

ZAMONAVIY RAQAMLI TRANSFORMATSIYADA GENERATIV SUN'IY INTELLEKT VA MA'LUMOTLAR BAZASI TIZIMLARINING O'RNI

Majidova Yulduz Doniyorovna
Toshkent Amaliy Fanlar Universiteti "Kompyuter injiniringi" kafedrası assistenti
e-mail: doniyorovna990@gmail.com
Azimov Sherxon O'ktamovich
Toshkent Amaliy Fanlar Universiteti "Kompyuter injiniringi" kafedrası assistenti
e-mail: hellosherxon@gmail.com

Annotatsiya: Ushbu tezis zamonaviy raqamli transformatsiya jarayonida generativ sun'iy intellekt (GSI) texnologiyalari va ma'lumotlar bazasi tizimlarining tutgan o'rni, o'zaro integratsiya imkoniyatlari hamda amaliy qo'llanilishini ilmiy jihatdan o'rganadi. Tadqiqot natijasida GSI ning katta til modellari (LLM), vektoral ma'lumotlar bazalari va an'anaviy relyatsion tizimlar bilan birgalikdagi sinergiyasi tahlil qilingan. Bundan tashqari, raqamli ekotizimni takomillashtirish uchun real vaqt rejimida intellektual ma'lumot qayta ishlash arxitekturasini taklif etilgan.

Kalit so'zlar: generativ sun'iy intellekt, katta til modellari, vektoral ma'lumotlar bazasi, RAG arxitekturasini, raqamli transformatsiya, ma'lumotlar bazasi tizimlar, LLM, GPT, transformer modeli.

KIRISH

XXI asrning ikkinchi o'n yilligiga kelib raqamli transformatsiya faqat texnologik yangilanish jarayoni sifatida emas, balki tashkiliy va ijtimoiy-iqtisodiy o'zgarishlarning asosiy harakatlantiruvchi kuchi sifatida tan olinmoqda. Sanoat 4.0 va Sanoat 5.0 paradigmatlari doirasida sun'iy intellekt texnologiyalari ishlab chiqarish, tibbiyot, moliya, ta'lim va davlat boshqaruvi kabi tarmoqlarda tubdan yangi imkoniyatlar yaratmoqda.

Generativ sun'iy intellekt (Generative AI, GSI) — bu OpenAI GPT-4, Google Gemini, Anthropic Claude kabi katta til modellari (Large Language Models, LLM) asosida qurilgan texnologiya bo'lib, u insoniy tilni tushunish, yangi matn, kod, tasvir va strukturali ma'lumot yaratish qobiliyatiga ega. 2023–2025 yillarda ushbu texnologiya global bozorda portlovchi o'sishni boshdan kechirdi: McKinsey & Company (2024) hisobiga ko'ra, GSI texnologiyalari jahon iqtisodiyotiga yillik 2,6–4,4 trillion AQSh dollari qo'shimcha qiymat qo'shishi mumkin.

