

## KATTA MA'LUMOTLARNI TAHLIL QILISH VA SAQLASH TIZIMLARI TAHLILI

*Alisher Madaminov*

*G'ayrat Matlatipov*

*Urganch Davlat Universiteti, Kompyuter ilmlari kafedrası*

[alisherqadambayevich@gmail.com](mailto:alisherqadambayevich@gmail.com)

**Annotatsiya:** Ushbu ilmiy ish katta ma'lumotlar (Big data)ni qayta ishlash tizimlarining xususiyatlarini va ta'sirini qisqacha o'rganadi. Kengaytirish, xatolarga chidamlilik va parallel ishlov berishni ta'kidlab, moliya, sog'liqni saqlash va chakana savdo kabi sohalarda ularning hal qiluvchi rolini o'rganadi. Inshoda ushbu tizimlar tashkilotlarga ma'lumotlarga asoslangan davrimizda ongli qarorlar qabul qilish uchun amaliy tushunchalarni olish imkoniyatini qanday berishini ta'kidlaydi.

**Kalit so'zlar:** Katta ma'lumotlar, big data tizimlari, big datani saqlash.

**Аннотация:** В данной научной работе кратко рассмотрены характеристики и эффекты систем обработки больших данных. Подчеркивается масштабируемость, отказоустойчивость и параллельная обработка, а также исследуется их решающая роль в таких отраслях, как финансы, здравоохранение и розничная торговля. В эссе подчеркивается, как эти системы позволяют организациям получать полезную информацию для принятия обоснованных решений в наш век данных.

**Ключевые слова:** Большие данные, системы больших данных, хранилище больших данных.

**Annotation:** This scientific work briefly examines the characteristics and effects of Big Data processing systems. Emphasizes scalability, fault tolerance, and parallel processing, and explores their critical role in industries such as finance, healthcare, and retail. The essay highlights how these systems enable organizations to gain actionable insights to make informed decisions in our data-driven age.

**Key words:** Big data, big data systems, big data storage.

**Kirish.**

## International Conference on Education and Innovation

Bugungi raqamli texnologiyalar asrida turli manbalardan olingan ma'lumotlarning eksponentsial o'sishi korxonalar va tashkilotlar uchun muhim tarkibiy qismlar sifatida katta ma'lumotlarni tahlil qilish va saqlash tizimlarining paydo bo'lishiga olib keldi. Ushbu tizimlar katta hajmdagi tuzilgan va tuzilmagan ma'lumotlarni qayta ishlash, boshqarish va tushunchalar olishda muhim ro'l o'ynaydi. Axborot ko'pligi davrida massiv ma'lumotlar to'plamini yaratish odatiy holga aylandi, bu esa samarali ma'lumotlarni qayta ishlash va tahlil qilish uchun ilg'or tizimlarni talab qiladi. Big datani qayta ishlash tizimlari keng va murakkab ma'lumotlar to'plamidan mazmunli tushunchalarni qayta ishlash va olishda muhim rol o'ynaydi. Ushbu insho Big datani qayta ishlash tizimlarini tahlil qiladi va ularning asosiy xususiyatlarini, me'moriy tarkibiy qismlarini va turli sohalarga ta'sirini o'rganadi.

### **Katta ma'lumotlarni tahlil qilish va saqlash tizimlarining ahamiyati:**

Katta ma'lumotlarni tahlil qilish va saqlash tizimlari turli sohalardagi tashkilotlar uchun asosli qarorlar qabul qilish, operatsion samaradorlikni oshirish va raqobatdosh ustunliklarga ega bo'lish uchun zarurdir. Ushbu tizimlar korxonalariga real vaqtda yoki real vaqt rejimida ma'lumotlarni olish, saqlash, qayta ishlash va tahlil qilish imkonini beradi, bu esa o'z vaqtida tushuncha va harakatlarni osonlashtiradi. Bundan tashqari, ular tashkilotlarga katta ma'lumotlar to'plamidagi faktlar, tendentsiyalar va korrelyatsiyalarni aniqlash uchun mashinani o'rganish, bashoratli modellashtirish va ma'lumotlarni qazib olish kabi ilg'or analitik usullarning kuchidan foydalanish imkoniyatini beradi.

### **Big datani qayta ishlash tizimlarining asosiy xususiyatlari:**

**Masshtablilik** – Big datani qayta ishlash tizimlari keng miqyosda ishlab chiqilgan bo'lib, ularga ish yukini bir nechta tugunlar yoki klasterlar bo'ylab taqsimlash orqali o'sib borayotgan ma'lumotlar to'plamlarini boshqarishga imkon beradi. Masshtablilik ish unumdorligini pasaytirmasdan ortib borayotgan ma'lumotlar hajmini sig'dira olishini ta'minlaydi.

