

ELEKTRON ENSIKLOPEDIK LUG‘ATNING ONTOLOGIK MAQOMI VA AXBOROT-QIDIRUV TIZIMLARIDAGI O‘RNI

<https://doi.org/10.70728/phil.v05.i06.004>

Abduqaxxorov Sardorbek Akramjon o‘g‘li
sardorbek.akramovich@gmail.com
Andijon davlat chet tillari instituti

Annotatsiya: Maqolada elektron ensiklopedik lug‘at (EEL) lug‘at va ensiklopediya oralig‘ida joylashgan ikki qavatli ontologik maqomga ega axborot manbasi sifatida tahlil qilingan. EELda bilimning leksik, semantik, konseptual, pragmatik va bog‘lovchi sathlarda tashkil etilishi hamda uning axborot-qidiruv tizimlari komponentlaridagi (indekslash, so‘rovni kengaytirish, semantik qidiruv, natijalarni tartiblash) vazifalari ochib berilgan. EELning bilim grafi, RAG va multimodal qidiruv tizimlari bilan integratsiyasi yoritilgan.

Kalit so‘zlar: elektron ensiklopedik lug‘at, ontologik maqom, axborot-qidiruv tizimi, bilim grafi, semantik qidiruv, RDF/OWL, RAG

ONTOLOGICAL STATUS OF ELECTRONIC ENCYCLOPEDIA DICTIONARY AND ITS PLACE IN INFORMATION SEARCH SYSTEMS

Abduqaxxorov Sardorbek Akramjon oğlu
sardorbek.akramovich@gmail.com
Andijan State Institute of Foreign Languages

Abstract: The article analyses the electronic encyclopedic dictionary (EED) as an information source with a two-layered ontological status situated between a dictionary and an encyclopedia. The organisation of knowledge across lexical, semantic, conceptual, pragmatic and connective levels, and the functions of the EED within information retrieval system components (indexing, query expansion, semantic search, ranking) are revealed. Integration with knowledge graphs, RAG and multimodal search systems is discussed.

Keywords: electronic encyclopedic dictionary, ontological status, information retrieval system, knowledge graph, semantic search, RDF/OWL, RAG

Kirish

Elektron ensiklopedik lug‘at (keyingi o‘rinlarda EEL) zamonaviy raqamli bilim infratuzilmasida lug‘at va ensiklopediyaning tutashgan nuqtasida shakllanadigan

murakkab axborot manbasidir. Lugʻat odatda til birligi, uning maʼnosi, grammatik va uslubiy xususiyatlarini tavsiflaydi; ensiklopediya esa nomlangan predmet, hodisa, shaxs, joy yoki nazariya haqidagi bilimni tizimli bayon qiladi. EEL bu ikki yoʻnalishni raqamli muhitda birlashtiradi: u bir tomondan leksik kirish nuqtasi orqali ishlaydi, ikkinchi tomondan obyektning bilim modelini saqlaydi.

EELni faqat “qogʻoz lugʻatning elektron nusxasi” yoki “maqolalar toʻplami” deb talqin qilish yetarli emas. Bunday talqin raqamli resursning eng muhim imkoniyatlarini - mashinada qayta ishlanish, semantik indekslash, koʻp mezonli qidiruv, konseptlararo bogʻlanish, avtomatik tavsiya, soʻrovni kengaytirish va bilim grafiga ulanib ishlashni cheklab qoʻyadi.

Zamonaviy tilshunoslikda lugʻatning maqomi haqidagi munozaralar metaleksikografiyaning markaziy muammolaridan biriga aylangan. V.G.Gak taʼkidlaganidek, lugʻat shunchaki soʻzlar toʻplami emas, balki tilning leksik tizimini maʼlum nazariy konsepsiya asosida tartibga soluvchi modeldir. Yu.Karaulov lugʻatning gnoseologik vazifasini, N.Mahmudov esa uning leksikografik amaliyot bilan tilshunoslik nazariyasi orasidagi vositachi rolini taʼkidlaydi.

Maqolaning maqsadi - EELning ontologik maqomini aniqlash va uning axborot-qidiruv tizimlaridagi oʻrnini asoslashdan iborat. Bu masala lugʻat maqolasi qanday birliklardan tashkil topishi, bu birliklar qaysi turdagi mavjudliklarni ifodalashi va axborot-qidiruv tizimi ular bilan qanday operatsiyalar bajarishi mumkinligini aniqlashdan boshlanadi.

