

BIG DATA SOHASIDAGI XALQARO LOYIHALAR**Onarkulov Maksadjon Karimberdiyevich***Farg'ona davlat universiteti amaliy matematika va informatika kafedrası dotsenti*maxmaqsad@gmail.com**Tojimamatov Israiljon Nurmamatovich***Farg'ona davlat universiteti amaliy matematika va informatika kafedrası katta o'qituvchisi*isik80@mail.ru**Yusupov Mirsaid Abdulaziz o'g'li***Farg'ona davlat universiteti amaliy matematika va informatika kafedrası o'qituvchisi*mirsaidbeky@gmail.com**Usmonaliyev Ulug'bek Ismoiljon o'g'li***Farg'ona davlat universiteti*ulugbekusmonaliyevo4@gmail.com

Annotatsiya. Big Data (Katta Ma'lumotlar) - bu hajmi, tezligi va turli shakllari bilan an'anaviy ma'lumotlar bazalarini ishlatish uchun juda katta va murakkab bo'lgan ma'lumotlar to'plamidir. Katta ma'lumotlar texnologiyalari biznes, ilm-fan, tibbiyot, xavfsizlik, va boshqa sohalarda qarorlar qabul qilishni yaxshilash, prognozlash, innovatsiyalar yaratish, va operatsion samaradorlikni oshirishda keng qo'llanilmoqda. Mavzuda, Big Data'nun asosiy xususiyatlari — hajm, tezlik va turli xillik, hamda global miqyosda amalga oshirilayotgan ba'zi katta loyihalar, masalan, Hadoop, Google BigQuery, CERN, IBM Watson va aqlli shaharlar loyihalari ko'rib chiqiladi. Ushbu texnologiyalar yordamida ilmiy kashfiyotlar amalga oshirilmoqda, iqlim o'zgarishlari prognoz qilinmoqda, sog'liqni saqlashni rivojlantirish va shahar infratuzilmasini boshqarish mumkin bo'lmoqda.

Kalit so'zlar:Katta Ma'lumotlar (Big Data), Hadoop, Apache Spark, Google BigQuery, Inson Genomi Loyihasi, CERN, IBM Watson, Aqlli Shaharlar, Sog'liqni Saqlash, Iqlim O'zgarishi, Prognozlash, Innovatsiya, Ma'lumotlarni Tahlil Qilish

Аннотация. Big Data (Большие Данные) — это огромные и сложные наборы данных, которые не могут быть обработаны традиционными базами данных из-за их объема, скорости и разнообразия. Технологии Big Data активно применяются в различных сферах, таких как бизнес, наука, медицина, безопасность, для улучшения принятия решений, прогнозирования, создания инноваций и повышения операционной эффективности. В статье рассматриваются основные характеристики Big Data — объем, скорость и разнообразие, а также некоторые крупные международные проекты, такие

как Hadoop, Google BigQuery, CERN, IBM Watson и умные города. Эти технологии позволяют делать научные открытия, прогнозировать изменения климата, развивать здравоохранение и управлять инфраструктурой городов.

Ключевые слова: Большие Данные (Big Data), Hadoop, Apache Spark, Google BigQuery, Проект Геном Человека, CERN, IBM Watson, Умные Города, Здравоохранение, Изменение Климат, Прогнозирование, Инновации, Анализ Данных

Annotatsiya. Big Data refers to vast and complex datasets that cannot be processed using traditional databases due to their volume, velocity, and variety. Big Data technologies are widely applied across various fields such as business, science, healthcare, security, and urban management, improving decision-making, forecasting, driving innovation, and increasing operational efficiency. This topic discusses the core characteristics of Big Data — volume, velocity, and variety, as well as international projects like Hadoop, Google BigQuery, CERN, IBM Watson, and Smart Cities. These technologies enable scientific discoveries, climate change predictions, advancements in healthcare, and the optimization of city infrastructure.