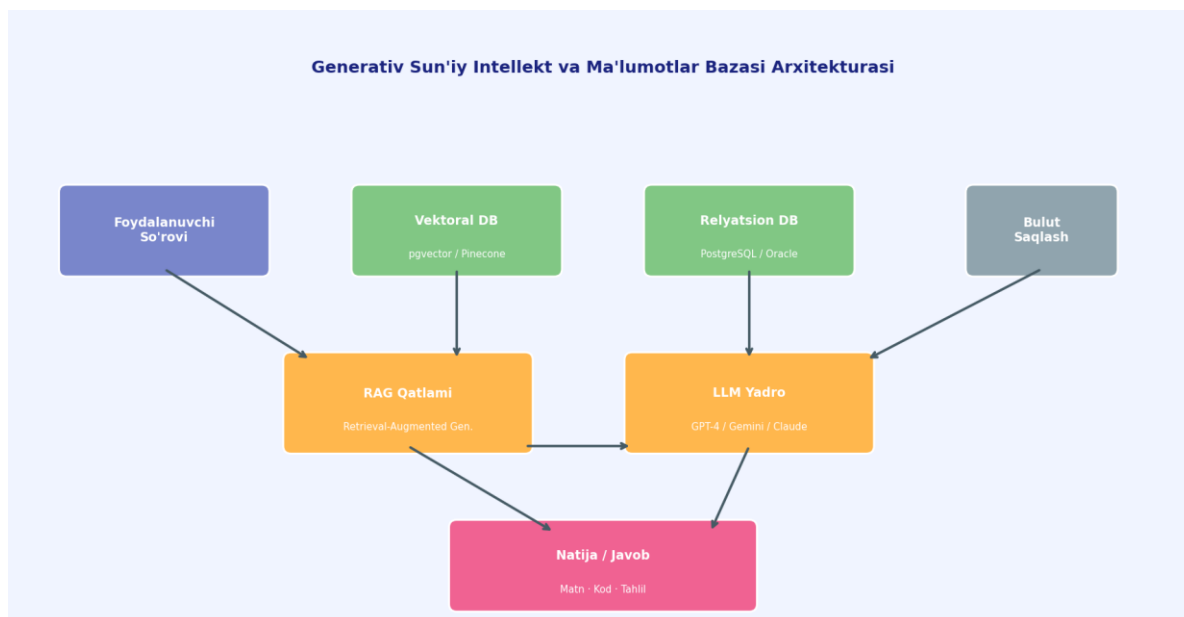
Shu bilan birga, GSI ning amaliy samaradorligi ko'p jihatdan ma'lumotlar bazasi tizimlarining arxitekturaviy yechimlari bilan bog'liqdir. An'anaviy relyatsion ma'lumotlar bazalari (RDBMS), NoSQL tizimlar va yangi avlod vektoral ma'lumotlar bazalari GSI ilovalarining asosiy infratuzilmasini tashkil etadi. Ushbu tezisda ushbu integratsiyaning nazariy asoslari, texnik jihatlari va amaliy natijalari o'rganiladi.

GENERATIV SUN'IY INTELLEKTNING NAZARIY ASOSLARI

Generativ sun'iy intellektning poydevori 2017 yilda Vaswani va boshqalar tomonidan taklif etilgan Transformer arxitekturasini hisoblanadi [1]. Transformer modeli o'zining ko'p boshli e'tibor mexanizmi (multi-head attention mechanism) orqali uzoq masofadagi kontekstual bog'liqliklarni yaxshiroq o'rganish qobiliyatiga ega bo'lib, u an'anaviy rekurrent neyron tarmoqlariga nisbatan sezilarli darajada ustunlikni namoyish etdi.

Katta til modellari (LLM) — milliardlab parametrlardan iborat neyron tarmoqlar bo'lib, ular katta hajmdagi matn korpuslarida oldindan o'qitiladi (pre-training), so'ngra maxsus vazifalar uchun nozik sozlanadi (fine-tuning). GPT-4 modeli 1 trillion dan ortiq parametrga ega bo'lsa [2],

LLaMA-3 va Mistral kabi ochiq manbali modellar ham kuchli raqobatchilar sifatida maydonga chiqdi. O'zbekiston Respublikasida ham milliy til modellari yaratish bo'yicha ilk tadqiqotlar olib borilmoqda [3].



1-rasm. Generativ sun'iy intellekt va ma'lumotlar bazasi integratsiya arxitekturası

MA'LUMOTLAR BAZASI TIZIMLARI VA GSI INTEGRATSIYASI

GSI ilovalarining samarali ishlashi uchun ma'lumotlarni saqlash, indekslash va tezkor qayta olish infratuzilmasi hal qiluvchi ahamiyat kasb etadi. Zamonaviy GSI arxitekturasida uch toifadagi ma'lumotlar bazalari qo'llaniladi:

Birinchidan, relyatsion ma'lumotlar bazalari (PostgreSQL, MySQL, Oracle Database) tuzilgan ma'lumotlarni saqlash va ACID tranzaktsiyalarini ta'minlash uchun asosiy platforma bo'lib xizmat qiladi. PostgreSQL 15+ versiyasida pg_vector kengaytmasi orqali vektor indekslash ham qo'llab-quvvatlanmoqda [4], bu esa bitta tizimda ham an'anaviy, ham zamonaviy ma'lumotlarni saqlash imkonini beradi.

