

## **Katta hajmdagi ma'lumotlarni (Big Data) qayta ishlash uchun tahliliy tizim yaratish**

*Aminov Bahrom Burxonovich – Samarqand davlat  
pedagogika instituti “Informatika” kafedrasida assistenti,  
Urunova Xilola Najmiddin qizi – Samarqand  
davlat pedagogika instituti “Aniq va amaliy  
fanlar” fakulteti “Matematika va informatika”  
yo‘nalishi 4-kurs talabasi. urunovaxilola2004@gmail.com*

### **Annotatsiya**

Bugungi kunda raqamli texnologiyalar rivojlanishi natijasida turli sohalarda juda hajmdagi ma'lumotlar shakllanmoqda. Ushbu ma'lumotlarni oddiy usullar bilan qayta ishlash tobora qiyinlashib bormoqda. Shu sababli katta hajmdagi ma'lumotlarni yig'ish, saqlash, qayta ishlash va tahlil qilishni ta'minlaydigan zamonaviy tahliliy tizimlarni yaratish dolzarb masalaga aylanmoqda. Mazkur maqolada katta hajmdagi ma'lumotlarni yig'ish, qayta ishlashga mo'ljallangan tahliliy tizimning arxitekturasi va amaliy realizatsiyasi yoritiladi. Shuningdek, tizimning web-asosidagi monitoring interfeysi misolida ma'lumotlarni yig'ish, qayta ishlash va vizual nazorat qilish mexanizmlari ko'rib chiqiladi.

### **Аннотация**

В настоящее время, с развитием цифровых технологий, в различных сферах формируется огромный объем данных. Обработка этих данных традиционными методами становится все более сложной. Поэтому создание современных аналитических систем, обеспечивающих сбор, хранение, обработку и анализ больших объемов данных, становится актуальной задачей. В данной статье рассматривается архитектура и практическая реализация аналитической системы, предназначенной для сбора и обработки больших данных. Также на примере веб-ориентированного интерфейса мониторинга рассматриваются механизмы сбора, обработки и визуального контроля данных.

### **Annotation**

Nowadays, with the development of digital technologies, vast amounts of data are being generated across various fields. Processing this systems that ensure the data using traditional methods is becoming increasingly difficult. Therefore, the creation of modern analytical collection,

storage, processing, and analysis of large datasets has become a pressing issue. This article discusses the architecture and practical implementation of an analytical system designed for collecting and processing big data. Additionally, mechanisms for data collection, processing, and visual monitoring are examined using a web-based monitoring interface as an example.

**Kalit so‘zlar:** Big Data, tahliliy tizim, Data Lake, ETL, monitoring, analitika, web-ilova.

**Kirish.** Axborot texnologiyalari jadal sur‘atlarda rivojlanayotgan hozirgi davrda ma’lumotlar hajmi yil sayin keskin ortib bormoqda. Ijtimoiy tarmoqlar, onlayn savdo tizimlari, bank operatsiyalari, davlat xizmatlari va turli sensor qurilmalar doimiy ravishda katta miqdorda axborot ishlab chiqarmoqda. Ushbu ma’lumotlar hajmining ortishi ularni saqlash va tahlil qilish jarayonini murakkablashtiradi. An’anaviy ma’lumotlar bazasi tizimlari bunday yuklamani har doim ham samarali boshqara olmaydi.

Big Data atamasi hajmi katta, tezkor ravishda yangilanib turadigan va turli formatlarda keladigan ma’lumotlarni anglatadi. Odatda u hajm, tezlik va turli xillik kabi uchta asosiy xususiyat bilan tavsiflanadi. Biroq amaliyot shuni ko‘rsatadiki, ma’lumotlarning ishonchliligi va ulardan olinadigan foyda ham katta ahamiyatga ega. Shuning uchun katta hajmdagi ma’lumotlar bilan ishlashda faqat texnik yechim emas, balki puxta o‘ylangan tizimli yondashuv talab etiladi.