Metodlar

Tadqiqotda ontologik modellashtirish, qiyosiy va struktur-funksional tahlil metodlari qoʻllanildi. Nazariy bazani T.Gruber, R.Studer, N.Guarino ontologiya taʼriflari, J.Sowaning leksikon-ontologiya munosabati, V.G.Gak va Yu.Karaulovning leksikografik maqom haqidagi qarashlari, K.Manning va boshqalarning axborot qidiruvi modeli, R.Baeza-Yates va B.Ribeiro-Netoning resurs tasnifi hamda S.Nielsen va R.Lewning leksikografik axborot xarajatlari nazariyasi tashkil etdi.

Anʼanaviy va elektron ensiklopedik lugʻat sakkiz mezon boʻyicha qiyoslandi; bilimning ifodalanish sathlari, zamonaviy elektron resurslarning ontologik tipologiyasi va EELning AQT komponentlaridagi vazifalari jadval shaklida modellashtirildi. N.Guarinoning “mavjudlik darajalari” konsepsiyasi (konseptual, lingvistik, texnik sath) tahlil asosi sifatida olindi.

Natijalar

Tahlil EELning ikki qavatli ontologik maqomga ega ekanligini koʻrsatdi: u bir tomondan leksik kirish nuqtasi (lema, termin, nom, qisqartma, variant yozuv, transliteratsiya) orqali ishlaydi, ikkinchi tomondan shu birlik ifodalagan obyektning bilim modelini, oʻzaro aloqalarini va qidiruvga yaroqli semantik belgilarini saqlaydi.

J.Sowa ta'kidlaganidek, har qanday lug'at yashirin ontologiyaga ega bo'lib, leksikografiyaning ilmiy bosqichi uni eksplisit shaklga keltirishni talab qiladi.

V.P.Berkov lingvistik va ensiklopedik lug'atlar farqini aniq belgilaydi: birinchisi til birliklarining lisoniy ma'nolarini, ikkinchisi shu birliklar bilan ifodalanadigan tushunchalar va voqeliklar haqidagi bilimni taqdim etadi. N.Guarino tasnifiga ko'ra EELni domen ontologiyalari qatoriga qo'shish mumkin, chunki u muayyan bilim sohasi yoki millat madaniyatining tushunchaviy strukturasini formal ifodalaydi.

EELda bilim besh sathda tashkil etilishi aniqlandi: leksik sath (bosh so'z, grammatik forma - XML/JSON tegi), semantik sath (ta'rif, ma'no izohi - RDF triple, SKOS konseptlari), konseptual sath (tushuncha ierarxiyasi - OWL klasslari), pragmatik sath (misol, frazeologizm, sitata - korpus tizimlari) va bog'lovchi sath (ichki/tashqi havola - Linked Data, URI). Har bir sath o'ziga xos qidiruv strategiyasi va indekslash texnologiyasini talab qiladi.

An'anaviy va elektron lug'atning sakkiz mezon bo'yicha qiyosi EEL bosma nashrdan nafaqat texnologik, balki ontologik jihatdan ham tubdan farq qilishini ko'rsatdi: dinamik kengayuvchi hajm, ko'p o'lchamli semantik tartiblash, real vaqtli yangilanish, multimodal ma'lumot turlari, eksplisit ontologiya va Linked Data orqali global integratsiya uning differensial belgilaridir.

Zamonaviy elektron resurslarning ontologik tipologiyasi olti turni qamrab oladi: onlayn lug'at, elektron ensiklopediya, wiki-platforma, leksik tarmoq (WordNet), bilim grafi (DBpedia, Wikidata) va domen ontologiyasi (SNOMED CT). EEL ikki qutb - ma'lumot taqdimi va strukturalashgan bilim ifodasi - o'rtasidagi oraliq mavqeni egallab, leksikografiya va bilim muhandisligi sohalari o'rtasidagi bog'lovchi vosita vazifasini bajaradi.

EELning AQT komponentlaridagi vazifalari aniqlandi: indekslashda lemmatizatsiya va polisemiyaning hal etish, so'rov tahlilida semantik tushunish, so'rovni kengaytirishda sinonim/gipernim/derivat qo'shish, semantik qidiruvda ontologiyaga asoslangan tushunchaviy qidiruv, natijalarni tartiblashda avtoritetli manba sifatida ulush berish hamda bilim grafi orqali javobni boyitish (knowledge panel). S.Nielsen nazariyasiga ko'ra lug'at samaradorligi foydalanuvchidan talab qilinadigan kognitiv va vaqt xarajatlari bilan teskari proporsional aloqada bo'ladi.