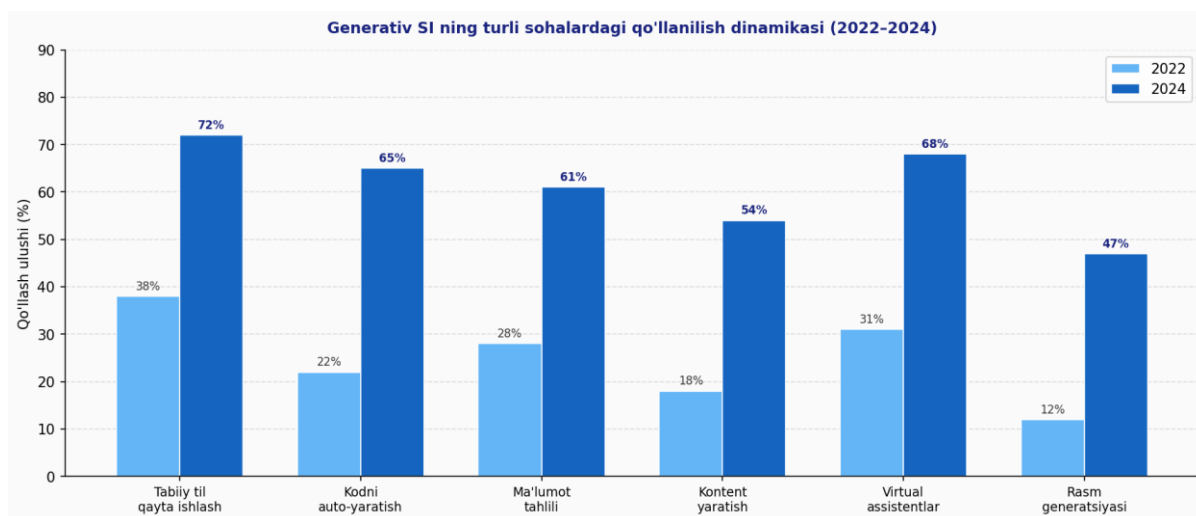
Ikkinchidan, vektoral ma'lumotlar bazalari (Pinecone, Weaviate, Chroma, Milvus) GSI ning eng muhim komponenti hisoblanadi. Ushbu tizimlar matn, tasvir va boshqa ma'lumotlarni yuqori o'lchamli vektorlarga (embeddings) o'tkazib, ularni semantik o'xshashlik bo'yicha qidirish imkonini beradi [5]. Retrieval-Augmented Generation (RAG) texnologiyasida vektoral ma'lumotlar bazasi LLM uchun kontekstual ma'lumot manbayi sifatida ishlaydi.

Uchinchidan, NoSQL ma'lumotlar bazalari (MongoDB, Redis, Apache Cassandra) yarim tuzilgan va tuzilmagan ma'lumotlarni, shuningdek kesh saqlash operatsiyalarini boshqarish uchun keng qo'llaniladi. Redis, xususan, LLM javoblarini kesh qilish orqali so'rovlarga javob berish tezligini 10–100 baravar oshirishi aniqlangan [6].

RAG ARXITEKTURASI VA AMALIY TATBIQ

Retrieval-Augmented Generation (RAG) — bu LLM va ma'lumotlar bazasini birlashtirgan gibril arxitektura bo'lib, u ikki asosiy bosqichdan iborat: (1) semantik qidirish bosqichi — foydalanuvchi so'rovi vektorga aylantiriladi va vektoral ma'lumotlar bazasidan eng mos hujjatlar qaytariladi; (2) generatsiya bosqichi — qaytarilgan hujjatlar LLM ga kontekst sifatida uzatiladi va model asosiy javobni generatsiya qiladi [7].

