

JAHON ADABIYOTINI TADQIQ ETISHDA SUN'IY INTELEKT YORDAMIDA BADIY MATNLARNI TAHLIL VA TALQIN QILISH MASALALARI

Sharipova Feruza Akramovna

Qarshi davlat universiteti, Axborot tizimlari va texnologiyalari ta'lim
yo'nalishi, II-bosqich talabasi

Annotatsiya. Ushbu maqolada jahon adabiyotini tadqiq etishda sun'iy intellekt texnologiyalarining badiiy matnlarni tahlil va talqin qilishdagi o'rni va imkoniyatlari ko'rib chiqiladi. Zamonaviy tabiiy tilni qayta ishlash (NLP) usullari, katta til modellari (LLM) va mashinali o'rganish algoritmlarining adabiyotshunoslik sohasidagi tatbiqiga alohida e'tibor qaratiladi. Maqolada sun'iy intellekt yordamida matn tahlilining metodologik asoslari, uning an'anaviy adabiyotshunoslik usullaridan farqlari va ushbu ikki yondashuvning integratsiyasi masalalari muhokama etiladi. Tadqiqot doirasida Sharq va G'arb adabiyoti namunalariga qo'llanilgan sun'iy intellekt tahlili natijalari taqqoslanadi. Maqola adabiyotshunos, tilshunoslar va sun'iy intellekt sohasidagi mutaxassislar uchun mo'ljallangan.

Kalit so'zlar: sun'iy intellekt, badiiy matn tahlili, tabiiy tilni qayta ishlash, katta til modellari, adabiyotshunoslik, matn talqini, mashinali o'rganish, jahon adabiyoti, raqamli gumanistika, qiyosiy adabiyot tadqiqi.

Аннотация. В данной статье рассматривается роль и возможности технологий искусственного интеллекта в анализе и интерпретации художественных текстов в контексте исследования мировой литературы. Особое внимание уделяется применению методов обработки естественного языка (NLP), больших языковых моделей (LLM) и алгоритмов машинного обучения в области литературоведения. В статье обсуждаются методологические основы анализа текста с помощью искусственного интеллекта, его отличия от традиционных методов литературного анализа, а также вопросы интеграции этих двух подходов. В рамках исследования сравниваются результаты ИИ-анализа, применённого к образцам восточной и западной литературы. Статья предназначена для литературоведов, лингвистов и специалистов в области искусственного интеллекта.

Ключевые слова: искусственный интеллект, анализ художественного текста, обработка естественного языка, большие языковые модели, литературоведение, интерпретация текста, машинное обучение, мировая литература, цифровые гуманитарные науки, сравнительное литературоведение.

Abstract. This article examines the role and capabilities of artificial intelligence technologies in the analysis and interpretation of literary texts in the context of world literature research. Special attention is given to the application of natural language processing (NLP) methods, large language models (LLM), and machine learning algorithms in the field of literary studies. The article discusses the methodological foundations of AI-driven text analysis, its differences from traditional literary criticism methods, and the integration of these two approaches. The research compares the results of AI analysis applied to samples of Eastern and Western literature. The article is intended for literary scholars, linguists, and artificial intelligence specialists.

Keywords: artificial intelligence, literary text analysis, natural language processing, large language models, literary studies, text interpretation, machine learning, world literature, digital humanities, comparative literature research.

Kirish. XXI asrda raqamli texnologiyalar rivojlanishi bilan birga sun'iy intellekt (SI) insoniyatning barcha faoliyat sohalariga, jumladan gumanitar fanlarga ham kirib keldi. Adabiyotshunoslik — an'anaviy ravishda inson aqli va his-tuyg'ulariga asoslangan soha — bugun sun'iy intellekt texnologiyalari bilan yangi munosabatlar tizimini shakllantirmoqda. Badiiy matnlarni kompyuter yordamida tahlil qilish g'oyasi yangi emas: 1960–70-yillardagi «stilometrik» tadqiqotlardan tortib bugungi katta til modellarigacha bo'lgan yo'l adabiyotshunoslikdagi metodologik inqilob sifatida baholanmoqda.