Tahliliy tizimni loyihalash. Katta hajmdagi ma’lumotlar uchun mo‘ljallangan tahliliy tizim oddiy saqlash vositasi emas, balki to‘liq ishlaydigan analitik infratuzilma bo‘lishi kerak. Bunday tizim ma’lumotlarni qabul qilishdan tortib, yakuniy natijani foydalanuvchiga yetkazishgacha bo‘lgan barcha jarayonlarni qamrab oladi. Tizim odatda bir nechta asosiy bosqichlardan iborat bo‘ladi.

Birinchi bosqich – ma’lumotlarni yig‘ish. Bu yerda ma’lumotlar turli manbalardan olinadi va ular strukturaviy, yarim strukturaviy yoki tuzilmasiz ko‘rinishda bo‘lishi mumkin. Masalan, jadval ko‘rinishidagi ma’lumotlar bilan bir qatorda matn, rasm yoki audio fayllar ham qayta ishlanishi mumkin.

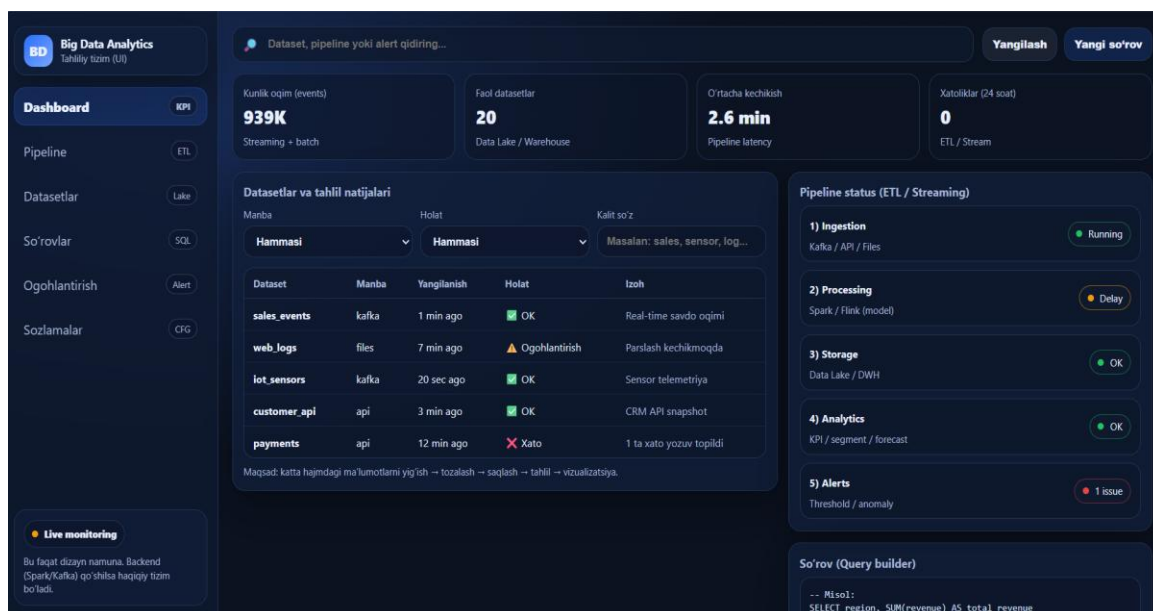
Ma’lumotlarni qayta ishlash jarayoni alohida ahamiyatga ega. Bu bosqichda ma’lumotlar tozalanadi, xatolar tuzatiladi, ortiqcha yoki takroriy qiymatlar olib tashlanadi. Qayta ishlash ikki

xil rejimda amalga oshirilishi mumkin: paket va oqimli. Paket rejimida katta hajmdagi tarixiy ma'lumotlar ustida murakkab hisob-kitoblar bajariladi. Oqimli rejim esa real vaqt ma'lumotlarini tezkor tahlil qilish imkonini beradi. Amaliyotda ushbu ikki rejimni birgalikda qo'llash eng samarali natijani beradi.

Keyingi bosqich – ma'lumotlarni saqlash. Katta hajmdagi ma'lumotlarni markazlashtirilgan bitta serverda saqlash samarali emas, shu sababli taqsimlangan saqlash muhitidan foydalaniladi. An'anaviy Data Warehouse tizimlari qat'iy tuzilma asosida ishlasa, Data Lake ko'proq moslashuvchanlik beradi. So'nggi yillarda ushbu ikki yondashuvni birlashtiruvchi lakehouse modeli keng qo'llanilmoqda. Bu model ma'lumotlarni erkin saqlash bilan birga ularni tartibli boshqarish imkonini ham yaratadi.