Muhokama

EELning zamonaviy axborot ekotizimidagi o'rni multifunksional: u endi mustaqil mahsulot emas, balki ko'p sathli tizimning tarkibiy qismidir. 2012-yilda Google joriy etgan bilim grafi qidiruv natijalarini boyituvchi panellar uchun Wikipedia, Wikidata va WordNet manbalaridan foydalanadi; shu tarzda EEL javob shakllantiruvchi vositaga aylangan va leksikografiyaning rolini qayta belgilashni taqozo etadi.

EELning AQT bilan integratsiyasi foydalanuvchi tajribasini sezilarli o'zgartiradi: semantik qidiruv tufayli javob bevosita natijalarning yuqori qismida struktura

ko'rinishida taqdim etiladi, bu esa axborot olish vaqtini o'rtacha 60-70 foizga qisqartiradi. O'zbek tilidagi EELning xalqaro qidiruv tizimlari (Google Knowledge Graph, Bing Entities) bilan integratsiyasi millat madaniyati va tarixi haqidagi ma'lumotlarning xalqaro miqyosda tarqalishiga zamin yaratadi.

Maxsus e'tiborga molik jihat - EELning katta til modellari va generativ qidiruv (RAG - Retrieval-Augmented Generation) tizimlari bilan o'zaro ta'siri. Kichik tillar, jumladan o'zbek tili bo'yicha sifatli, struktura asosli ma'lumotlar yetishmasligi modellar javobining aniqligini pasaytiradi. Sifatli EEL mavjudligi RAG-tizimlari uchun ishonchli, normativ manba bo'lib, tizim javoblarining haqqoniyligini ta'minlaydi.

EELning multimodal qidiruv tizimlari bilan integratsiyasi ham istiqbolli yo'nalishdir: zamonaviy lug'atlar maqolalarni tasvir, audio talaffuz va video illyustratsiyalar bilan boyitib taqdim etadi, bu kontent qidiruv tizimi tomonidan parallel indekslanishi mumkin.

EELning AQTdagi vazifalarini bajarishi uchun ma'lumotlarning standartlashganligi (TEI, LMF, OntoLex-Lemon), API orqali ochiqqligi, barqaror identifikatorlar tizimi (har bir maqolaga maxsus URI), versiya boshqaruvi va Linked Data tamoyiliga muvofiqligi talab etiladi. O'zbek tili uchun bunday EEL yaratish davlat axborot infrastrukturasi strategik vazifasi hisoblanadi va O'zbekiston Respublikasi Prezidentining til siyosatiga oid farmonlari bilan bevosita bog'liqdir.

Xulosa (Conclusion)

Elektron ensiklopedik lug'atning ontologik maqomi uning lug'at va ensiklopediya oralig'idagi ikki qavatli tabiatida namoyon bo'ladi: u leksik birlikning shakli, ma'nosi, variantlari va qo'llanishini tavsiflashi bilan birga, shu birlik orqali ifodalanuvchi obyektning bilim modelini semantik aloqalar tizimida ochib beradi.

EELning samarali ishlashi leksik, semantik, konseptual, pragmatik va bog'lovchi sathlarning izchil tashkil etilishiga bog'liqligi isbotlandi. Bunday ko'p sathli tuzilma lug'atning ilmiy aniqligi va qidiruv tizimlaridagi funksional qiymatini kuchaytiradi hamda uni axborot ekotizimi va sun'iy intellekt rivojlanishida strategik resurs darajasiga olib chiqadi.

Foydalanilgan adabiyotlar

1. Gruber T. Toward Principles for the Design of Ontologies // Int. J. Human-Computer Studies. - 1995.
2. Studer R., Benjamins V.R., Fensel D. Knowledge Engineering: Principles and Methods // Data & Knowledge Engineering. - 1998.
3. Guarino N., Oberle D., Staab S. What Is an Ontology? // Handbook on Ontologies. - Springer, 2009.
4. Sowa J. Knowledge Representation. - Brooks/Cole, 2000.

5. Berners-Lee T., Hendler J., Lassila O. The Semantic Web // Scientific American. - 2001.
6. Manning C.D., Raghavan P., Schütze H. Introduction to Information Retrieval. - Cambridge: CUP, 2008.
7. Baeza-Yates R., Ribeiro-Neto B. Modern Information Retrieval. - Addison-Wesley, 2011.
8. Nielsen S. The Effect of Lexicographical Information Costs // Lexikos. - 2008.
9. Lew R. Online Dictionaries of English // Electronic Lexicography. - Oxford: OUP, 2012.
10. Vrandečić D., Krötzsch M. Wikidata // Communications of the ACM. - 2014.