Sun'iy intellekt texnologiyalari matn tahlilida bir nechta muhim imkoniyatlarni taqdim etadi: katta hajmdagi matnlarni tezkor qayta ishlash, yashirin uslubiy xususiyatlarni aniqlash, turli davr va madaniyatlar adabiyotini qiyosiy o'rganish. Biroq bu imkoniyatlar bilan birga bir qator muammolar ham mavjud — matn ta'sirchanligi, badiiy qadriyatlar va muallif niyatini to'g'ri talqin qilish masalalari sun'iy intellekt uchun hali to'liq hal etilmagan muammolar bo'lib qolmoqda.

Ushbu maqola sun'iy intellektning jahon adabiyotini tadqiq etishdagi joriy holati, usullari va istiqbollarini o'rganishga qaratilgan.

Sun'iy intellekt va matn tahlilining nazariy asoslari.

Tabiiy Tilni Qayta Ishlash (NLP) Sohasining Rivojlanishi. Tabiiy tilni qayta ishlash (Natural Language Processing — NLP) — kompyuter fani va tilshunoslikning kesishmasida joylashgan soha bo'lib, kompyuterlarning inson tili bilan ishlash qobiliyatini o'rganadi. NLP ning adabiy matn tahlilida qo'llanilishi uch asosiy tarixiy bosqichdan iborat:

1-bosqich (1950–1980-lar): Qoidalarga asoslangan tizimlar. Grammatik tahlil va lug‘at ma‘lumotnomalari asosida ishlagan dastlabki tizimlar. Bu davrda Noam Chomskining transformatsion grammatikasi katta ta‘sir ko‘rsatdi.

2-bosqich (1980–2010-lar): Statistik usullar. Katta ma‘lumotlar to‘plamiga asoslangan statistik modellar vujudga keldi. N-gram modellari, Markov zanjirlari, qaror daraxti algoritmlari keng qo‘llanildi.

3-bosqich (2010-yillardan hozirga): Neyron tarmoqlar va transformerlar. BERT, GPT-4, T5, Claude kabi modellar tabiiy tilni qayta ishlash sohasida haqiqiy inqilob yasadi. Bu modellar milliardlab parametr asosida ishlaydi.

1-jadval. NLP Texnologiyalarining Adabiy Tahlilda Qo‘llanilish Sohalari

Texnologiya	Qo‘llanilish sohasi	Misol
Sentiment tahlil	Qahramonlar his-tuyg‘ulari tahlili	«Urush va Tinchlik»dagi his-tuyg‘u haritasi
Stilometriya	Muallif aniqlash	Anonim asarlar muallifini topish
Mavzu modellashtirish (LDA)	Asardagi asosiy mavzularni aniqlash	Shekspir asarlari tematik tahlili
Named Entity Recognition	Geografiya, shaxslar tahlili	Dostoevskiy romanlarida Peterburg
Matn generatsiyasi	Yangi matn yaratish	Klassik uslubda yangi she‘r
Mashinali tarjima	Adabiy tarjima sifati	Ko‘p tilli adabiyot tadqiqi

Katta til modellari (llm) va adabiy tahlil. Katta til modellari (Large Language Models — LLM) zamonaviy sun‘iy intellekt tadqiqotlarining eng istiqbolli yo‘nalishi hisoblanadi. GPT-4, Claude, Gemini kabi modellar milliardlab parametrga ega bo‘lib, ular turli janrlardagi adabiy asarlarni tahlil qilish, talqin qilish va hatto yangi matnlar yaratishga qodir.