**Xatolarga chidamlilik** - Big Data operatsiyalari ko'lamini hisobga olgan holda, apparatdagi nosozliklar yoki tarmoq muammolari ehtimoli yuqoriroq. Xatolarga

chidamlilik mexanizmlari uzluksiz ishlov berishni ta'minlash uchun juda muhimdir. Ma'lumotlarning takrorlanishi va tugunlar bo'ylab hisob-kitoblar nosozliklar yuz berganda tizim yaxlitligini saqlashga yordam beradi.

**Parallel ishlov berish** - Big Data tizimlari vazifalarni bir vaqtning o'zida bajarilishi mumkin bo'lgan kichik vazifalarga ajratish uchun parallel ishlov berishdan foydalanadi, bu esa ishlov berish vaqtini sezilarli darajada kamaytiradi. Parallel ishlov berishga Apache Hadoop va Apache Spark kabi taqsimlangan hisoblash tizimlari orqali erishiladi.

#### **Arxitektura komponentlari:**

**Ma'lumotlarni yig'ish** - Katta ma'lumotlar tizimlari sensorlar, ijtimoiy media, korporativ tizimlar va narsalar interneti (IoT) kabi turli manbalardan ma'lumotlarni to'playdi. Bu jarayon ma'lumotlar bazalaridan tuzilgan ma'lumotlarni, XML yoki JSON fayllaridan yarim tuzilgan ma'lumotlarni va matnli hujjatlar, tasvirlar va videolardan tuzilmagan ma'lumotlarni olishni o'z ichiga oladi.

**Ma'lumotlarni saqlash** - Katta ma'lumotlarni saqlash tizimlari bir nechta tugunlar yoki serverlarda katta hajmdagi ma'lumotlarni saqlash uchun Hadoop Distributed File System (HDFS) kabi taqsimlangan fayl tizimlari klasterdagi tugunlar bo'ylab katta hajmdagi ma'lumotlarni saqlaydi va boshqaradi. Apache Cassandra va MongoDB kabi NoSQL ma'lumotlar bazalari turli xil ma'lumotlar turlari uchun moslashuvchan va kengaytiriladigan saqlash yechimlarini taqdim etadi.

**Qayta ishlash ramkalari** - Apache Spark, Apache Flink va Hadoop MapReduce - bu Big Data to'plamlarida taqsimlangan hisoblashni ta'minlaydigan mashhur qayta ishlash ramkalari. Bu kabi katta ma'lumotlarni qayta ishlash tizimlari tovar apparatlari klasterlari bo'ylab keng ko'lamli ma'lumotlar to'plamlarini taqsimlangan hisoblash imkonini beradi. Ushbu ramkalar turli xil ma'lumotlarni tahlil qilish vazifalarini samarali bajarish uchun ommaviy ishlov berish, oqimni qayta ishlash va interaktiv so'rovlarni qo'llab-quvvatlaydi va ishlab chiquvchilar uchun yuqori darajadagi abstraktsiyalarni taqdim etadi.

## International Conference on Education and Innovation

**Resurslarni boshqarish** - Apache YARN va Kubernetes kabi resurs menejerlari resurslarni taqsimlangan muhitda samarali taqsimlaydi va boshqaradi, natijada ishlashni optimallashtiradi. Dinamik resurslarni taqsimlash hisoblash resurslarining ish yuki talablari asosida taqsimlanishini ta'minlaydi.

**Biznes va moliya** - Big Datani qayta ishlash tizimlari moliyaviy operatsiyalar, firibgarlikni aniqlash va risklarni boshqarish uchun real vaqt tahlilini osonlashtiradi.

**Chakana savdo va elektron tijorat** - Big Data tizimlari shaxsiylashtirilgan tavsiyalarni yaxshilash va ta'minot zanjiri boshqaruvini optimallashtirish uchun mijozlarning xatti-harakatlari, afzalliklari va tendentsiyalarini tahlil qiladi. Inventarizatsiyani boshqarish va talabni prognozlash keng ko'lamli tranzaksiya ma'lumotlarini qayta ishlashdan foyda keltiradi.