Key words: Big Data, Hadoop, Apache Spark, Google BigQuery, Human Genome Project, CERN, IBM Watson, Smart Cities, Healthcare, Climate Change, Forecasting, Innovation, Data Analysis

Big Data (Katta Ma'lumotlar) — bu hajmi, tezligi va turli xil shakllari bilan an'anaviy ma'lumotlar bazalarini ishlatish uchun juda katta va murakkab bo'lgan ma'lumotlar to'plamidir. Big Data - bu texnologiyani ulkan sohasi bo'lib, uni deyarli har qanday ehtiyoj uchun turli xil yechimlarga aylantirish va optimallashtirish mumkin

Big Data odatda quyidagi uchta asosiy xususiyat bilan tavsiflanadi:

1. **Volume (Hajm):** Ma'lumotlarning juda katta miqdori. Masalan, kompaniyalar kundalik ish faoliyatida millionlab foydalanuvchilarni va tranzaksiyalarni qayd etadi, shuningdek, turli manbalardan (ijtimoiy tarmoqlar, sensorlar, internet, va hokazo) ma'lumotlar yig'iladi.

2. **Velocity (Tezlik):** Ma'lumotlarning tezligi, ya'ni ular qanday tezlikda yaratilayotganini yoki oqayotganini anglatadi. Masalan, ijtimoiy tarmoqlarda yangi postlar, sharhlar va boshqa kontentlar tezda yangilanadi va bu ma'lumotlar real vaqtda tahlil qilinishi kerak.

3. **Variety (Turli Xillik):** Ma'lumotlarning xilma-xilligi. Big Data turli shakllarda bo'lishi mumkin: strukturalangan (masalan, ma'lumotlar bazalari), strukturalanmagan (matn, rasm, video) va yarim strukturalangan (masalan, log fayllari).

Big Data (Katta Ma'lumotlar)ni ishlatishning asosiy maqsadi — katta hajmdagi, turli xil va tezda o'zgarib turadigan ma'lumotlardan samarali foydalanish, tahlil qilish va ulardan foydali ma'lumotlar olishdir. Katta ma'lumotlar tahlili orqali turli sohalarda

samaradorlikni oshirish, qarorlar qabul qilish, yangi imkoniyatlar yaratish va muammolarni hal qilish mumkin.

Big Data'dan foydalanishning ba'zi asosiy sabablarini quyidagilar

1. Tahlil qilish va qarorlar qabul qilishni yaxshilash

Misol: Banklar, mijozlarining moliyaviy odatlarini kuzatib, ularga maxsus moliyaviy xizmatlar taklif qiladi.

2. Prognozlash va kelajakni oldindan bilish

Misol: Sog'liqni saqlash sohasida, katta ma'lumotlarni tahlil qilish orqali, kasalliklarning oldini olish yoki ularning rivojlanishini prognoz qilish mumkin.

3. Ijtimoiy tarmoqlarni tahlil qilish

Misol: Markalar, ijtimoiy tarmoqlardagi postlar va izohlarga asoslanib, mijozlarning fikrlarini o'rganadi va marketing strategiyalarini shakllantiradi.

4. Tez va samarali qarorlar qabul qilish

Misol: Moliyaviy bozorlarda ham Big Data real vaqt rejimida tahlil qilinadi, bu orqali investitsiya qarorlarini tezroq va aniqroq qabul qilish mumkin.

5. Mijozlarni yaxshiroq tushunish va moslashtirilgan xizmatlar taklif qilish

Misol: Raqobatchilarni tahlil qilib, yangi mahsulotlarni ishlab chiqish yoki mavjud mahsulotni yaxshilash uchun yangi g'oyalar olish mumkin.

6. Operatsion samaradorlikni oshirish

Misol: Ishlab chiqarish kompaniyalari mashinalar va uskunalarning ishlash holatini tahlil qilib, nosozliklarni oldindan aniqlash va ularni tuzatish uchun oldindan choralar ko'rishlari mumkin (bu jarayon "Prediktiv Texnika" deb ataladi).

7. Innovatsiyalarni yaratish

Misol: Avtonom transport vositalari uchun Big Data orqali yo'l harakati ma'lumotlarini tahlil qilish orqali xavfsiz va samarali tizimlar ishlab chiqilmoqda.

8. Xavfsizlikni ta'minlash

Misol: Kiberhujumlarni aniqlash va ularga qarshi kurashish uchun Big Data tizimlari foydalanuvchilarning faoliyatini real vaqtda kuzatadi va shubhali harakatlarni aniqlaydi.