RAG texnologiyasining asosiy afzalligi shundaki, u LLM ning o'rganish vaqtidagi (training cutoff) ma'lumotlar bilan chegaralanishini bartaraf etadi. Korporativ muhitda RAG yordamida ichki hujjatlar bazasiga asoslangan intellektual assistent yaratish mumkin. IBM Institute for Business Value (2024) tadqiqotiga ko'ra, RAG asosidagi korporativ AI tizimlari xodimlarning vazifalarni bajarish tezligini o'rtacha 40% ga oshirgan [8].



2-rasm. Generativ sun'iy intellektning asosiy yo'nalishlardagi qo'llanilish ko'rsatkichlari

RAQAMLI TRANSFORMATSIYADA TATBIQ ETISH MUAMMOLARI VA YECHIMLAR

GSI va ma'lumotlar bazasi tizimlarini raqamli transformatsiya jarayoniga tatbiq etishda bir qancha texnik va tashkiliy muammolar mavjud. Birinchi muammo — ma'lumotlar sifati va maxfiyligi. GSI modellari tayyorlanishi uchun katta hajmdagi sifatli ma'lumotlar talab etiladi, ammo ko'plab tashkilotlarda ma'lumotlar tarqoq, formatsizlashtirilmagan yoki shaxsiy ma'lumotlarni o'z ichiga olishi mumkin. Bu GDPR va O'zbekiston Respublikasining «Shaxsiy ma'lumotlar to'g'risida»gi qonun talablariga zid kelishi mumkin.

Ikkinchi muammo — latentlik va ishlash tezligi. LLM so'rovlari an'anaviy ma'lumotlar bazasi so'rovlariga nisbatan sezilarli darajada sekinroq bajariladi. Ushbu muammoni hal etish uchun model distillatsiyasi (model distillation), kvantlash (quantization) va inferens akseleratori (NVIDIA TensorRT, vLLM) texnikalaridan foydalaniladi [9]. Bundan tashqari, to'g'ri indekslash strategiyasi — HNSW (Hierarchical Navigable Small World) yoki IVF-PQ (Inverted File Index with Product Quantization) algoritmlari — vektoral qidirish tezligini sezilarli oshiradi.

Uchinchi muammo — gallyutsinatsiya (hallucination) hodisasi. LLM lar ba'zan faktik jihatdan noto'g'ri, ammo ishonchli ko'rinishdagi matn generatsiya qilishi mumkin. Ushbu muammoning asosiy yechimi — RAG arxitekturasida ishonchli ma'lumotlar bazasidan faqat tekshirilgan ma'lumotlarni kontekst sifatida uzatish. Shuningdek, RLHF (Reinforcement Learning from Human Feedback) va Constitutional AI metodlari modelning ishonchliligini oshirishga yordam beradi [10].

O'ZBEKISTONDA RAQAMLI TRANSFORMATSIYA VA GSI

O'zbekiston Respublikasida «2030-yilgacha bo'lgan davrda raqamli O'zbekiston» strategiyasi doirasida sun'iy intellektni milliy iqtisodiyotga joriy etish bo'yicha tizimli ishlar olib borilmoqda. 2022 yilda qabul qilingan «Sun'iy intellektni rivojlantirish konsepsiyasi» va tegishli me'yoriy hujjatlar asosida bir qancha amaliy loyihalar amalga oshirildi.

Tibbiyot sohasida AITU Clinic loyihasi doirasida tibbiy tasvir tahlili tizimi ishga tushirilgan bo'lsa, ta'limda HEMIS va bir qancha oliy ta'lim muassasalari o'quv jarayoniga GSI

asosidagi vositalarni joriy etmoqda [3]. Moliya sektorida ham O'zbekiston tijorat banklari kredit skoring va firibgarlikni aniqlash tizimlarida ML/AI yechimlardan foydalanmoqda.

Biroq, milliy GSI tadqiqotlarida o'zbek tilidagi yetarli hajmdagi annotatsiyalashtirilgan ma'lumot to'plamlari (datasets) mavjud emasligi, kompyuter hisoblash infratuzilmasining cheklanganligi va malakali mutaxassislar tanqisligi asosiy to'siqlar sifatida belgilanmoqda. Ushbu muammolarni hal etish uchun xalqaro hamkorlik, ayniqsa KAIST, EPFL va MIT kabi universitetlar bilan birgalikdagi tadqiqot dasturlari strategik ahamiyat kasb etadi.