Oxirgi bosqich – analitika va vizualizatsiya bo'lib, bunda qayta ishlangan va saqlangan ma'lumotlar asosida turli analitik ko'rsatkichlar hisoblanadi, ma'lumotlar guruhlanadi va muayyan mezonlar asosida tahlil qilinadi. Vizualizatsiya tizimning muhim qatlami bo'lib, u murakkab hisob-kitob natijalarini sodda va tushunarli ko'rinishga keltiradi. Natijada foydalanuvchi tizimning joriy holatini tezkor baholashi, muammo yuzaga kelgan bosqichni aniqlashi va zarur boshqaruv qarorlarini qabul qilishi mumkin bo'ladi.

Amaliy realizatsiya va tizim interfeysi. Mazkur maqola doirasida katta hajmdagi ma'lumotlarni qayta ishlash jarayonini vizual tarzda namoyish etuvchi web asosidagi tahliliy tizim interfeysi ishlab chiqildi. Ushbu tizim Big Data arxitekturasi asosiy komponentlarini modellashtiradi va foydalanuvchiga ma'lumotlar oqimini monitoring qilish imkonini beradi.



**1-rasm. Katta hajmdagi ma'lumotlarni analiz qilish tizimining bosh oynasi.**

Ishlab chiqilgan tizim bir nechta funksional modullardan iborat: Dashboard (asosiy boshqaruv paneli) – tizimning asosiy ko'rsatkichlarini aks ettiradi. Unda kunlik ma'lumotlar oqimi, faol datasetlar, o'rtacha kechikishlar va oxirgi 24 soatdagi xatoliklar sonini kuzatish mumkin; datasetlar va monitoring moduli – ma'lumotlar jadval ko'rinishida bo'ladi. Har bir dataset uchun quyidagi parametrlar ko'rsatiladi: ma'lumot manbasi, oxirgi yangilanish vaqti, ishlash holati (OK, ogohlantirish, xato), qo'shimcha izoh; Pipeline (ETL / Streaming) nazorati – tizim arxitekturasi quyidagi bosqichlarni vizual tarzda aks ettiradi: ingestion – ma'lumotlarni yig'ish (Kafka, API, fayl tizimi), processing – ma'lumotlarni qayta ishlash (Spark/Flink modeli asosida), storage – Data Lake yoki Data Warehouse qatlamida saqlash, analytics – agregatsiya, segmentatsiya va prognozlash, alerts – ogohlantirishlar shakllantirish, ogohlantirishlar tizimi, so'rovlar moduli. Bu monitoring va tizim barqarorligini ta'minlashda muhim rol o'ynaydi.

### Xulosa

Maqolada katta hajmdagi ma'lumotlarni qayta ishlash va tahlil qilish uchun web-asosidagi analitik tizim ishlab chiqilgan. Tizim Dashboard, datasetlar monitoringi, Pipeline nazorati, ogohlantirishlar va so'rovlar modullaridan iborat bo'lib, foydalanuvchiga ma'lumotlar oqimini real vaqt rejimida kuzatish va tahlil qilish imkonini beradi. Vizualizatsiya murakkab natijalarni sodd

va tushunarli ko‘rinishga keltiradi, bu esa tizim samaradorligini oshiradi va boshqaruv qarorlarini tezkor qabul qilishni ta‘minlaydi.

### **Foydalanilgan adabiyotlar**

1. Abdurahmonov O. Katta hajmdagi ma’lumotlarni qayta ishlashda tarqatilgan hisoblash texnologiyalari. Toshkent, 2022.
2. Axmedov S. va Boqijonov J. Katta ma’lumotlar: Big data va analitika. Toshkent (2019).
3. Inmon W. H. Building the Data Warehouse. John Wiley & Sons, 2005.
4. Ismoilov D. Ma’lumotlar bazalari va ularni boshqarish tizimlari. Toshkent (2020).
5. Kleppmann M. Designing Data-Intensive Applications. O’Reilly Media, 2017.