

МА'LUMOTLAR BAZASIDA SO'ROVLAR YARATISH

Tojimatov Israiljon Nurmamatovich

Farg'ona davlat universiteti

israiltojimatov@gmail.com**Abdubannonova O'g'ilxon**

Farg'ona davlat universiteti

abdubannonova1306@gmail.com**Annotatsiya**

Ushbu maqola tabiiy tildagi so'rovlarni ma'lumotlar bazasiga mos keladigan so'rovlarga aylantirish tizimining umumiy tuzilmasi va foydalaniladigan usullarini tasvirlaydi. Tabiiy til so'rovlarini tuzishda yuzaga kelishi mumkin bo'lgan muammolar tahlil qilingan va bunday so'rovlarni o'zgartirish uchun cheklovlar belgilangan. Shuningdek, tabiiy til so'rovlarini rasmiy tilga aylantirish uchun ishlab chiqilgan metodologiyaning asosiy tamoyillari keltirilgan.

Kalit so'zlar: ma'lumotlar bazasi, tabiiy til, mavzu sohasi, kompyuter lingvistikasi, xavfsizlik, relyatsion bazalar, grafik bazalar, sintaktik daraja, morfologik daraja, grafematik daraja

Аннотация

Он описывает общую структуру и инструменты, используемые для сопоставления запросов на естественном языке с базой данных. Анализируется проблема согласования запросов на естественном языке в контракте и установления таких запросов. За бессловесные иллюстрации методологии создания естественных, естественных запросов к официальному языку.

Ключевые слова: база данных, естественный язык, предметная область, лингвистика, компьютер, реляционные базы данных, графические базы данных, синтаксический уровень, морфологический уровень, графематический уровень.

Annotation

It describes the general framework and tools used to match natural language queries to a database. The contract discusses the discussion of natural language queries and the problem of setting up such queries. For wordless illustrations of natural, natural query production methodology for formal language.

Key words: database, natural language, subject area, linguistics, computer, relational databases, graphic databases, syntactic level, morphological level, graphematic level.

Kirish

Texnologik jarayonlarni avtomatlashtirish vazifasi muloqot tizimlari tarkibidagi ma'lumotlarni qayta ishlashning maxsus texnikalarini talab qiladi. Xususan, avtomatlashtirilgan loyihalash tizimlari (ALTI) uchun foydalanuvchilarni o'qitish davrini qisqartirish maqsadida tabiiy tildagi so'rovlarni ma'lumotlar bazasiga moslashtirish masalasi dolzarbdir. Bu masalani foydalanuvchi so'rovlarini tabiiy til so'rovlaridan ma'lumotlar bazasi rasmiy tili so'rovlariga avtomatik o'zgartirish orqali hal qilish mumkin.

Ishning maqsadi tabiiy tildagi jummalarni ma'lumotlar bazasiga mos keladigan so'rovlarga aylantirish metodlarini ishlab chiqishdan iborat. Ushbu maqsadga erishish uchun quyidagi vazifalar bajarilishi zarur:

Tabiiy tildagi jummalarni mashina tomonidan qayta ishlashga yaroqli axborotga aylantirish.

Matndan ajratib olingan ma'lumotlar asosida yaratilgan so'rovlarni ma'lumotlar bazasiga translyatsiya qilish.

Ushbu usullar rasmiy tillarni maxsus o'rganmagan foydalanuvchilarga mo'ljallangan interfeyslar yaratish va kompyuter bilan tabiiy til orqali muloqotni modellashtirish uchun qo'llanilishi mumkin. Mazkur texnologiyalar tibbiyot, ta'lim,