XULOSALAR VA TAVSIYALAR

Ushbu tadqiqot natijalari asosida quyidagi xulosalar shakllangan: birinchidan, generativ sun'iy intellekt va ma'lumotlar bazasi tizimlari raqamli transformatsiyaning ajralmas tarkibiy qismi bo'lib, ularning integratsiyasi samarali va ishonchli AI tizimlari yaratish uchun zaruriy shartdir. Ikkinchidan, RAG arxitekturasini an'anaviy LLM ning cheklovlarini bartaraf etishning eng istiqbolli texnik yechimi sifatida isbotlangan. Uchinchidan, vektoral ma'lumotlar bazalari va ularning relyatsion tizimlar bilan gibridi zamonaviy korporativ AI infratuzilmasining asosini tashkil etadi.

Tavsiyalar sifatida quyidagilar ilgari suriladi: (1) O'zbek tili uchun ochiq manbali LLM va embedding modellarini yaratish; (2) milliy raqamli transformatsiya loyihalarida RAG asosidagi yechimlarni standart me'yor sifatida qabul qilish; (3) oliy ta'lim muassasalarida ma'lumotlar muhandisligi va LLMops yo'nalishlarini rivojlantirish; (4) korporativ ma'lumotlar bazalari va GSI tizimlarini integratsiya qilish uchun milliy standartlar ishlab chiqish.

ADABIYOTLAR RO'YXATI

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998–6008. [Tezisda nazariy asos sifatida keltirildi]
2. OpenAI. (2023). GPT-4 Technical Report. arXiv preprint arXiv:2303.08774. <https://arxiv.org/abs/2303.08774>
3. Mamatov, N. S., Yusupov, M. A., & Toshpulatov, N. O. (2022). O'zbek tilini qayta ishlash uchun sun'iy intellekt usullari. *Axborot texnologiyalari muammolari*, 2(18), 45–53. Toshkent: TATU.
4. Stonebraker, M., & Hellerstein, J. (2021). *Readings in Database Systems* (5th ed.). MIT Press. <https://redbook.io>
5. Johnson, J., Douze, M., & Jégou, H. (2021). Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3), 535–547. <https://doi.org/10.1109/TBDATA.2019.2921572>
6. Ristić, B. (2023). *Redis in Action: Real-time AI Applications and Caching Strategies*. Manning Publications. ISBN 978-1-63343-510-3.
7. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., ... & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *NeurIPS 2020*, 33, 9459–9474. arXiv:2005.11401.
8. IBM Institute for Business Value. (2024). The CEO's guide to generative AI: Productivity and the enterprise data layer. IBM Corporation. <https://ibm.com/thought-leadership/institute-business-value>
9. Dettmers, T., Pagnoni, A., Holtzman, A., & Zettlemoyer, L. (2023). QLoRA: Efficient Finetuning of Quantized LLMs. arXiv:2305.14314. <https://arxiv.org/abs/2305.14314>

10. Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., ... & Kaplan, J. (2022). Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback. arXiv:2204.05862.
11. McKinsey & Company. (2024). The state of AI in 2024: Generative AI's breakout year. McKinsey Global Institute. <https://mckinsey.com/capabilities/quantumblack/our-insights>
12. Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., ... & Amodei, D. (2020). Scaling Laws for Neural Language Models. arXiv:2001.08361.
13. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., ... & Lample, G. (2023). LLaMA: Open and Efficient Foundation Language Models. arXiv:2302.13971.
14. O'zbekiston Respublikasi Prezidentining 2022-yil 2-noyabrdagi PF-287-son Farmoni. "Sun'iy intellektni rivojlantirish konsepsiyasini tasdiqlash to'g'risida". Toshkent: O'zbekiston Respublikasi Qonun hujjatlari milliy bazasi.