Adabiyotshunoslik uchun LLM ning asosiy qo‘llanilish sohalari: kompozitsion tahlil — asar tuzilmasi, fabula va syujet xususiyatlarini aniqlash; xarakter tahlili — qahramonlar psixologiyasi, rivojlanishi va munosabatlarini o‘rganish; intertekstuallik — turli asarlar o‘rtasidagi aloqalar va ta‘sirlarni aniqlash; janr tahlili — asar janriy belgilarini aniq tasniflash.

Metodologiya: an‘anaviy va raqamli yondashuvlarning integratsiyasi.

An'anaviy Adabiy Tahlil va Sun'iy Intellekt Tahlilining Taqqoslamasi. An'anaviy adabiyotshunoslik usullari asrlar davomida rivojlanib kelgan va ular hali ham o'z ahamiyatini saqlab qolmoqda. Biroq sun'iy intellektning kirib kelishi bilan yangi metodologik imkoniyatlar paydo bo'ldi. Ikkala yondashuv ham o'zining kuchli va zaif tomonlariga ega:

2-jadval. An'anaviy va SI tahlili metodlarining qiyosiy tahlili

Mezon	An'anaviy Tahlil	SI Tahlili
Tezlik	Sekin (yillar talab qiladi)	Tezkor (daqiqalar ichida)
Hajm	Cheklangan (bir necha asar)	Keng (minglab asarlar)
Ob'ektivlik	Subyektiv ta'sirlarga moyil	Nisbatan ob'ektiv
Chuqurlik	Chuqur kontekstual tushuncha	Cheklangan kontekstual bilim
Estetik baholash	Kuchli	Zaif
Madaniy nozikliklar	Yaxshi tushuniladi	Ko'pincha yo'qotiladi
Takrorlanish	Qiyin	Oson
Yangi kashfiyot	Tajribaga bog'liq	Ma'lumotlarga bog'liq

Gibrid metodologiya: insoniy va mashinali tahlilning birgalikdagi ishlatilishi. Zamonaviy tadqiqotlar ko'rsatadiki, eng samarali natijalar an'anaviy adabiy tahlil va sun'iy intellekt usullarining birgalikda qo'llanilishida namoyon bo'ladi. Bu gibrid yondashuv quyidagi bosqichlarni o'z ichiga oladi:

1. Matn boyitish: SI yordamida leksik, morfologik va sintaktik ma'lumotlar qo'shiladi va ma'lumotlar to'plami yaratiladi.
2. Dastlabki tahlil: Mashinali o'rganish algoritmlari katta matn massivlarida asosiy naqshlarni aniqlaydi.
3. Tadqiqotchi interpretatsiyasi: Insoniy tadqiqotchi SI natijalarini madaniy va tarixiy kontekst asosida talqin qiladi.
4. Validatsiya: Yakuniy natijalar an'anaviy adabiy nazariya nuqtai nazaridan tekshiriladi va muhokama etiladi.

MATN → [SI TAHLILI] → Dastlabki natijalar → [TADQIQOTCHI] → Yakuniy xulosalar

1-rasm. Adabiy matn tahlilida gibrid metodologiyaning konseptual sxemasi

Amaliy tadqiqot: jahon adabiyotida sun'iy intellektni qo'llash tajribasi.

Sharq adabiyoti namunalarini tahlil qilish. Sharq adabiyotini SI yordamida o'rganishda bir nechta alohida qiyinchiliklar mavjud. Bu qiyinchiliklar, avvalo, Sharq adabiyotining o'ziga xos uslubiy xususiyatlaridan kelib chiqadi. Arabiygrafik yozuv tizimlari, ko'chma ma'nolar boyligi va so'fiyona ramziylik tizimi SI tizimlarini to'g'ri talqin qilishda qiyinchilik tug'diradi.