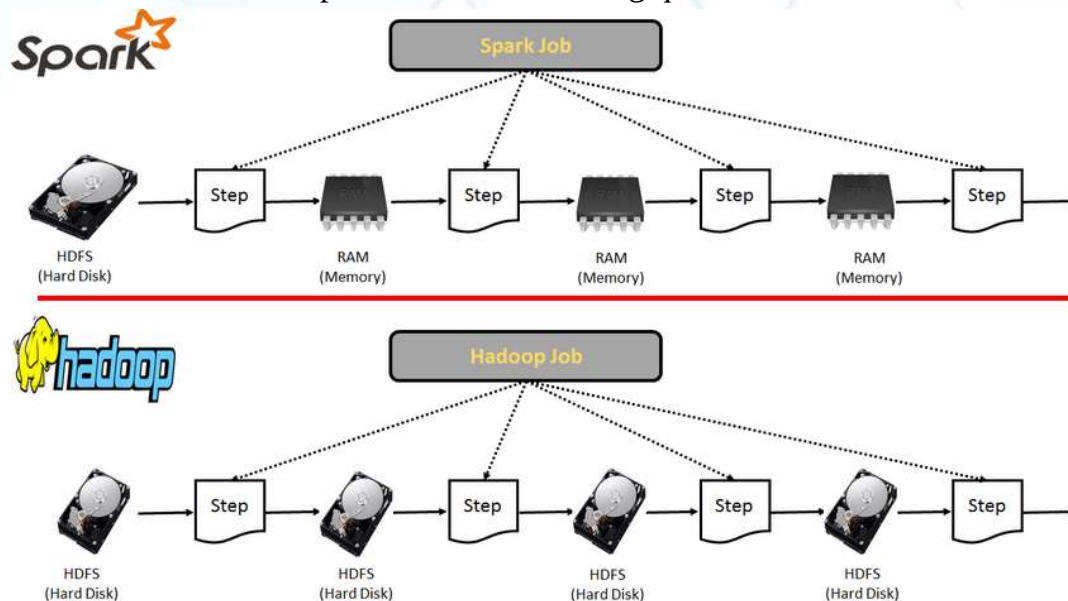
Sohadagi xalqaro loyihalar

1. Hadoop va Apache Spark (Apache Software Foundation)

Apache Hadoop — bu katta hajmdagi ma'lumotlarni saqlash va ishlov berish uchun mo'ljallangan ochiq manbali ramkadir. Bu platforma Big Data sohasida eng mashhur texnologiyalardan biri bo'lib, katta hajmdagi ma'lumotlarni tahlil qilish uchun parallel hisoblashni qo'llaydi.

- **Apache Spark** — bu Hadoop platformasining tezlashtirilgan versiyasi bo'lib, real vaqt rejimida ma'lumotlarni tahlil qilish imkonini beradi. Spark ma'lumotlarni tezda qayta ishlash va to'g'ridan-to'g'ri memoriya orqali tahlil qilishni osonlashtiradi.

Loyihaning maqsadi: Katta hajmdagi ma'lumotlarni saqlash, tahlil qilish va uzluksiz yangilanishi kerak bo'lgan real vaqtda ishlov berish. Spark va Hadoop Big Data tahlili va biznesni optimallashtirishda keng qo'llaniladi.



2. Google BigQuery (Google Cloud Platform)

Data processing in the cloud



Google BigQuery — bu Google Cloud Platformda ishlaydigan katta hajmdagi ma'lumotlarni saqlash va tahlil qilish uchun mo'ljallangan bulutli xizmatdir. BigQuery foydalanuvchilarga terabyte yoki petabyte o'lchamdagi ma'lumotlar ustida tezkor va samarali analizlar o'tkazishga imkon beradi.

• **Loyihaning maqsadi:** Google foydalanuvchilariga katta hajmdagi ma'lumotlarni tezkor tahlil qilish va ulardan foydali ma'lumotlarni olish imkoniyatini taqdim etadi. Bu platforma real vaqt tahlilini amalga oshiradi va biznes tahlilchilari uchun qulay vositadir.

3. Human Genome Project (HGP)

Human Genome Project (HGP) — bu inson genomini to'liq xaritalashni maqsad qilgan xalqaro loyiha bo'lib, Big Data texnologiyalari yordamida genomik ma'lumotlar to'plandi va tahlil qilindi.