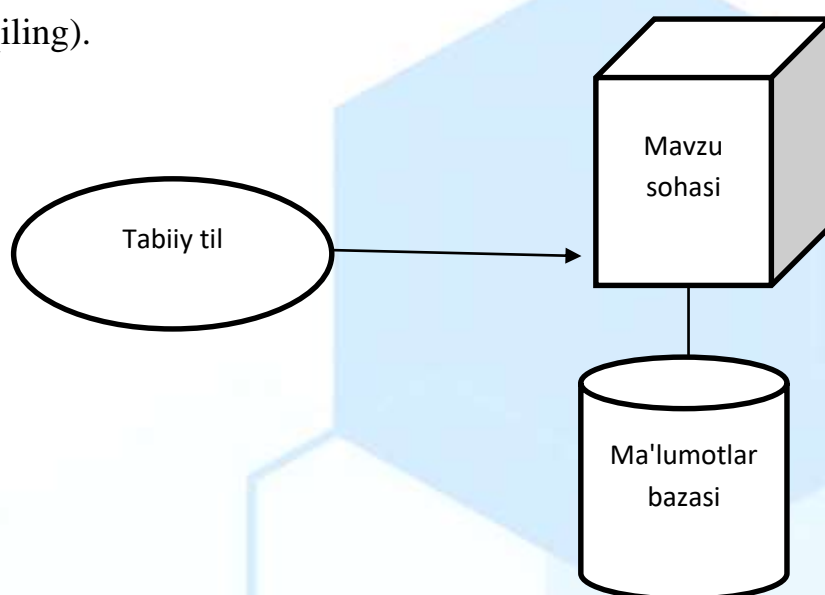
kutubxonalar kabi ko'plab sohalarda, foydalanuvchilarning ma'lumotlar bazasiga kirishiga yordam beruvchi vositalar talab etiladigan joylarda qo'llanilishi mumkin.

2. Tabiiy tildagi so'rovlarni qayta ishlash tizimini tavsifi

Mazkur ish inson bilimlarining ko'plab yo'nalishlariga tegishlidir. Tabiiy til so'rovlarida predmet sohasiga tegishli tushunchalarni ajratib olish talab etiladi, bu esa tabiiy til matnlarini tahlil qilishni taqozo etadi. Tushunchalarni aniqlash uchun tabiiy tilni qayta ishlash tizimining umumiy tuzilmasini tushunish zarur. Tabiiy til ma'lum bir predmet sohasiga kuchli bog'liq bo'lganligi sababli (masalan, "semafor" tushunchasi kompyuter fanlarida va dengizchilikda farq qiladi), avvalo, rivojlantiriladigan tizim qaysi predmet sohasiga mo'ljallanganligini aniqlash kerak. Bu jarayonda tahlil faqat modelga tegishli muhim jihatlar bilan cheklanadi.

Mazkur ish umumiy xarakterdagi masalani ko'rib chiqadi – kelgusida muayyan predmet sohasiga moslashishi mumkin bo'lgan metodlarni ishlab chiqish. Umuman olganda, predmet sohasi ob'ektlari orasidagi murakkablik va aloqalar soni cheklanmagan. Predmet sohasining tushunchalarini ajratish masalasi maxsus metodlar yordamida hal etiladi, bunda tushuncha formal ta'rifi bilan belgilanadi, bu boshqa sohalardagi bir xil nomli tushunchalarning ta'riflariga mos kelmaydi.

Agar predmet sohasidan ajratib olingan tushunchalardan iborat ma'lumotlar bazasi mavjud bo'lsa, tabiiy tildagi so'rovlarni ma'lumotlar bazasiga mos so'rovlarga aylantirish metodlarini ishlab chiqish haqida gapirish maqsadga muvofiqdir. Ishlab chiqilgan tizim shartli quyidagi sxema orqali tasvirlanishi mumkin. (Rasmga murojaat qiling).



1-rasm. Ma'lumotlar bazasi sxemasi

Tadqiqot vazifalarini baholaganda, quyidagi bilim sohalarini tahlil qilish zarurligi qayd etilgan:

Kompyuter lingvistikasi metodlari, bu tabiiy til ma'lumotlarini "mashina" formatidagi ma'lumotlarga o'zgartirish imkonini beradi.

So'rovlarga qo'yiladigan cheklovlarni shakllantirish, bu esa kiritilgan va chiqarilgan ma'lumotlarni bir ma'noli aniqlash imkonini beradi.