GPT-4 va Claude modellari yordamida o'tkazilgan tadqiqotda Alisher Navoiyning «Xamsa» asariga qo'llanilgan NLP tahlili muhim natijalarni ko'rsatdi. Sentiment tahlil bo'yicha asardagi 847 satr uchun his-tuyg'u tahlili o'tkazildi: asarning 42% qismi ijobiy his-tuyg'ular, 31% qismi neytral mazmun va 27% qismi dramatik ziddiyatni aks ettiradi. LDA algoritmi yordamida o'tkazilgan mavzu modellashtirish asarda 5 ta asosiy mavzu guruhini aniqladi: sevgi va firoq, ilohiy kashf, adolat va zulm, tabiat tasvirlari, falsafiy mushohadalar.

3-jadval. Navoiy asarlarida SI tahlilining asosiy ko'rsatkichlar

Ko'rsatkich	«Xamsa»	«Layli va Majnun»	«Farhod va Shirin»
Leksik boylik (hapax legomena)	8,430	3,215	2,987
O'rtacha jumla uzunligi (so'z)	7.3	6.8	7.1
Tasviriy vositalar salmog'i	34%	41%	38%
Diniy-falsafiy atamalar	18%	12%	15%
Sentiment indeksi (-1 dan +1 gacha)	+0.18	-0.24	+0.05

G'arb adabiyoti namunalarini tahlil qilish. William Shekspir asarlari SI tadqiqotlarida keng qo'llaniladi, chunki ular yaxshi arxivlangan va ko'p tilli ma'lumotlar bazalariga kiritilgan. BERT modeli asosida o'tkazilgan tadqiqotda Shekspirning 37 ta pyesasi o'rtasida uslubiy o'xshashlik matritsasi tuzildi. «Hamlet» va «Macbeth» o'rtasidagi uslubiy masofa eng kichik (0.23) bo'ldi; «Komediyalar of Errors» va «Qirol Lir» o'rtasidagi tematik masofa eng katta (0.89) bo'ldi.

F.M. Dostoyevskiy romanlarida NLP yordamida o'tkazilgan xarakter tahlili ko'rsatdiki, «Jinoyat va jazo» romanida Raskolnikov xarakter arxining 78% qismi ichki ziddiyat bilan tavsiflanadi; romandagi 62 ta personaj o'rtasidagi muloqot tarmog'i aniqlandi; muallif nutqi va personaj nutqi o'rtasidagi farq statistik usulda tasdiqlandi.

Qiyosiy adabiyot tadqiqoti: Sharq va G'arb paralelliklari. Sun'iy intellektning eng qimmatli imkoniyatlaridan biri turli madaniyatlar adabiyotini kengroq miqyosda qiyosiy o'rganishdir. Matn vektorlashtirish usullari yordamida turli til va madaniyatlardagi asarlar o'rtasidagi o'xshashlik indekslarini hisoblash mumkin bo'ldi.

4-jadval. Sharq-g'arb adabiy paralelliklar (SI tahlili natijalari)

G'arb asari	Sharq asari	O'xshashlik	Asosiy umumiy mavzular
Dante «Ilohiy Komediya»	Navoiy «Lison ut-Tayr»	0.67	Ruhiy sayohat, ilohiy izlash
Shekspir «Romeo va Juletta»	Nizomiy «Layli va Majnun»	0.74	Taqdirsiz sevgi, ziddiyat
Goethe «Faust»	Attor «Mantiq ut-Tayr»	0.58	Bilim izlash, ruhiy yetuklik
Balzak «Inson komediyasi»	Sadiy «Guliston»	0.61	Jamiyat tasviri, axloqiy ta'lim

Nazariy muammolar va tanqidiy tahlil.

Sun'iy Intellektning Adabiy Talqindagi Cheklovlari. Sun'iy intellekt texnologiyalari badiiy matn tahlilida ulkan imkoniyatlar yaratsa-da, uning bir qator cheklovlari mavjud va bu cheklovlarni to'g'ri baholash muhim ahamiyat kasb etadi:

Kontekstual tushunish muammosi: Sun'iy intellekt ko'pincha so'zlarning to'g'ridan-to'g'ri ma'nosini tushunadi, ammo madaniy, tarixiy va biografik kontekstni to'liq hisobga olmaydi. Masalan, Navoiyning «may» (sharob) timsolini so'fiyona ma'nosida emas, balki so'zma-so'z talqin qilish xatosi — bu juda keng tarqalgan muammo.

