Дәріс 5

Үлкен деректерді сақтау технологиясына шолу

Үлкен деректер саласы ауқым, алуантүрлілік, жылдамдық, деректердің дұрыстығы, жинақталған ақпараттардың құндылығы сияқты белгілермен сипатталады.

Көп жағдайда үлкен деректермен жасалатын жұмысқа өңделмеген деректерді жинастырудан қолдануға жарамды ақпарат алуға дейінгіқалыпты жұмыс процесі кіреді.

Үлкен деректермен жұмыс жасаудағы негізгі мақсат – бұл олардың негізінде іс жүзінде қолдану үшін құнды аналитикалық тұжырымдар алу.

Негізінде үлкен деректер барлық мүдделі тараптар үшін қолжетімді бола бастауы тиіс, бұл олардың өз бетінше қолдануға есептелген теңгерілетін визуализация және бизнес-сараптама құралдарының көмегімен деректер пакетін жеңіл әрі жылдам зерттеп-білулеріне мүмкіндік береді.Сараптама түріне байланысты, ақырғы пайдаланушыларға статикалық «болжамдар» (болжамды сараптама жағдайында) түріндегі дайын нәтижелер немесе ұсынылатын іс-әрекеттер (белгіленген сараптама жағдайында) берілуі мүмкін.

Үлкен деректер бизнестің көптеген салаларында кеңінен қолданыс тапқан. Оларды денсаулық сақтау, телекоммуникация, сауда, логистика салаларында, қаржы компанияларында, сонымен қатар мемлекеттік басқаруда пайдаланады.

Үлкен деректер технологиясы негізінен клиенттік ортаны сараптау үшін қолданылады. Мысалы, HSBC Үлкен деректер технологиясын пластикалық карталармен болатын алаяқтық операцияларға қарсы қолданады. Big Data көмегімен компания қауіпсіздік қызметінің тиімділігін 3 есе, ал алаяқтық жағдайларды тануды 10 есеге арттырды. Аталмыш технологияларды енгізуден алынған экономикалық тиімділік 10 млн АҚШ долларын құрады.

Сарапшылардың айтуынша, Үлкен ді танымал әрі таралған міндеттермен қатар, түрлі аурулармен күресте және эпидемияның өсуін бақылап отыруда пайдалануға болады. Мысалы, Германияда бүгіннің өзінде Үлкен деректердің арқасында онкологиялық аурулар немесе оларға бейімділік пациенттер мен донорлардың қан талдамасына қарап анықталады.

Ақпараттардың үлкен көлемін сақтауға мүмкіндік беретін заманауи технологиялар бизнестің жаңа формасын қалыптастырады. Мысалы, International Data Corporation зерттеулерінің нәтижесі бойынша жыл сайын сақталатын ақпараттар көлемі 40%-ға өсуде.Мұндай жағдайларда ақпарат жеке дара құндылық болудан қалады, алдыңғы қатарға оны өңдеу және пайдалану тәсілдері шығады.

Big Data технологияларды пайдалану нәтижесінде компаниялар маңызды ақпаратты бірнеше секундта алуға қол жеткізеді. Бұл өз кезегінде экономикалық шешімдердің тиімділігін арттыруға, клиенттердің іс-әрекетіндегі өзгерістерге жылдамырақ жауап қайтаруға, нақты уақыт жағдайында ең ерте сатыларда нарықтық үдерістерді анықтауға мүмкіндік береді.

Big Data дамуында жеке және заңды тұлғалар туралы деректерің үлкен көлеміне ие мемлекет те үлкен рөл ойнайды. Мысалы, Еуропала ақпараттың үлкен көлемі ашық, бұл зерттеушілерге және кәсіпкерлерге жақсы мүмкіндік.

Мысалы, кейбір еуропалық мемлекеттерде мемлекеттік органдар жалға берілетін тұрғын үйлер туралы деректерге ие. Олардың негізінде нақты уақыт режимінде бос пәтерлер бар екендігін көрсететін интернет-қосымшалар пайда болды. Сонымен қатар мұнда полиция бөлімшелерінің қандай да бір аудандағы қылмыстық деңгейді көрсететін деректері, басқа да қызметтердің қоршаған инфрақұрылым туралы деректері кіреді. Нәтижесінде адам өзіне ең ыңғайлысын алуға қол жеткізеді.

Қазіргі таңда Қазақстандағы мемлекеттік органдар BigData и OpenData (ашық деректер) концептілерін ендіруге күш салуда.