Ma'lumotlar bazasi bilan bog'liq bo'lgan ob'ektlarni tabiiy tildan ajratish, bu esa ma'lumotlarni tabiiy til so'rovlaridan bazaga mos so'rovlarga o'zgartirish uchun zarur bo'lgan faktlar va qoidalarni aniqlash imkonini beradi.

Ishlaydigan ma'lumotlar bazasi turini tanlash, shu jumladan, ushbu bazalarga nisbatan cheklovlar va ular bilan ishlashni amalga oshiradigan dasturiy vositalarni belgilash.

So'rovlar yaratish metodlari, bu foydalanuvchi yakuniy so'rovlarni qanday qilib hosil qilishini tushunishga yordam beradi.

Mazkur vazifalar turli metod va dasturiy vositalar yordamida hal etilishi mumkin, bu esa keyingi bo'limlarda muhokama qilinadi.

2.1. Amalga oshiriladigan ma'lumotlar bazasini tanlash

Bugungi kunda ma'lumotlar bazasining yagona aniq ta'rifi mavjud emas. Eng keng tarqalgan ta'rif quyidagicha:

Ma'lumotlar bazasi — ma'lumotlarni ma'lum qoidalarga muvofiq tashkil qilgan holda, ularni tasvirlash, saqlash va boshqarish uchun umumiy tamoyillarni o'z ichiga oladi. U amaliy dasturlardan mustaqil bo'lib, predmet sohasining axborot modeli hisoblanadi.

Ma'lumotlar bazasi tashkil etilishiga qo'yiladigan asosiy talablar quyidagilar:

Ma'lumotlarning ortiqcha emasligi (bir nusxada mavjud bo'lishi);

Ko'p foydalanuvchilar uchun qo'shma foydalanish imkoniyati;

Ma'lumotlarga samarali kirish (tez javob berish vaqti);

Butunlikni ta'minlash (ichki mantiqqa moslik);

Xavfsizlik;

Nosozliklardan keyin tiklash;

Amaliy dasturlardan mustaqillik.

Ma'lumotlar bazasi modellariga asoslangan tizimlarning afzallik va kamchiliklarini o'rganish zarur. Hozirgi kunda ma'lumotlar bazasi boshqaruv tizimlari (MBBT) quyidagicha tasniflanadi:

1. Dorasyonal bazalar (invert qilingan ro'yxatlar, ierarxik va tarmoqli tuzilmalar).

2. Relyatsion bazalar, tarkibida struktura, yaxlitlik va boshqaruv komponentlarini o'z ichiga oladi.

3. Postrelyatsion bazalar, ya'ni zamonaviy talablar uchun kengaytirilgan tuzilmalar.

3. Dorelyatsion bazalar

Relyatsion bazalarning kashshoflari hisoblanadi. Ularning asosiy xususiyatlari:

Ma'lumotlarni boshqarishning rivojlangan usullari, past darajadagi tashqi xotira bilan ishlash.

Foydalanuvchiga qo'lda samarali tizimlarni yaratish imkoniyati.

Kamchiliklari orasida qiyin foydalanish va fizik tashkilotni bilish zaruriyati mavjud.

Bugungi kunda dorelyatsion bazalar keng ishlatilmaydi, ularning o'rnini relyatsion va boshqa zamonaviy tizimlar egallagan.

4. Relyatsion bazalar

Relyatsion baza o'zaro bog'liq jadvallar to'plamidan iborat. Har bir jadval ma'lum bir ob'ekt haqida ma'lumotni saqlaydi, satrlar esa alohida ob'ektlarni ifodalaydi.

Afzalliklari:

Ma'lumotlarni qulay va tushunarli shaklda aks ettiradi.

Operatsiyalarni matematik jihatdan aniq ta'riflash imkonini beradi.

Ma'lumotlar bilan deklarativ va protseduraviy usulda ishlash imkonini beradi.

Kamchiliklari:

Eng sekin kirish tezligi.

