



## IMOM G‘AZZOLIY ASARLARINING SEMANTIK TARMOG‘I: NLP YONDASHUVI

**Barnoyeva Zarina Sayfiddinovna,**  
tayanch doktoranti  
[barnoyevazarina072@gmail.com](mailto:barnoyevazarina072@gmail.com)  
BuxDU

**Annotatsiya.** Mazkur maqolada Imom G‘azzoliy asarlari asosida yaratilgan korpusning semantik tuzilmasi zamonaviy kompyuter lingvistikasi yondashuvlari asosida tahlil qilinadi. Tadqiqot jarayonida semantik teglash, ontologik modellashtirish hamda word2vec algoritmi orqali so‘zlarning yashirin semantik bog‘liqliklari aniqlangan. Natijada diniy, axloqiy va ijtimoiy semantik maydonlardan iborat kompleks semantik model ishlab chiqildi. Ushbu yondashuv diniy-falsafiy matnlarni avtomatlashtirilgan tahlil qilish va konseptual qidiruv tizimlarini yaratishda muhim ahamiyat kasb etadi.

**Kalit so‘zlar:** *semantik tarmoq, NLP, korpus lingvistikasi, semantik teglash, diniy matnlar, word2vec.*

**Annotation.** In this article, the semantic structure of the corpus created based on the works of Imam Ghazzali is analyzed based on modern computational linguistics approaches. During the research, the hidden semantic relationships of words were identified using semantic tagging, ontological modeling, and the word2vec algorithm. As a result, a complex semantic model consisting of religious, moral, and social semantic fields was developed. This approach is of great importance in the automated analysis of religious and philosophical texts and the creation of conceptual search systems.

**Keywords:** *semantic network, NLP, corpus linguistics, semantic tagging, religious texts, word2vec.*

**Kirish.** So‘nggi yillarda kompyuter lingvistikasi sohasida yuz berayotgan rivojlanishlar tilni faqat grammatik jihatdan emas, balki semantik darajada ham



o'rganishni taqozo etmoqda. Ayniqsa, diniy-falsafiy matnlar murakkab ma'no tizimiga ega bo'lib, ularni an'anaviy lingvistik usullar yordamida to'liq ochib berish har doim ham mumkin emas. Imom G'azzoliy asarlari mazmunan chuqur, konseptual jihatdan murakkab va o'zaro bog'langan tushunchalarga boy bo'lib, unda diniy, axloqiy hamda ijtimoiy kategoriyalar bir butun tizimni tashkil etadi. Shu bois mazkur asarlarni zamonaviy NLP metodlari asosida o'rganish dolzarb ilmiy muammolardan biri hisoblanadi.

Mazkur tadqiqotning asosiy maqsadi – G'azzoliy asarlarida mavjud semantik tizimni aniqlash, uni ontologik model asosida tasniflash va kompyuter lingvistikasi vositalari yordamida semantik tarmoq ko'rinishida modellashtirishdan iborat.

Mualliflik korpuslari tilshunoslik va NLP (Natural Language Processing)da muhim ahamiyatga ega bo'lib, ular til birliklarini keng ko'lamda avtomatik tahlil qilish va turli lingvistik qonuniyatlarni aniqlash imkonini beradi. Xususan, muayyan muallif ijodiga tegishli mualliflik korpusi til uslubi, lug'at boyligi va g'oyaviy xususiyatlarni chuqurroq o'rganish uchun qo'l keladi. Ushbu tadqiqotda buyuk alloma Imom Abu Homid al-G'azzoliy (1058–1111) asarlarining o'zbek tilidagi tarjimalari asosida yaratilgan mualliflik korpusi taqdim etildi. Imom G'azzoliy o'zining “*Ihyo ul-ulum ad-din*” kabi nodir asarlari orqali islom falsafasi va axloqiy ta'limotlarga ulkan hissa qo'shgan; uning asarlarining o'zbek tiliga tarjimalari mazmun va til xususiyatlarini kompyuter lingvistikasi nuqtai nazaridan tadqiq qilish dolzarbdir.

Mazkur korpusni yaratishdan maqsad – Imom G'azzoliy asarlari matnlarini raqamli ko'rinishda to'plash va ularni morfologik, sintaktik hamda semantik jihatdan teglash orqali izlanishlar uchun qulay shaklga keltirishdir. Korpus bir necha qatlamli annotatsiyaga ega bo'lib, har bir so'zning lug'aviy shakli (lemma), so'z turkumi (morfologik tegi), grammatik xususiyatlari (masalan, kelishik, zamon, nisbat) va gapdagi sintaktik vazifasi (bog'lanish turi) belgilandi. Bunday lingvistik



teglash matnlarni chuqur lingvistik tahlil qilish, kerakli birliklarni tezkor topish, turli asarlarni qiyosiy o'rganish imkoniyatini yaratadi. Shuningdek, korpusga kiritilgan semantik teqlash yordamida muhim tushuncha va nomlar (masalan, Alloh, Payg'ambar, jannat, dunyo kabi tushunchalar) alohida belgilab chiqildi, bu esa mazkur diniy-falsafiy matnlarda asosiy tushunchalarning qo'llanishini tahlil etish imkonini oshiradi.