- **Loyihaning maqsadi:** Inson genetik kodini xaritalash va uning tahlilini amalga oshirish orqali kasalliklarning oldini olish va yangi davolash usullarini ishlab chiqish. HGP genomik ma'lumotlarni saqlash va tahlil qilish uchun katta hajmdagi ma'lumotlarni yig'ishga asoslangan.

Big Data'dan foydalanish: Genomik ma'lumotlar juda katta hajmga ega bo'lib, ulardan foydalangan holda tibbiyot va biotexnologiyalar sohasida yangi ilmiy kashfiyotlar amalga oshirildi.

4. CERN (The European Organization for Nuclear Research)

CERN — bu Yevropa yadroviy tadqiqotlar tashkiloti bo'lib, u katta hajmdagi ilmiy ma'lumotlarni saqlash va tahlil qilish uchun Big Data texnologiyalaridan foydalanadi. CERNning **Large Hadron Collider (LHC)** loyihasi, asosan, atomlararo to'qnashuvlar natijasida yuzaga keladigan ma'lumotlarni tahlil qilishga qaratilgan.

- **Loyihaning maqsadi:** Katta hajmdagi ilmiy ma'lumotlarni yig'ish, saqlash va tahlil qilish, masalan, kvant fizikasi va yadro tadqiqotlari uchun ma'lumotlar. CERNda Big Data texnologiyalari orqali subatomik zarralar to'qnashuvlaridan olingan ma'lumotlar tahlil qilinadi.

Big Data'dan foydalanish: CERNning LHC loyihasida yuzlab petabayt ma'lumotlar yig'iladi va tahlil qilinadi. Ushbu ma'lumotlar orqali ilmiy kashfiyotlar amalga oshiriladi, masalan, **Higgs bozonini** kashf etish.

4. IBM Watson



IBM Watson — bu mashinani o'rganish va tabiiy tilni qayta ishlash texnologiyalaridan foydalangan holda Big Data ma'lumotlarini tahlil qilish uchun yaratilgan tizimdir. Watsonning ko'plab qo'llaniladigan sohalari bor, jumladan sog'liqni saqlash, biznes tahlili va mijozlarga xizmat ko'rsatish.

- **Loyihaning maqsadi:** IBM Watsonni sog'liqni saqlashda, jumladan, saratonni davolashda, ilmiy tadqiqotlarni tahlil qilishda va biznes analitikalarida qo'llash.

Watson katta ma'lumotlarni tahlil qilish orqali mijozlarga tavsiyalar berish, kasalliklarni aniqlash va ilg'or tadqiqotlar o'tkazishda yordam beradi.

Big Data'dan foydalanish: IBM Watson sog'liqni saqlashda yirik tibbiy ma'lumotlar to'plamini tahlil qiladi, masalan, bemorlar tarixini va ilmiy maqolalarni o'rganish, saraton kasalligini aniqlash uchun klinik tahlillarni taqdim etadi.

6. NASA's Earth Science Data (NASA)

NASA Earth Science Data — bu NASA tomonidan boshqariladigan, Yer yuzasidagi tabiiy hodisalarni kuzatish uchun yig'ilgan katta hajmdagi ma'lumotlar to'plami. NASA Yer ilmiy ma'lumotlar markazi orqali iqlim o'zgarishlari, atmosfera tahlili va boshqa ekologik jarayonlar haqida ma'lumotlar yig'adi.

- **Loyihaning maqsadi:** Yer yuzasidagi tabiiy hodisalar va iqlim o'zgarishlarini kuzatib borish. NASA Big Data texnologiyalaridan iqlimni modellashtirish, o'rmonlarning holatini kuzatish, atmosfera holatini tahlil qilish va boshqa ekologik o'zgarishlarni o'rganish uchun foydalanadi.

Big Data'dan foydalanish: NASA Yerning atmosfera tizimi va tabiiy hodisalarini tahlil qilish uchun katta hajmdagi ma'lumotlar to'plamini yig'adi va bu ma'lumotlar yordamida iqlim o'zgarishlarini prognoz qilish mumkin.