Rivojlantirish jarayonining mehnatsevarligi.

5. Postrelyatsion bazalar (NoSQL)

Postrelyatsion texnologiyalar katta hajmdagi ma'lumotlar va ko'p foydalanuvchilarni qo'llab-quvvatlash uchun mo'ljallangan. Ushbu bazalar quyidagi turlarga bo'linadi:

Hujjatga yo'naltirilgan.

"Kalit-qiyamat" saqlash tizimlari.

Grafik bazalar.

Grafik bazalar tabiiy grafik tuzilmalar bilan ishlashda samaraliroq bo'lishi mumkin. Ushbu bazalar, ayniqsa, so'rovlarni sintaktik daraxtlar yordamida avtomatlashtirishda qulay hisoblanadi.

5.1. Kompyuter lingvistika

Kompyuter lingvistika – bu tilshunoslikning amaliy sohasidan bo'lib, kompyuter texnologiyalari va dasturlaridan foydalanish orqali tilni modellashtirishga qaratilgan. Kompyuter lingvistikasining maqsadi — ma'lum bir sharoitda yoki muammolar sohasida tilning ishlashini simulyatsiya qilishdir. Bu soha kompyuterlarni tilshunoslik vazifalarini hal qilishda, xususan tabiiy tilni qayta ishlashda qo'llaydi.

Kompyuter lingvistikasining ob'ekti — tilning ma'nosi va tuzilishini ifodalovchi turli xil matnlar (yazma yoki og'zaki nutq) bo'lib, ular ma'lum bir ma'lumotni o'z ichiga oladi va ma'lum bir struktura bilan tashkil etiladi. Bunday matnlar sintaktik va morfologik tahlilni talab qiladi.

Hozirgi kunda tabiiy tildagi matnlarni tahlil qilishda quyidagi strukturalar farqlanadi:

Sintaktik daraja — jumlar yoki gaplar tuzilishi.

Morfologik daraja — so'z shakllari, grammatik shakllar (masalan, "aktrisa", "aktrisaning" va boshqalar).