**Ma'lumotlarni tahlil qilish** - Katta ma'lumotlarni tahlil qilish platformalari tadqiqot ma'lumotlarini tahlil qilish, statistik modellashtirish va mashinalarni o'rganish uchun vositalar va kutubxonalarini taqdim etadi. Ushbu platformalar ma'lumotlardan amaliy tushunchalar va bashoratlarni olish uchun ma'lumotlarni vizualizatsiya qilish, xususiyat muhandisligi, modellarni o'qitish va baholashni osonlashtiradi. Ularning afzalliklariga qaramay, katta ma'lumotlarni tahlil qilish va saqlash tizimlari bir qator qiyinchiliklarga duch keladi, jumladan:

**Masshtablilik:** Keng miqyosli taqsimlangan tizimlarni boshqarish va masshtablash unumdorlik va ishonchlilikni ta'minlash uchun murakkab infratuzilma va resurslarni boshqarish usullarini talab qiladi.

**Ma'lumotlar integratsiyasi:** Turli formatlar, sxemalar va sifatlar bilan turli xil manbalardan olingan ma'lumotlarni integratsiyalash ma'lumotlarni tozalash, o'zgartirish va o'zaro ishlashda qiyinchiliklarni keltirib chiqaradi.

**Ma'lumotlar xavfsizligi:** Maxfiy ma'lumotlarni ruxsatsiz kirish, ma'lumotlar buzilishi va kiber-hujumlardan himoya qilish ma'lumotlar maxfiyligini saqlash va tartibga solish talablariga muvofiqligini ta'minlash uchun juda muhimdir.



**Ma'lumotlarni boshqarish:** Ma'lumotlarni boshqarish, sifatni ta'minlash va muvofiqlik uchun siyosat, jarayonlar va nazoratni o'rnatish ma'lumotlar aktivlarining yaxlitligi, izchilligi va ishonchliligini ta'minlash uchun muhim ahamiyatga ega.

**Xulosa:** Xulosa qilib aytganda katta ma'lumotlarni tahlil qilish va saqlash tizimlari tashkilotlarga strategik qarorlar qabul qilish, innovatsiyalar va raqobatdosh ustunlik uchun ma'lumotlar qiymatidan foydalanish imkonini berishda hal qiluvchi ro'l o'ynaydi. Asosiy muammolarni hal qilish va rivojlanayotgan tendentsiyalarni qabul qilish orqali korxonalariga yangi tushunchalarni ochish, operatsiyalarni optimallashtirish va raqamli iqtisodiyotda o'sishni rag'batlantirish uchun ushbu tizimlardan foydalanishlari mumkin. Katta ma'lumotlarni qayta ishlash tizimlari keng va murakkab ma'lumotlar to'plamiga xos bo'lgan qiymatni ochish uchun ajralmas hisoblanadi. Ularning kengayishi, nosozliklarga chidamliligi va parallel ishlov berish imkoniyatlari turli sohalardagi tashkilotlarga amaliy tushunchalarni olish, jarayonlarni optimallashtirish va ma'lumotlarga asoslangan qarorlar qabul qilish imkonini beradi. Texnologiya rivojlanishda davom etar ekan, Katta ma'lumotlarni qayta ishlash tizimlarini tahlil qilish ushbu tizimlar ma'lumotlarga asoslangan dunyoning tobora ortib borayotgan talablariga qanday moslashishini tushunish uchun muhim bo'lib qoladi.

#### FOYDALANILGAN ADABIYOTLAR RO'YXATI :

1. G. Matlatipov and Z. Vetulani, "Representation of uzbek morphology in prolog," in *Aspects of Natural Language Processing*, Springer, 2009, pp. 83–110.
2. M. Sharipov and O. Yuldashov, "UzbekStemmer: Development of a Rule-Based Stemming Algorithm for Uzbek Language," *arXiv preprint arXiv:2210.16011*, 2022.
3. I. I. Baltaeva, I. D. Rakhimov, and M. M. Khasanov, "Exact traveling wave solutions of the loaded modified Korteweg-de Vries equation," *Известия Иркутского государственного университета. Серия Математика*, vol. 41, no. 0, pp. 85–95, 2022.

## International Conference on Education and Innovation

4. Big data fundamentals: Concepts, drivers & techniques” by Thomas Erl, Wajid Khattak, Paul Buhler. Pearson, 2016. -240p.

5.B. Kutlimuratova, E. Kuriyozov, and M. Tillaeva, “TEACHING ENGLISH AS A FOREIGN LANGUAGE FOR PRIMARY SCHOOL CHILDREN: LITERATURE REVEIW,” in *FOREIGN LANGUAGE TEACHING AND APPLIED LINGUISTICS*, 2022, pp. 161–171.

6. G. U. Urazboev, A. A. Reyimberganov, and I. D. Rakhimov, “The soliton solutions for the nonlinear Schrödinger equation with self-consistent sources,“,” *ACTUAL PROBLEMS OF STOCHASTIC ANALYSIS*, pp. 242–244, 2021.