Мысалы, Қазақстан Республикасының Ақпарат және коммуникациялар министрлігі ақпараттың үлкен көлемін сақтау және өңдеудегі жаңа технологияны енгізуді жоспарлауда. Сондай-ақ ведомство BigData жаңа технологиясын енгізу және Қазақстанды

АКТ дүниежүзілік салада алдыңғы қатарға шығару бойынша алдарына өте үлкен міндет қойып отырғандықтарын айтты.

Үлкен деректер технологиясын денсаулық сақтау, банктер, бөлшек сауда, байланыс операторлары және т.б. көптеген салаларда, яғни көбісі стандартталған немесе ұлттық, сондай-ақ халықаралық деңгейлерде белсенді стандартталудағы салаларда пайдалану үлкен деректерді енгізу мәселесін туындатады.

Қазіргі таңда бірнеше негізгі стандарттау институты, атап айтқанда: Стандарттау жөніндегі халықаралық ұйым және Халықаралық электротехникалық комиссия (ISO/IEC),Халықаралық электр байланысыодағы(ITU), Британия стандарттар институты (BSI), АҚШ Ұлттық стандарттар және технологиялар институты(NIST) үлкен деректерге арналған стандартты әзірлеу бойынша жұмыстарға тартылған.

Стандарттау жөніндегі халықаралық ұйым және Халықаралық электротехникалық комиссия(ИСО/ХЭК) келесідей технологияларды: үлкен деректер (ISO/IEC JTC1/WG 9 «Үлкен деректер»), заттар интернеті (ISO/IEC

JTC1/WG 10 «Заттар интернеті») және ақылды қалаларды (ISO/IEC JTC1/WG 11 «Ақылды қалалар») стандарттауға бағытталған 3 жұмыс тобын құрды.

ИСО стандартына сәйкес, үлкен деректер жөніндегі Жұмыс тобы деректерді стандарттау үлкен бағдарламасының басты тақырыбын айқындаушылар ретінде және стандарттау саласындағы олқылықтарды анықтау бойынша қызмет атқаратын болмақ. Ол негізгі стандарттарды, соның ішінде эталонды сәулетті жасап шығарады.

Бүгінгі күні ISO/IEC JTC1/WG 9 «Үлкен деректер» – стандарттау жөніндегі халықаралық жұмыс тобы халықаралық стандарттардың келесідей жоабаларын әзірлеуде: үлкен деректердің эталонды сәулетіне арналған стандарттар жиынтығы (стандарттар сериясы ISO/IEC 20547) және терминдер мен анықтамаларға арналған стандарттар (ISO/IEC 20546). Аталмыш жобалар жобаны дайындау сатысында (саты коды – 30) тұр.

ITU-да үлкен деректерге қатысты белсенділіктің бірнеше саласын атап көрсетуге болады. ITU құжаттарында белсенділіктің келесі салалары аталады:

– сенімділігі өте жоғары, икемді және жоғары өткізгіштік қабілетілігі бар, төменгі бөгелісті, масштабталатын желілік инфрақұрылым.

– деректер жиынтығын біріктіру және жою.

2015 жылдың соңында ХЭО мүшелері үлкнен деректерге арналған халықаралық стандарт туралы келісімге келді. ХЭО-Т V.3600 ұсынған жаңа стандарт «Үлкен деректер – бұлттық есептеулер негізіндегі талаптар мен мүмкіндіктер» деп аталады.

Стандарт бұлттық есептеу жүйелерінің Big Data қызметтерін ұсынуда қалай қолданылуы мүмкін екендіктерін сипаттайды. Ең бастысы, ол үлкен деректер негізіндегі бұлттық есептеулерге қойылатын талаптарды (деректерді жинауға, деректерді алдын ала өңдеу және деректерді сақтауға қойылатын талаптар, сараптау, визуализациялау және басқару талаптары және деректер қауіпсіздігі мен қорғауға қойылатын талаптар) сипаттайды.

АҚШ Ұлттық стандарттар және технологиялар институты (NIST) үлкен деректер бойынша көп жұмыс жасалған NIST Big Data Interoperability Framework V1.0 – стандарттар легін ұсынады. Оған төмендегі құжаттар кіреді:

- NISTАрнайыбасылым 1500-1, 1-том, Анықтамалар;
- NISTАрнайыбасылым 1500-2, 2-том, Үлкен деректер таксономиясы;

- NIST Арнайыбасылым 1500-3, 3-том, Пайдалану мысалдары және жалпы талаптар;

- NIST Арнайыбасылым 1500-4, 4-том, Қауіпсіздікжәнеқұпиялылық;

- NISTАрнайыбасылым 1500-5, 5-том, Сәулет;

- NISTАрнайыбасылым 1500-6, 6-том, Эталонды сәулет;

- NISTАрнайыбасылым 1500-7, 7-том, Жол картасының стандарттары.