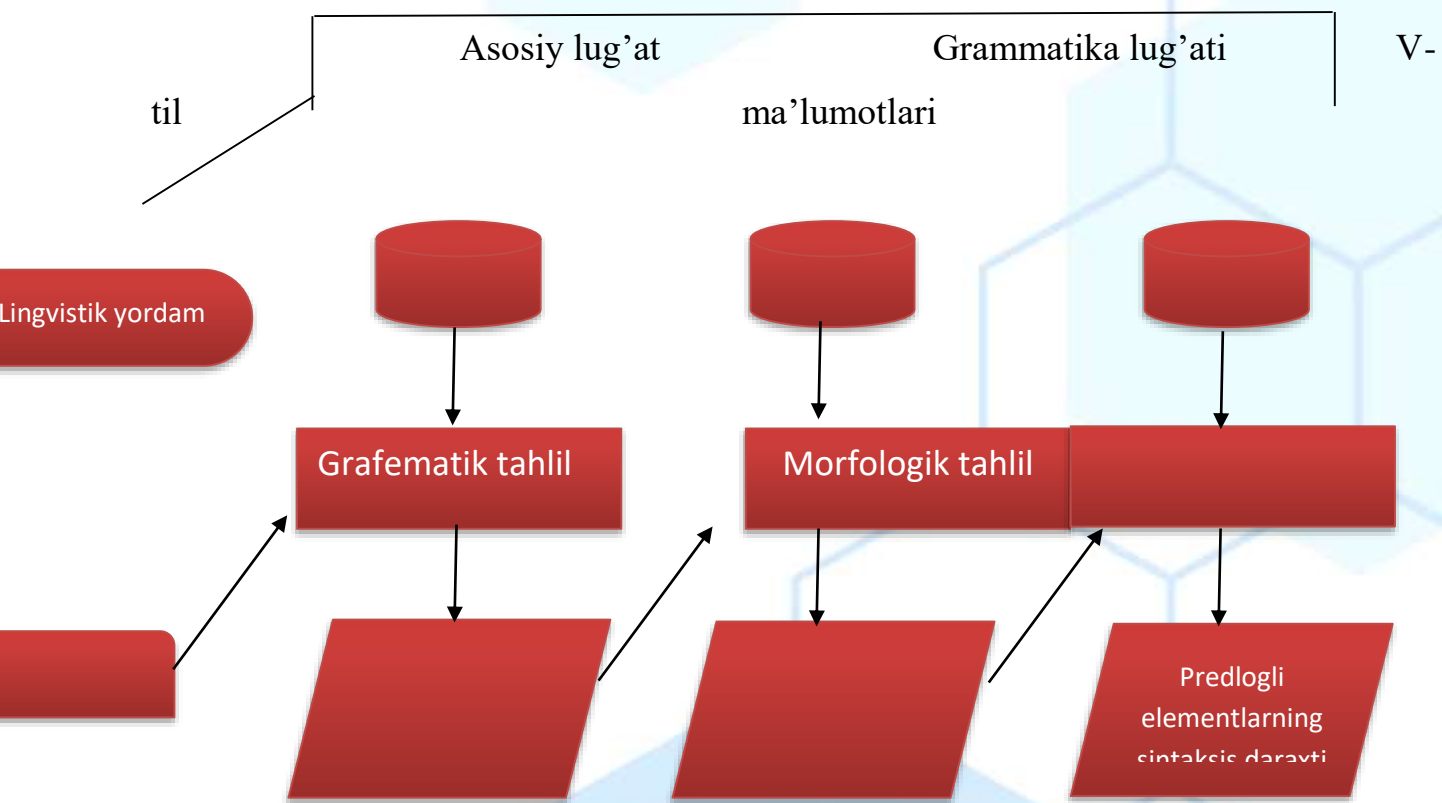
Grafematik daraja — matnning strukturalarini (so'zlar, tinish belgilari) tokenlarga ajratish.

Matnlarni qayta ishlash jarayonida maqsad — ularning grammatik va sintaktik tuzilishini aniqlash, shuningdek, ma'no va kontekstni tushunishdir. Bu jarayon ko'plab tahlil bosqichlarini o'z ichiga oladi, jumladan:

Grafematik tahlil — matni belgilarga ajratish (so'zlar, tinish belgilari va boshqa tuzilmalar).

Morfologik tahlil — soʻz shakllarini aniqlash va ularning grammatik ma'nosini tushunish.

Sintaktik tahlil — jumlar va gaplar o'rtasidagi munosabatlarni aniqlash, ya'ni sintaktik daraxtni qurish.



2-rasm. Maʼlumotlar tahlili grafikasi

Shuningdek, tabiatda o'xshash bo'lgan (leksik, morfologik, sintaktik) omonimiyalik (bitta shakldagi so'zlarning turli ma'nolarini tushunish) kabi masalalarni hal qilish ham muhimdir.

5.2. Ma'lumotlar bazasiga so'rovlar yaratish

Tabiiy til so'rovlarini formal so'rovlarga aylantirishdan so'ng, keyingi bosqichda ularni ma'lumotlar bazasiga to'g'ri so'rovga aylantirish zarurati tug'iladi. Tabiiy til so'rovini formal so'rovga aylantirishda ikkita asosiy yo'l mavjud:

1. Semantik tarmoqlar — bu ob'ektlar va ularning aloqalarini aks ettiruvchi tarmoq modelidir. Semantik tarmoq orqali so'rovlarni yaratish ma'lumotlar bazasi bilan ishlashda qulaylik yaratadi, chunki u ma'lumotlar orasidagi munosabatlarni aniq va oddiy tarzda tasvirlaydi.

2. Mantiqiy dasturlash vositalari — mantiqiy dasturlash tilida (masalan, Prolog) sintaktik daraxtlarni mantiqiy qoidalarga aylantirish mumkin. Bu usul ko'p hollarda grafik va relyatsion bazalar bilan ishlashda samarali bo'ladi, chunki ularni mantiqiy modellarga o'tkazish orqali so'rovlar tuzilishini aniq belgilash mumkin.