O'zbek tilida hozirgi kungacha kompyuter lingvistikasi uchun resurslar cheklangan bo'lib, yaqindagina ilk Universal Dependencies (UD) formatidagi to'liq teglangan treebank taqdim etildi [10:134]. UD standarti turli tillar uchun bir xil morfo-sintaktik teqlash qoidalarini joriy etuvchi xalqaro tashabbus bo'lib, bugungi kunda 168 ta tilda jami 296 ta treebankni o'z ichiga oladi. O'zbek tilida birinchi UD treebanki (Uzbek-UT) 2025-yilda 500 gapdan iborat kichik hajmli namunaviy to'plam sifatida yaratilib, unda so'zning asosi, UPOS (umumiy so'z turkumi), morfologik xususiyatlari va sintaktik bog'lanishlari belgilangan [4:10-17]. O'tgan yillarda o'zbek tilini kompyuterda qayta ishlash borasida ayrim ishlar bajarilgan: masalan, Prolog dasturlash tilida birlamchi morfologik tahlil dasturi yaratilgan (Matlatipov va Vetulani, 2009)[8:83-110], biroq u murakkab so'z shakllarini to'liq qamrab olmagan. Keyinchalik o'zbek tilida kengaytirilgan teqlash to'plami va qoidaga asoslangan UzbPosTegger morfologik teglagichi ishlab chiqilib, 12 ta so'z turkumi bo'yicha teglangan namunaviy matnlarda sinovdan o'tkazildi[5:63-68]. Shuningdek, o'zbek tilida so'zlarni lemmalash (asosiy shaklga keltirish) va stemming qilish bo'yicha vositalar taklif etilib, maxsus algoritmlar yaratildi. 2023-yilda esa o'zbek tilida stemmer, lemmalovchi va POS-teglagichni birlashtirgan UzMorphAnalyzer dasturi taqdim etildi. Apertium platformasidagi o'zbek tili uchun mo'ljallangan lug'at va finite-state avtomatga asoslangan morfologik tahlilchi ham matnlarni avtomatik teqlashda qo'l keladi[6:238-243]. Matnni semantik razmetkalash yoki semantik teqlash deganda matndagi har bir muhim til birligini



muayyan semantik kategoriya, guruh yoki maydonga mansubligini ko'rsatuvchi belgilar bilan ajratish jarayoni tushuniladi. Bu jarayon natijasida matnning ichki ma'no tuzilmasi aniqroq ko'rinish oladi.

Semantik teglashning asosiy afzalligi shundaki, u oddiy leksik qidiruvdan farqli ravishda matnni konseptual darajada tahlil qilish imkonini beradi. Masalan, oddiy qidiruv tizimida foydalanuvchi “sabr” so'zini kiritrsa, faqat shu so'z mavjud bo'lgan joylar aniqlanadi. Semantik teglangan korpusda esa ushbu tushunchaga yaqin bo'lgan “chidam”, “bardosh”, “matonat” kabi birliklar ham natijaga qo'shiladi. Bu esa qidiruv samaradorligini sezilarli darajada oshiradi.

Xalqaro tajribada semantik teglash masalasi keng o'rganilgan. Xususan, WordNet tizimi so'zlarni sinonimik to'plamlar asosida bog'lab, ularning iyerarxik semantik tarmog'ini yaratishga xizmat qiladi. Shuningdek, UCREL USAS tizimi turli tillar uchun tayyor semantik kategoriyalar tizimini taklif etadi.