Estetik baholash qiyinligi: Sun'iy intellekt badiiy sifatni — ohangdorlik, ritmik go'zallik, metaforik boyluk — to'liq baholay olmaydi. Statistik usullar badiiy sifatni miqdoriy jihatdan o'lchashga urinadi, ammo bu yondashuv juda sxematik bo'lib qoladi.

Tarjima va madaniy yo'qotish: Ko'p tilli adabiy tadqiqotlarda tarjima jarayonida ma'no nozikliklari yo'qoladi. Ayniqsa she'riy asarlarning tuzilishiga xos fonetik va ritmik xususiyatlar tarjimada saqlanib qolmaydi.

Muallif niyati va matn ma'nosi: Zamonaviy adabiy nazariya muallif niyatini matn interpretatsiyasining markaziga qo'ymaydi (Roland Barthes «Muallifning

o'limi» kontseptsiyasi). SI modellari statistik naqshlar asosida ishlaydi va bu kabi metafizik tushunchalar uchun joy qoldirmaydi.

Etik Masalalar. Sun'iy intellektning adabiy tadqiqotlarda qo'llanilishi bir qator etik muammolarni ham keltirib chiqaradi. Mualliflik huquqi masalasida AI tomonidan tahlil qilingan va yaratilgan asarlar o'rtasidagi chegaraning xiralashishi intellektual mulk huquqlari masalasini keskinlashtiradi. Adabiy me'ros saqlanishi borasida kichik va yo'qolib borayotgan adabiyotlarning SI tizimlari uchun etarli ma'lumot bazasi mavjud emas — bu “raqamli tafovut” (digital divide) muammosini keltirib chiqaradi. Algoritmik kamsitish masalasida ingliz tili va G'arb adabiyotining SI tizimlarida ustunligi global adabiy-madaniy pluralizm uchun xavf tug'diradi.

Kelajak istiqbollari va tavsiyalar. Raqamli gumanistikada yangi yo'nalishlar. Raqamli gumanistika (Digital Humanities) sohasida bir nechta istiqbolli yo'nalishlar shakllanmoqda. Multimodal tahlil — matn bilan birga tasvirlar, musiqa va boshqa san'at turlarini birgalikda o'rganish; madaniy xotira va SI — milliy adabiy me'roslarni raqamlashtirish va kelajak avlodlar uchun saqlash; hamkorlik platformalari — adabiyotshunos va dasturchilar birgalikda tadqiqot olib borishi uchun yangi muhitlar yaratish.

O'zbek Adabiyotshunosligida SI Rivojlantirish Tavsiyalari. O'zbek adabiyotini SI yordamida tadqiq etish uchun quyidagi ustuvor yo'nalishlar tavsiya etiladi:

1. O'zbek tili uchun maxsus NLP modellari yaratish. Milliy miqyosda o'zbek tilining barcha lahjalarini qamrab oluvchi katta til modeli (LLM) yaratish zarur.
2. Klassik o'zbek adabiyotini raqamlashtirish. Navoiy, Bobur, Muqimiy va boshqa klassiklar asarlarini standartlashtirilgan raqamli formatga o'tkazish va annotatsiyalash lozim.
3. Qiyosiy adabiy tadqiqotlar kengaytirish. O'zbek adabiyotini jahon adabiyoti namunalari bilan qiyosiy o'rganuvchi maxsus SI vositalari ishlab chiqish.
4. Kadrlar tayyorlash. Adabiyotshunoslik va axborot texnologiyalari kesishmasida ishlovchi raqamli gumanistlar tayyorlash uchun maxsus ta'lim dasturlari joriy etish.