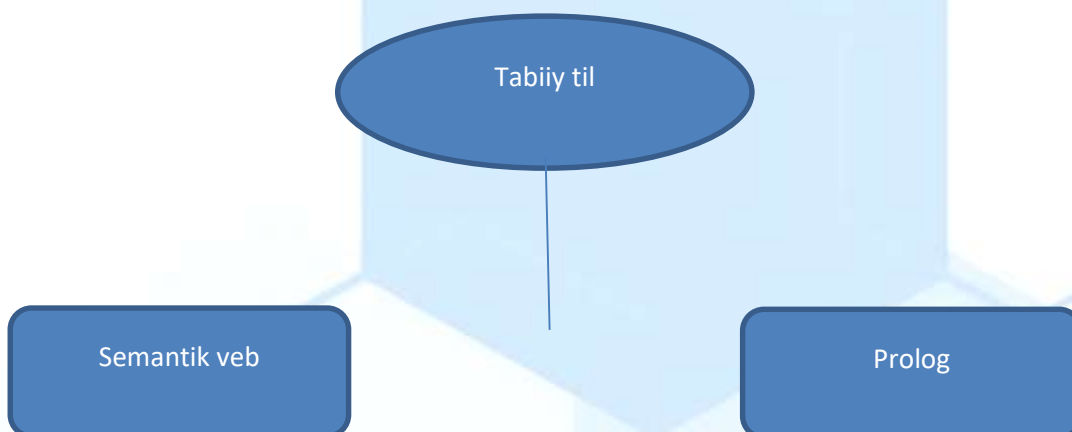
Prolog — mantiqiy dasturlash tili bo'lib, u obyektlar va ularning munosabatlarini ta'riflashda, shuningdek, masalalarni yechishda qo'llaniladi. Prolog tilidagi dasturlar predikat mantiqiga asoslangan bo'lib, ular yordamida ma'lumotlar bazasiga so'rovlar yozish va javoblarni olish mumkin.

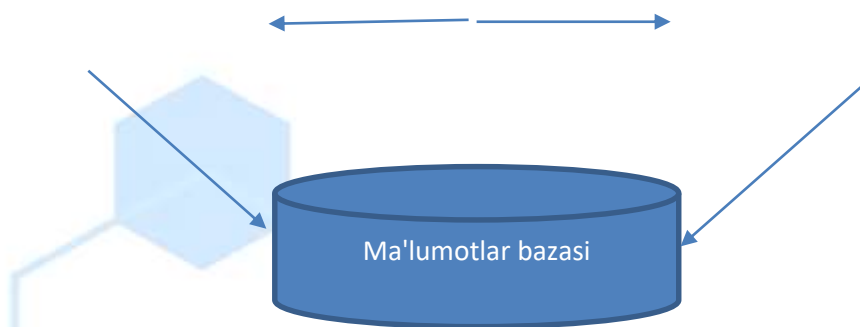
Prolog tilidan foydalanishning afzalliklari:

Obyektlar va munosabatlarni aniq ifodalash.

Ma'lumotlar bazalarini tez va samarali tahlil qilish imkoniyati.

So'rovlarni aniq va tuzilgan holda olish imkoniyati.





3-rasm Sintaksis daraxti

Bundan tashqari, Prolog tili bilan ishlashda tizimning ish faoliyatini kengaytirish va yangi so'rov turlarini qo'shish oson. Bunday metod yordamida avtomatik so'rovlar tuzishni osonlashtirish mumkin.

Xulosa

Ushbu tizimni ishlab chiqish jarayonida tabiiy tildagi so'rovlarni ma'lumotlar bazasiga formal so'rovlarga aylantirishda quyidagi xulosalar chiqarildi:

Tizimda predmet sohasining aniqligi muhim ahamiyatga ega bo'lib, so'rovlarni aniq va samarali tahlil qilishga yordam beradi.