Biroq o'zbek tili, ayniqsa diniy-falsafiy matnlar uchun maxsus ishlab chiqilgan semantik teglash tizimi hali mavjud emas. Shu jihatdan Imom G'azzoliy asarlari asosida bunday model yaratish ilmiy jihatdan muhim hisoblanadi. Tadqiqot uchun G'azzoliy asarlarining o'zbek tilidagi tarjimalari asosida maxsus korpus shakllantirildi. Korpus tarkibiga “Kimyoyi saodat”, “Ey farzand”, “Riyozatun nafs”, “Qalb gavhari”, “Ilm kitobi” kabi asarlar kiritildi. Yaratilgan korpusning umumiy hajmi taxminan 200 ming tokenni tashkil etadi. Matnlar dastlab texnik jihatdan tozalandi, ortiqcha belgilar olib tashlandi va yagona kodlash tizimiga keltirildi. Shundan so'ng ular maxsus dasturiy muhitda tokenlashtirildi. Korpus asosida semantik tahlil o'tkazish uchun maxsus ontologik model ishlab chiqildi. Ushbu modelda uchta asosiy semantik maydon ajratildi: Birinchisi – diniy maydon bo'lib, u Alloh, Qur'on, ibodat va oxirat kabi tushunchalarni o'z ichiga oladi. Bu maydon G'azzoliy asarlarining asosiy mazmuniy yadrosini tashkil etadi. Ikkinchisi – axloqiy maydon bo'lib, unda insonning ichki dunyosi, xulq-atvori va ma'naviy holatlari



bilan bog‘liq tushunchalar jamlangan. Masalan, sabr, taqvo, ixlos, riyo kabi birliklar shu guruhga kiradi. Uchinchi – ijtimoiy maydon bo‘lib, unda inson va jamiyat o‘rtasidagi munosabatlar aks etadi. Bu yerda inson, jamiyat, oila kabi tushunchalar muhim o‘rin tutadi. Umumiy hisobda 13 ta semantik guruh va 80 ga yaqin kichik semantik to‘dalar aniqlanib, ular iyerarxik tizimga joylashtirildi.

Korpusni semantik teglash qo‘lda teglangan qoidalar hamda avtomatik klasterlash algoritmlari uyg‘unligi asosida amalga oshirildi[7:12-23]. Dastlab, yuqorida tuzilgan ontologiya asosida har bir semantik guruh va to‘da uchun kalit so‘zlar ro‘yxati tayyorlandi. Bunda guruh mazmunini ifodalovchi 5–10 ta tipik so‘z tanlab olindi va Sketch Engine vositasida korpusdan qidirildi. Qidiruv natijalarida topilgan ko‘plab kontekstlar ichidan haqiqatan shu guruhga tegishli yangi so‘zlar (sinonimlar, so‘z birikmalari) aniqlanib, ro‘yxatga qo‘shib borildi. Masalan, dastlab “sabr” so‘zi asosida axloqiy fazilatlar guruhini belgilash boshlangan bo‘lsa, matndan unga yaqin ishlatilgan “chidam”, “bardosh”, “matonat” kabi sinonimlar hamda “qanoat”, “shukr” kabi semantik yaqin tushunchalar topildi va *Axloqiy fazilatlar* guruhiga birlashtirildi. Shu tariqa har bir guruh bo‘yicha kengaytirilgan lug‘aviy bazalar shakllantirildi. Keyingi bosqichda avtomatik semantik model tuzishga o‘tdik. Xususan, word2vec algoritmi yordamida so‘zlarning vektorli tasviri yaratildi. Ushbu model orqali so‘zlarning kontekstual yaqinligi aniqlanib, ularning semantik o‘xshashligi matematik jihatdan ifodalandi. Tadqiqot natijalari G‘azzoliy asarlarining murakkab semantik tizimga ega ekanligini ko‘rsatdi. Eng avvalo, diniy tushunchalar korpusda eng yuqori chastotaga ega ekanligi aniqlandi. Bu esa asarlarning asosiy yo‘nalishi diniy-falsafiy ekanligini tasdiqlaydi.

Axloqiy tushunchalar ham muhim o‘rin tutib, ular asosan nasihat va o‘git shaklida ifodalangan. Bu holat muallif uslubining didaktik xarakterga ega ekanligini ko‘rsatadi.

Ijtimoiy tushunchalar esa nisbatan kamroq uchrasa-da, ular jamiyat va inson o'rtasidagi munosabatlarni tushuntirishda muhim rol o'ynaydi.

Word2vec modeli natijalari shuni ko'rsatdiki, semantik jihatdan yaqin so'zlar bir-biriga yaqin joylashadi. Masalan, “Alloh”, “tavba”, “qalb” kabi so'zlar bir klasterda, “sabr”, “taqvo” kabi birliklar boshqa klasterda jamlangan. Olingan natijalar G'azzoliy asarlarining semantik tuzilmasi tizimli va iyerarxik xarakterga ega ekanligini ko'rsatadi. NLP metodlari yordamida bu tizimni modellashtirish diniy matnlarni chuqurroq anglash imkonini beradi. Bundan tashqari, semantik teglash konseptual qidiruv tizimlarini yaratishda muhim ahamiyatga ega. Bu esa kelajakda sun'iy intellekt asosida diniy bilim bazalarini yaratish imkonini beradi. Mazkur tadqiqotda Imom G'azzoliy asarlari asosida semantik tarmoq modeli ishlab chiqildi. Tadqiqot natijasida:

- semantik ontologiya yaratildi
- semantik teglash amalga oshirildi
- word2vec asosida modellashtirish bajarildi

Natijalar shuni ko'rsatdiki, diniy-falsafiy matnlar murakkab semantik tizimga ega bo'lib, ularni NLP yordamida samarali tahlil qilish mumkin.