Бұдан бөлек, Британия стандартар институтына (BSI)сәйкес, келесіде мәселелер үлкен деректерде стандарттауға жатады:

- Метадеректер стандарты.

- Деректермен жұмыс жасау талаптарының стандарттары.

- Деректерді жинау стандарттары.

- BigData жобаларына арналған түсіндіру стандарттары.

- BigData-ға арналған «Қалай жасау керек?» нұсқаулығы.

Бұл жерде ақылды қалалардағы деректерді анықтау (алуға) үшін «Заттар интернеті» BSI PAS-212:2016 стандартын көрсетуге болады.

Бастамалары бар, үлкен деректерге қатысты өзге де институттардан төмендегілерді атап көрсетеміз:

- ЭлектротехникажәнеЭлектроника инженерлер институты (IEEE)

- Халықаралық электротехникалық комиссия (IEC)

- Интернеттің инженерлік кеңесі (The Internet Engineering Task Force - IETF)

- Дүниежүзілік ғаламтор консорциумы (World Wide Web Consortium -W3C)

- Ашықгео-консорциум (Open Geospatial Consortium - OGC)және басқалары.

«Ақпаратты Қазақстан – 2020» мемлекеттік бағдарламасына сәйкес ұлттық стандарттау аясында биылғы жылы ҚРМС «Үлкен деректер. Талаптар мен мүмкіндіктер негізіндегі бұлттық есептеулер» (V.3600 негізінде: «Үлкен деректер. Талаптар мен мүмкіндіктер негізіндегі бұлттық есептеулер»)жобасын әзірлеу бойынша жұмыстар жүргізілуде.

ҚР МС «Үлкен деректер. Талаптар мен мүмкіндіктер негізіндегі бұлттық есептеулер» жобасы үлкен деректерді қолданудағы проблемаларды қанағаттандыруға арналған бұлттық есептеулерді пайдалануды қарастырады.

Бұдан бөлек, «Зерде» ұлттық инфокоммуникация холдингі» АҚ негізіндегі «Ақпараттық технологиялар» ТК 34 деректерді өңдеу бойынша келесідей стандарттарды үйлестіруді жүргізу үшін Мемлекеттік стандарттау жоспарына ұсыныстар жолданды.

2017 жылға:

- ҚР МС «Ақпараттық технология. Деректерді өңдеу орталықтары. Тиімділіктің негізгі көрсеткіштері. 1-бөлім. Шолу және жалпы талаптар» ISO/IEC 30134-1:2016негізінде;

- ҚР МС «Ақпараттық технология. Деректерді өңдеу орталықтары. Тиімділіктің негізгі көрсеткіштері. 2-бөлім.Энерготиімділік коэфициенті (PUE)» ISO/IEC 30134-2:2016 негізінде;

- ҚР МС «Ақпараттық технология. Деректерді өңдеу орталықтары. Тиімділіктің негізгі көрсеткіштері. 3-бөлім. Қайта қалпына келетін энергия (REF)» ISO/IEC 30134-3:2016негізінде;

2018-2020 жылдарға:

- ҚР МС «Ақпараттық технология. АТ үшін және АТ көмегімен төзімділік. Деректерді өңдеу орталықтарының интеллектуалды ресурстарын бақылау және қадағалау» ISO/IEC 19395:2015негізінде;

- ҚР МС «Ақпараттықтехнологиялар. Телекоммуникациялар және жүйелер арасында ақпарат алмасу». ISO/IEC 13239:2002 негізіндегі деректерді табыстау каналын басқарудың жоғары деңгейлі хаттамалары;

- ҚР МС «Ақпараттықтехнологиялар. Құжаттарды өңдеу және олармен байланысты деректерді табыстау. ISO/IEC 13673:2000 негізіндегі құжаттарды сипаттаудың стандартты жалпылық тіл жүйесінің конформдылығын тексеру»;

- ҚР МС «Ақпараттық технологиялар. Құжатты сипаттау және деректерді өңдеу тілдері. Гипермәтіндерді баяндау тілі» ISO/IEC 15445:200;0 негізінде.

- ҚР МС «Ақпараттықтехнологиялар. Жиілік менеджментіне арналған радиожиілікті сәйкестендіру. Бағдарламалық қамтамасыз ету жүйесінің инфрақұрылымы. 2-бөлім. Деректерді басқару».