkonini beradi [5;2–4-b.]. Kam resursli tillar uchun huquqiy QA tizimlarini yaratish bo‘yicha tadqiqotlarda korpus yetishmasligi, morfologik murakkablik va semantik moslikni aniqlash muammolari asosiy to‘siqlar sifatida ko‘rsatiladi [4;5–7-b.]. Shu bois zamonaviy ishlarda qidiruv va semantik modellarni birlashtiruvchi gibrid yondashuvlar eng samarali yechim sifatida taklif etilmoqda.

Tadqiqot metodologiyasi. Mazkur tadqiqotda o‘zbek tilida huquqiy savol-javob tizimini yaratish uchun lingvistik va texnologik yondashuvlarni birlashtiruvchi gibrid model ishlab chiqildi. Tadqiqot metodologiyasi savolni qayta ishlashdan boshlab javobni shakllantirishgacha bo‘lgan bir necha ketma-ket bosqichlarni o‘z ichiga oladi.

Birinchi bosqichda foydalanuvchi savoli lingvistik jihatdan tahlil qilinadi. Bu jarayonda savol tarkibi tozalanadi, ortiqcha belgilar olib tashlanadi va so‘zlar lemmatizatsiya qilinadi. Shuningdek, savolning turi (faktual, normativ yoki izohli) aniqlanadi hamda asosiy huquqiy birliklar ajratib olinadi. O‘zbek tilining agglutinativ xususiyatini hisobga olgan holda morfologik normalizatsiya alohida ahamiyat kasb etadi, chunki bir leksik birlik turli grammatik shakllarda namoyon bo‘lishi mumkin [4;5–7-b.].

Ikkinchi bosqichda huquqiy korpusdan dastlabki mos segmentlarni aniqlash uchun axborot qidiruv yondashuvi qo‘llaniladi. Bu jarayonda BM25 modeli asosiy qidiruv algoritmi sifatida ishlatiladi. Mazkur model hujjatlar ichidan savolga yaqin bo‘lgan matnlarni tez va samarali aniqlash imkonini beradi [7;336–340-b.].

Uchinchi bosqichda semantik moslashtirish amalga oshiriladi. Bu bosqichda transformer modellaridan foydalanilib, savol va topilgan huquqiy segmentlar o‘rtasidagi ma’no yaqinligi aniqlanadi. BERT modeli kontekstga asoslangan semantik ifodalarni yaratish orqali savol-javob tizimlarida yuqori samaradorlikni ta’minlaydi [8;4172–4175-b.]. O‘zbek tili uchun ishlab chiqilgan UzBERT va BERTbek modellar ushbu jarayonni milliy tilga mos holda amalga oshirish imkonini beradi [5;2–4-b.].

To‘rtinchi bosqichda qoida-asoslangan tekshiruv amalga oshiriladi. Bu bosqichda tanlangan javobning huquqiy norma bilan mosligi tekshiriladi va noto‘g‘ri yoki umumiy javoblar filtrlanadi. Ushbu yondashuv huquqiy savol-javob tizimlarida ishonchlilik va aniqlikni oshirishga xizmat qiladi.

Yakuniy bosqichda foydalanuvchiga tushunarli, qisqa va manbaga asoslangan javob shakllantiriladi. Shu tariqa taklif etilgan metodologiya quyidagi ketma-ketlikda ishlaydi:

Savol→lingvistik tahlil→BM25 qidiruv→semantik moslashtirish→tekshiruv→javob. Mazkur yondashuv lingvistik va texnologik komponentlarning integratsiyasini ta’minlab, huquqiy savol-javob tizimlarining aniqligi, ishonchliligi va samaradorligini oshirishga xizmat qiladi.

Tahlil va natijalar. Mazkur tadqiqotda taklif etilgan lingvistik va texnologik model o‘zbek tilida huquqiy savol-javob tizimlarini yaratish nuqtayi nazaridan tahlil qilindi. Modelning samaradorligini baholashda uning turli bosqichlari: lingvistik tahlil, qidiruv, semantik moslashtirish va javobni shakllantirish komponentlari alohida ko‘rib chiqildi.

Lingvistik tahlil bosqichida savol tarkibidagi asosiy huquqiy birliklarni aniqlash va ularni normalizatsiya qilish tizimning keyingi bosqichlari uchun muhim ahamiyat kasb etishi kuzatildi. Ayniqsa, o‘zbek tilining agglutinativ xususiyati sababli lemmatizatsiya va morfologik tahlil natijalari qidiruv samaradorligiga bevosita ta’sir ko‘rsatdi. Bu bosqichda noto‘g‘ri yoki yetarli darajada aniqlanmagan birliklar keyingi semantik moslashtirish bosqichida aniqlikning pasayishiga olib kelishi mumkinligi aniqlandi.