Relyatsion va postrelyatsion grafik ma'lumotlar bazalarini tanlash mumkin, ammo loyiha uchun eng samarali yechim sifatida relyatsion bazalar tanlandi.

Tabiiy tildagi so'rovlarni formal so'rovlarga aylantirishda mantiqiy dasturlash usullaridan foydalanish tizimni kengaytirishda va optimallashtirishda yordam beradi.

Semantik tarmoqlar va mantiqiy dasturlash vositalari yordamida so'rovlar samarali tarzda shakllantiriladi.

Tizimning to'g'ri ishlashi va samaradorligini ta'minlash uchun, so'rovlarning semantik tahlilini amalga oshirish zarurati bor, ayniqsa, omonimiyalik va ko'p ma'noli so'zlar bilan bog'liq muammolarni hal qilishda. Yuqoridagi metod va vositalarni hisobga olgan holda, tabiiy tildagi so'rovlarni ma'lumotlar bazalariga to'g'ri va samarali tarzda aylantirish tizimi ishlab chiqilishi kutilmoqda. Ushbu tizimning samaradorligi va samaradorligini baholash, qo'llaniladigan dasturiy vositalarning samaradorligini tekshirish, shuningdek, mavjud metodlarni yaxshilash va

optimallashtirish uchun imkoniyat yaratadi. Tizimni ishlab chiqishda mantiqiy dasturlash tilidan foydalanish va u bilan bog‘liq bilim bazasini yaratish tizimni yanada samarali va moslashuvchan qiladi.

Foydalanilgan adabiyotlar

1. Бородин Д.С., Строганов Ю.В., 2016 & К задаче составления запросов к базам данных на естественном языке, Новые информационные технологии в автоматизированных системах, МГТУ им. Н.Э Баумана.
2. Tojimatov, I. N., Topvoldiyeva, H., Karimova, N., & Inomova, G. (2023). GRAFIK MA'LUMOTLAR BAZASI. Евразийский журнал технологий и инноваций, 1(4), 75-84.
3. Nurmamatovich, T. I., & Azizjon o'g, N. A. Z. (2024). The SQL server language and its structure. American Journal of Open University Education, 1(1), 11-15.
4. Nurmamatovich, T. I. (2024). MY SQL MISOLIDA LOYIHA YARATISH. Ta'limda raqamli texnologiyalarni tadbiiq etishning zamonaviy tendensiyalari va rivojlanish omillari, 31(2), 82-90.
5. Ro'zimatov, J. I., & Nurmamatovich, T. I. (2024). SQL tili tarixi, vazifasi, turlari va rejimlari.
6. Nurmamatovich, T. I. (2024). NORMALLASHTIRISH. NORMAL FORMALAR. worldly knowledge conferens, 7(2), 597-599.
7. Isroil, T. (2023). NOSQL MA'LUMOTLAR BAZASI: TANQIDIY TAHLIL VA TAQQOSLASH. IJODKOR O'QITUVCHI, 3(28), 134-146.
8. Qodirjonova, N., Tursunova, N., Parpiboyev, N., & Tojimatov, I. (2023). BIR KOMPYUTERDA KATTA MA'LUMOTLAR BILAN ISHLASH. Центральноазиатский журнал образования и инноваций, 2(4), 104-111.
9. Tojimatov, I., & Doniyorbek, A. (2023). KATTA HAJMLI MA'LUMOTLAR AFZALLIKLARI VA