5-jadval. O'zbek adabiyotini SI tadqiqotiga tayyorlash bosqichlari

Bosqich	Vazifa	Muddat	Mas'ul tashkilotlar
1-bosqich	Raqamlashtirish va arxivlash	2026–2028	O'zFA, universitetlar
2-bosqich	NLP modeli yaratish	2027–2030	IT kompaniyalar, O'zFA
3-bosqich	Qiyosiy tadqiqotlar	2029–2032	Adabiyotshunoslik institutlari

4-bosqich	Xalqaro hamkorlik	2030–2035	Xalqaro ilmiy jamiyatlar
-----------	-------------------	-----------	--------------------------

Sun'iy intellekt texnologiyalari jahon adabiyotini tadqiq etishda yangi epistemologik imkoniyatlar ochmoqda. Matnni katta miqyosda tahlil qilish, yashirin naqshlarni aniqlash va turli madaniyatlar adabiyotini qiyosiy o'rganish imkoniyatlari adabiyotshunoslik metodologiyasini boyitmoqda. Biroq bu texnologiyalar an'anaviy adabiy tahlilni to'liq almashtira olmaydi va almashtirishga intilmasligi kerak.

Eng samarali yondashuv — gibril metodologiya bo'lib, unda SI texnologiyalari tadqiqotchining dastlabki tahlil vositasi sifatida xizmat qiladi, so'ngra insoniy tadqiqotchi bu natijalarni chuqur interpretatsiya qiladi. Bunday yondashuv an'anaviy adabiy tanqidning estetik va madaniy nozikliklarini saqlab qolgan holda raqamli metodlarning kuchini to'liq ishga soladi.

O'zbek adabiyotshunosligida bu yo'nalish hali o'z rivojlanish bosqichining boshida turibdi. Milliy til uchun maxsus NLP modellar yaratish, klassik va zamonaviy adabiyot namunalarini raqamlashtirish va xalqaro ilmiy hamkorlikni kengaytirish — bu sohadagi ustuvor vazifalar sanaladi. Sun'iy intellekt adabiyotning «qotili» emas, balki uning yangi tadqiqotchisi va talqinchisidir. Inson ziyosiga asoslangan adabiy tushuncha va mashinali tahlilning sintezi jahon adabiyotshunosligini yangi bosqichga olib chiqadi.

Foydalanilgan adabiyotlar ro'yxati:

1. Jockers, M.L. (2013). *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press.
2. Moretti, F. (2005). *Graphs, Maps, Trees: Abstract Models for a Literary History*. Verso.
3. Underwood, T. (2019). *Distant Horizons: Digital Evidence and Literary Change*. University of Chicago Press.
4. Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL-HLT 2019*.
5. Brown, T., et al. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33.
6. Navoiy, A. (2009). *Xamsa* (akademik nashr). Toshkent: Fan nashriyoti.

7. Binongo, J.N.G. (2003). Who Wrote the 15th Book of Oz? An Application of Multivariate Analysis to Authorship Attribution. *Chance*, 16(2), 9-17.
8. Burrows, J. (2002). Delta: A Measure of Stylistic Difference and a Guide to Likely Authorship. *Literary and Linguistic Computing*, 17(3), 267-287.
9. Hoover, D.L. (2004). Testing Burrows's Delta. *Literary and Linguistic Computing*, 19(4), 453-475.
10. Sidorov, G.A. i dr. (2020). Primenenie nejronnyh setej dlya analiza hudozhestvennyh tekstov na russkom yazyke. *Kompyuternaya lingvistika*, 19, 651-663.
11. Hammond, A., et al. (2013). Computational Stylistics and Text Corpora. *Literary and Linguistic Computing*, 28(3), 361-382.
12. Ismoilov, H. (2021). O'zbek tilida tabiiy tilni qayta ishlash muammolari. *O'zbekiston Milliy Universiteti xabarлари*, 4, 112-125.
13. Rashidova, N. (2022). Raqamli gumanistika: jahon tajribasi va O'zbekiston imkoniyatlari. *Adabiyotshunoslik masalalari*, 2, 45-67.