BM25 asosidagi qidiruv bosqichi dastlabki huquqiy segmentlarni tez aniqlashda samarali natija berdi. Ushbu yondashuv katta hajmdagi huquqiy korpusdan savolga yaqin bo‘lgan matnlarni qisqa vaqt ichida ajratib olish imkonini beradi. Biroq tahlil shuni ko‘rsatdiki, BM25 faqat leksik moslikka asoslangani sababli ba’zi hollarda semantik jihatdan mos, lekin boshqa shaklda ifodalangan segmentlarni aniqlay olmaydi. Bu holat, ayniqsa, sinonimlar yoki grammatik variantlar mavjud bo‘lgan savollarda yaqqol namoyon bo‘ldi.

Semantik moslashtirish bosqichida transformer modellaridan foydalanish qidiruv natijalarining aniqligini sezilarli darajada oshirdi. BERT asosidagi model savol va huquqiy segmentlar o‘rtasidagi ma’no yaqinligini aniqlash orqali leksik jihatdan mos kelmagan, ammo semantik jihatdan to‘g‘ri javoblarni topishga yordam berdi. Natijada gibrid model faqat qidiruvga asoslangan tizimga nisbatan aniqroq va ishonchli natijalar ko‘rsatdi.

Qoida-asoslangan tekshiruv bosqichi tizim natijalarining ishonchliligini oshirishda muhim rol o‘ynadi. Ushbu bosqichda tanlangan segmentlarning huquqiy norma bilan mosligi tekshirildi va umumiy yoki noto‘g‘ri javoblar filtrlandi. Bu esa foydalanuvchiga taqdim etiladigan javoblarning huquqiy asoslanganligini ta’minlashga xizmat qildi.

Umuman olganda, olib borilgan tahlillar shuni ko‘rsatadiki, lingvistik va texnologik komponentlarni birlashtirgan gibrid model huquqiy savol-javob tizimlarida yuqori samaradorlikka ega. Ayniqsa, lingvistik tahlil va semantik moslashtirish bosqichlarining uyg‘un ishlashi tizim aniqligini oshirishda asosiy omil sifatida namoyon bo‘ldi. Shuningdek, model o‘zbek tilining morfologik xususiyatlarini hisobga olgan holda ishlab chiqilgani sababli milliy til uchun mos va amaliy jihatdan qo‘llash mumkin bo‘lgan yechim sifatida baholanishi mumkin.

Xulosa va takliflar. Mazkur tadqiqotda o‘zbek tilidagi huquqiy diskurs xususiyatlarini hisobga olgan holda savol-javob tizimlari uchun lingvistik va texnologik model ishlab chiqildi. Tahlillar shuni ko‘rsatdiki, huquqiy matnlarning normativligi, terminologik aniqligi va sintaktik murakkabligi ularni avtomatik qayta ishlash jarayonida maxsus yondashuvni talab qiladi. Ayniqsa, o‘zbek tilining agglutinativ tabiati savol va huquqiy matn o‘rtasidagi moslikni aniqlashni murakkablashtiradi va lingvistik tahlilning ahamiyatini oshiradi.

Tadqiqot natijalari asosida lingvistik tahlil, BM25 asosidagi qidiruv, transformer modellarga asoslangan semantik moslashtirish va qoida-asoslangan tekshiruvni birlashtiruvchi gibrid model samarali yondashuv sifatida asoslandi. Ushbu model savolni chuqur tahlil qilish, huquqiy korpusdan mos segmentlarni aniqlash, semantik yaqinlikni baholash va manbaga asoslangan javobni shakllantirish imkonini beradi. Modelning asosiy afzalligi shundaki, u faqat leksik moslikka emas, balki semantik va lingvistik moslikka ham tayanadi hamda huquqiy javoblarning ishonchliligini ta’minlaydi.

Olib borilgan tahlillar shuni ko‘rsatdiki, qidiruv va semantik yondashuvlarni alohida qo‘llash yetarli natija bermaydi, balki ularning integratsiyasi yuqori aniqlik va barqarorlikni ta’minlaydi. Shuningdek, qoida-asoslangan tekshiruv komponenti tizimda xatoliklarni kamaytirish va huquqiy javoblarning asoslanganligini oshirishda muhim rol o‘ynaydi.

Tadqiqot natijalari asosida quyidagi takliflar ilgari suriladi:

•o‘zbek tilida huquqiy QA tizimlari uchun maxsus annotatsiyalangan korpuslar yaratish zarur;

•morfologik tahlil va lemmatizatsiya modullarini yanada takomillashtirish lozim;
•huquqiy matnlarni semantik segmentlarga ajratish usullarini rivojlantirish maqsadga muvofiq;
•o‘zbek tiliga mos transformer modellarni kengaytirish va optimallashtirish zarur;
•real amaliyotda qo‘llash uchun huquqiy chatbot va maslahat tizimlarini ishlab chiqish tavsiya etiladi.