**Xulosa.** Semantik teglash hozircha juda cheklangan toifalarni qamrab oladi. Faqatgina asosiy nomlar va diniy atamalar belgilandi xolos. Kelgusida bu qatlamni kengaytirish mumkin: masalan, mavzuga oid teglash (topic tagging) joriy etish yoki har bir gap uchun uning emosional-ohangini (masalan, nasihat, ogohlantirish, dalillash kabi) belgilash kabi yo'nalishlar qiziqarli bo'lardi. Shu bilan birga, ushbu korpusni boshqa mualliflar – masalan, *Yassaviy*, *Bahouddin Naqshband*, *Alisher Navoiy* asarlari korpuslari bilan boyitish va taqqoslash ham mumkin bo'ladi. Bunday kengaytma ko'p muallifli diniy adabiyot korpusini shakllantirib, o'zbek tilida turli davr va uslubdagi matnlarni solishtirishga imkon yaratadi.



Imom G‘azzoliy korpusi semantik teglash orqali diniy, axloqiy va ijtimoiy so‘zlar bo‘yicha tahlil qilindi, bunda Word2vec modeli yordamida avtomatik klasterlash va qo‘lda ekspert nazorati kombinatsiyasi qo‘llanildi. Natijalar shuni ko‘rsatdiki, diniy atamalar eng ko‘p (~66%), axloqiy tushunchalar muhim (~28%), ijtimoiy so‘zlar esa nisbatan kamroq (~5%) uchraydi. Teglash jarayoni so‘zlar orasidagi konseptual bog‘lanishlarni aniqlashga yordam berdi, masalan, “ilm-ma’rifat” yoki “dunyo-oxirat” kabi oraliq klasterlar shakllandi. Ushbu metodologiya nafaqat matnni chuqur tahlil qilish, balki kelajakda avtomatik semantik teglash, intellektual qidiruv va NLP loyihalarida qo‘llanish imkonini beradi.

### Foydalanilgan adabiyotlar ro‘yxati

1. Al-G‘azzoliy A.H. Ihya ulum ad-din (Dinning tirilishi). – Toshkent: O‘zbekiston, 2019. – 512 b.
2. Al-G‘azzoliy A.H. Mukoshafat ul-qulub (Qalbni yoritish). – Toshkent: Adolat, 2002. – 288 b.
3. Abdullayev I. Tasavvuf va ma’naviy poklanish. – Samarqand: Sharq, 2018.
4. Akhundjanova A. Harmonizing annotation of Turkic postverbal constructions: A comparative study of UD treebanks // Proceedings of the 18th Workshop on Building and Using Comparable Corpora (BUCC). – 2025. – P. 10–17.
5. Boltayevich E.B., Samariddinovich S.S., Mirdjonovna K.S., Adali E., Yuldashevna X.Z. POS tagging of Uzbek text using Hidden Markov Model // 2023 8th International Conference on Computer Science and Engineering (UBMK). – 2023. – P. 63–68.
6. Benko V., Garabik R., Mirdjonovna K.S. Araneum Uzbekicum: A gigaword web-crawled Uzbek corpus // 2024 9th International Conference on Computer Science and Engineering (UBMK). – 2024. – P. 238–243.



7. Elov B., Hamroyeva S., Hamroqulova M. NLPda semantik teglash usullari // Uzbekistan: Language and Culture. – 2024. – T. 1, №1. – B. 12–23.
8. Matlatipov G., Vetulani Z. Representation of Uzbek morphology in Prolog // Aspects of Natural Language Processing: Essays Dedicated to Leonard Bolc on the Occasion of His 75th Birthday. – 2009. – P. 83–110.
9. Normatov Sh. Tasavvuf va uning falsafiy jihatlari. – Toshkent: O‘zbekiston Milliy ensiklopediyasi, 2012.
10. Nivre J., De Marneffe M.C., Ginter F., Hajič J., Manning C.D., Pyysalo S., Zeman D. Universal Dependencies v2: An ever-growing multilingual treebank collection // arXiv preprint. – 2020. – arXiv:2004.134.