- КАМШИЛКЛАРИ. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 18(6), 66-70.
10. Ne'matillayev, A. N., Abduqahhorov, I. I., & Tojimatov, I. (2023). BIG DATA TEXNOLOGIYALARI VA UNING MUAMMOLARI. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 19(1), 61-64.
 11. Tojimatov, I. N., Olimov, A. F., Khaydarova, O. T., & Tojiboyev, M. M. (2023). CREATING A DATA SCIENCE ROADMAP AND ANALYSIS. PEDAGOGICAL SCIENCES AND TEACHING METHODS, 2(23), 242-250.
 12. Gulhayo, M., Gulnoza, A., & Isroil, T. (2023). MA'LUMOTLARNI QAYTA ISHLASHDA ERP TIZIMLARI. MA'LUMOTLARNI QAYTA ISHLASHDA SAP TIZIMLARI. Journal of Integrated Education and Research, 2(4), 87-89.
 13. Isroil, T. (2023). NOSQL MA'LUMOTLAR BAZASI: TANQIDIY TAHLIL VA TAQQOSLASH. IJODKOR O'QITUVCHI, 3(28), 134-146.
 14. Saidjamolova, B. M., & Tojimatov, I. N. (2023). BIZNESDA «BIG DATA» TEXNOLOGIYALARI VA ULARNING AHAMIYATI. Лучшие интеллектуальные исследования, 11(4), 56-63.
 15. Tojimatov, I. N., Topvoldiyeva, N., Karimova, N., & Inomova, G. (2023). GRAFIK MA'LUMOTLAR BAZASI. Евразийский журнал технологий и инноваций, 1(4), 75-84.
 16. Тожимамаатов, И. Н. (2023). ЗАДАЧИ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ. PEDAGOG, 6(4), 514-516.
 17. Mamasidiqova, I., Husanova, O., Madaminova, A., & Tojimatov, I. (2023). Data Mining Texnologiyalari Metodlari Va Bosqichlari Hamda Data Science Jarayonlar. Центральноазиатский журнал образования и инноваций, 2(3 Part 2), 18-21.
 18. Tojimatov, I. N., Olimov, A. F., Khaydarova, O. T., & Tojiboyev, M. M. (2023). CREATING A DATA SCIENCE ROADMAP AND

ANALYSIS. PEDAGOGICAL SCIENCES AND TEACHING METHODS, 2(23), 242-250.

19. Tojimatov, I. N., Topvoldiyeva, H., Karimova, N., & Inomova, G. (2023). GRAFIK MA'LUMOTLAR BAZASI. Евразийский журнал технологий и инноваций, 1(4), 75-84.
20. Ne'matillayev, A. H., Abduqahhorov, I. I., & Tojimatov, I. (2023). BIG DATA TEXNOLOGIYALARI VA UNING MUAMMOLARI. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 19(1), 61-64.
21. Tojimatov, I., Usmonova, S., Muhammadmusayeva, M., & Xoldarova, S. (2023). DATA MINING MASALALARI VA ULARNING YECHIMLARI. "TRENDS OF MODERN SCIENCE AND PRACTICE", 1(2), 60-63.
22. Nurmamatovich, T. I., & Azizjon o'g, N. A. Z. (2024). The SQL server language and its structure. American Journal of Open University Education, 1(1), 11-15.
23. Tojiddinov, A., Gulsumoy, N., Muntazam, H., & Tojimatov, I. (2023). BIG DATA. Journal of Integrated Education and Research, 2(3), 35-42.
24. Tojimatov, I. N., Asilbek, S., Abdumajid, S., & Mohidil, S. (2023, March). KATTA HAJMDAGI MA'LUMOTLARDA HADOOP ARXITEKTURASI. In INTERNATIONAL SCIENTIFIC AND PRACTICAL CONFERENCE "THE TIME OF SCIENTIFIC PROGRESS" (Vol. 2, No. 4, pp. 78-88).
25. Hakimjonov, O. U., Muhammadjonova, S. I., & Tojimatov, I. N. (2023). MA'LUMOTLARNI INTELEKTUAL TAHLIL QILISHDA DATA MINING QO'LLASH. Scientific progress, 4(3), 132-137.
26. Karimberdiyevich, O. M., Mahamadamin o'g'li, Y. A., & Abdulaziz o'g'li, Y. M. (2023). MASHINALI O'QITISH ALGORITMLARI ASOSIDA BASHORAT QILISH USULLARINI YARATISH. Journal of new century innovations, 22(2), 165-167