Xulosa qilib aytganda, taklif etilgan model o‘zbek tilida huquqiy savol-javob tizimlarini yaratish uchun nazariy va amaliy asos bo‘lib xizmat qiladi hamda kelgusida ushbu yo‘nalishda olib boriladigan tadqiqotlar uchun muhim ilmiy poydevor vazifasini bajaradi.

Foydalanilgan adabiyotlar ro‘yxati:

(1). Zhong H., Xiao C., Tu C., Zhang T., Liu Z., Sun M. How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 5218–5230.

(2). Martinez-Gil J. A survey on legal question-answering systems. Computer Science Review, 2023, vol. 48, Article 100552.

(3). Ariai F., Demartini G. et al. Natural Language Processing for the Legal Domain: A Survey of Tasks, Datasets, Models, and Challenges. ACM Computing Surveys, 2025.

(4). Rakhimova D., Turarbek A., Karyukin V., Sarsenbayeva A., Alieyev R. Legal AI in Low-Resource Languages: Building and Evaluating QA Systems for the Kazakh Legislation. Computers, 2025, vol. 14, № 9, Article 354.

(5). Mansurov B., Mansurov A. UzBERT: Pretraining a BERT Model for Uzbek. arXiv, 2021, arXiv:2108.09814.

(6). Kuriyozov E., Vilares D., Gómez-Rodríguez C. BERTbek: A Pretrained Language Model for Uzbek. Proceedings of SIGUL 2024 / LREC-COLING 2024, 2024.

(7). Robertson S., Zaragoza H. The Probabilistic Relevance Framework: BM25 and Beyond. Foundations and Trends in Information Retrieval, 2009, vol. 3, № 4, pp. 333–389.

(8). Devlin J., Chang M.W., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proceedings of NAACL-HLT, 2019, pp. 4171–4186.

Bakirova Gulasal Davlatovna (Denov Tadbirkorlik va pedagogika instituti tayanch doktorant (PhD), Surxandaryo, O‘zbekiston)

JEYMS JOYSNING “ULISS” ROMANIDA OKKAZIONALIZMLARNING MORFOLOGIK-STILISTIK VA KOGNITIV XUSUSIYATLARI

Annotatsiya. *Mazkur maqolada “Uliss” romanida morfologik stilistik vosita sifatida okkazionalizmlarning markaziy o‘rni tahlil qilinadi. Tadqiqot asarda uchraydigan muallifga xos, normadan chetga chiqqan til birliklarining strukturaviy va funksional xususiyatlarini o‘rganishga qaratilgan. Okkazionalizmlarning qo‘shma birliklar, semantik deformatsiya, morfologik kompressiya va takrorlash kabi turlari ilmiy asosda ko‘rib chiqiladi. Tahlil natijalari shuni ko‘rsatadiki, okkazionalizmlar nafaqat stilistik bezak, balki ma’no yaratish, emotsionallikni kuchaytirish va kognitiv jarayonlarni ifodalashning muhim vositasi hisoblanadi. Ular matnning obrazlilik darajasini oshiradi, tilning ekspressiv imkoniyatlarini kengaytiradi hamda inson tafakkurining assotsiativ va noan’anaviy xususiyatlarini aks ettiradi. Kognitiv-stilistik yondashuv nuqtai nazaridan, ushbu birliklar o‘quvchini faol interpretatsiyaga jalb etadi va matnni dinamik semantik tizimga aylantiradi.*

Kalit so‘zlar: *okkazionalizm, morfologik stilistika, lingvopoetika, Jeyms Joys, Uliss, modernizm, kognitiv stilistika, til eksperimenti, obrazlilik, semantik deformatsiya.*

Бакирова Гуласал Давлатовна (Докторант (PhD), Денауский институт предпринимательства и педагогики, Сурхандарья, Узбекистан)

МОРФОЛОГИЧЕСКО-СТИЛИСТИЧЕСКИЕ И ПОЗНАВАТЕЛЬНЫЕ ХАРАКТЕРИСТИКИ ОККАЗИОНАЛИЗМОВ В РОМАНЕ ДЖЕЙМСА ЖОЙСА «УЛИСС».

Аннотация. *Анализ романа «Улисс» показывает, что его лингвопоэтическая структура в корне сформирована взаимодействием модернистского и авангардного движений. Интегрируя разнообразные стилистические и когнитивные элементы, Джойс создает текст, бросающий вызов традиционным представлениям о языке, повествовании и значении. Использование в романе потока сознания, ассоциативного мышления и лингвистических экспериментов отражает сложность человеческого познания и динамичный характер восприятия. В то же время, его взаимодействие с различными художественными движениями обогащает его эстетические и интерпрета-*