7. Smart Cities (Aqlli shaharlar)

Aqlli shaharlar (Smart Cities) — bu shaharlar uchun Big Data texnologiyalaridan foydalanadigan global tashabbuslar bo'lib, transportni, energetikani, xavfsizlikni va boshqa shahar infratuzilmasini boshqarish uchun katta hajmdagi ma'lumotlarni yig'ish va tahlil qilishga qaratilgan.

- **Loyihaning maqsadi:** Shaharlar aholisining ehtiyojlarini o'rganish, transport tizimlarini optimallashtirish, energiya sarfini kamaytirish va xavfsizlikni yaxshilash uchun katta ma'lumotlar to'plamini ishlatish.

Big Data'dan foydalanish: Aqlli shaharlar loyihalarida real vaqt rejimida ma'lumotlar yig'iladi, masalan, sensorlar yordamida harakatlar va energiya sarfi kuzatiladi. Bu ma'lumotlar shaharni samarali boshqarish va aholiga yuqori sifatli xizmatlar taqdim etish imkonini beradi.

Xulosa:

Big Data (Katta Ma'lumotlar) — bu hajmi, tezligi va turli shakllari bilan an'anaviy ma'lumotlar bazalarini ishlatish uchun juda katta va murakkab bo'lgan ma'lumotlar to'plamidir. Big Data texnologiyalari turli sohalarda samarali qarorlar qabul qilish, tahlil qilish va innovatsiyalar yaratish uchun qo'llaniladi. Katta ma'lumotlar odatda uchta asosiy xususiyatga ega: Hajm (millionlab foydalanuvchilar va tranzaksiyalar), Tezlik (real vaqtda yangilanadigan ma'lumotlar), va Turli xillik (strukturali va strukturali bo'lmagan ma'lumotlar). Big Data tahlili orqali qarorlar qabul qilishni yaxshilash, kelajakni prognoz qilish, ijtimoiy tarmoqlardagi xatti-harakatlarni tahlil qilish, mijozlarga moslashtirilgan xizmatlar taklif qilish, innovatsiyalar yaratish va xavfsizlikni ta'minlash mumkin. Xalqaro miqyosda bir qancha katta loyihalar mavjud,

masalan, Hadoop, Google BigQuery, Human Genome Project, CERN, IBM Watson, NASA Earth Science Data va aqlli shaharlar loyihalari. Bu loyihalar katta hajmdagi ma'lumotlarni tahlil qilib, ilmiy kashfiyotlar, iqlim o'zgarishlarini prognoz qilish, sog'liqni saqlashni rivojlantirish va shaharlarni boshqarishda samaradorlikni oshiradi. Big Data texnologiyalari nafaqat ilm-fan va tadqiqotlarda, balki biznes va kundalik hayotda ham keng qo'llanilmoqda.

FOYDALANILGAN ADABIYOTLAR:

1. Mamasidiqova, I., Husanova, O., Madaminova, A., & Tojimamatov, I. (2023). DATA MINING TEXNALOGIYALARI METODLARI VA BOSQICHLARI HAMDA DATA SCIENCE JARAYONLAR. Центральноеазиатский журнал образования и инноваций, 2(3 Part 2), 18-21.
2. Tojimamatov, I. N., Mamalatipov, O. M., & Karimova, N. A. (2022). SUN'Y NEYRON TARMOQLARINI O'QITISH USULLARI. Oriental renaissance: Innovative, educational, natural and social sciences, 2(12), 191-203.
3. Nurmamatovich, T. 1. (2021). RAQAMLI IQTISODIYOTNING GLOBALLASHUV JARAYONIDA IQTISOD TARMOQLARIDA QO'LLANILISHINING ASOSIY YO'NALISHLARI. H34 Наука U инновации Международной, 291. 6 XXI веке: Материалы
4. Tsuchievich, B. M., & Nurmamatovich, T. I. (2021). JAMIYATDA RAQAMLI QIZIQAT. H34 21-asrda fan va innovatsiyalar: Proceedings International, 189.
5. Kizi, A. Z. I., & Nurmamatovich, T. I. (2021). ZAMONAVIY DASTURLASH FANINI O'QITISHDA PYTHON DASTURLASH VOSITALARI YORDAMIDA

Web-sahifalar

1. MongoDB: <https://www.mongodb.com>
2. Cassandra: <https://cassandra.apache.org>
3. Neo4j: <https://neo4j.com>
4. Couchbase: <https://www.couchbase.com>