



ӘОЖ 004.932.2

ҒТАХА 81.93.29

https://doi.org/10.53364/24138614_2025_39_4_9

А.Т.Кубигенова^{1*}, А.У.Ақтаева², А.А.Шарипбай³, В.А.Сухомлин⁴, Р.Н.Молдашева⁵

¹С. Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті,
 Астана қ., Қазақстан

² А.Мырзахметов атындағы Көкшетау университеті,
 Көкшетау қ., Қазақстан Республикасы

³Л.Н.Гумилев атындағы Еуразия ұлттық университеті,
 Астана қ., Қазақстан Республикасы

⁴ М.В.Ломоносов атындағы Мәскеу мемлекеттік университеті,
 Мәскеу қ., Ресей Федерациясы

⁵Х. Досмұхамедов атындағы Атырау университеті,
 Атырау қ., Қазақстан Республикасы

¹E-mail: akku_kubigenova@mail.ru *

ҮЛКЕН ДЕРЕКТЕР ТЕХНОЛОГИЯСЫНА ШОЛУ

Аңдатпа. АТ саласындағы үрдістер үнемі өзгеріп отырады және жаңа оқиғаларды алдын ала болжау ұйымның цифрлық ресурстарды, соның ішінде үлкен деректерді басқарудағы сәттілігінің негізгі факторы болып табылады. Бұл зерттеудің мақсаты нақты уақыт режимінде қауіпсіздікті қамтамасыз ету үшін үлкен деректерді зерттеу және талдау.

Библиометриялық талдау нәтижесінде зерттеуге қатысты кілт сөздер енгізілген 137 ақпарат көзінен алынған 323 ғылыми мақала алынды. Зерттеуге журналдық мақалалар, кітаптардың бөлімдері және патенттер кірді. Ғылыми мақалалар Scopus және WoS дерекқорларында алғашқы және қайталама кілт сөздер бойынша сұратылды. Зерттеуде келесі әдістер қолданылды: (а) 2013-2023 жылдар аралығында жарияланған мақалаларға негізделген жүйелі әдеби шолу және (б) 2000-2023 жылдар аралығында мақалаларға библиометриялық талдау жүргізілді, RStudio және Bibliometrix бағдарламалары қолданылды.

Жүргізілген талдау негізінде авторлар алдағы онжылдықтарда цифрлық қоғамның іскерлік әлемі адам мінез-құлқы мен жасанды интеллект туралы терең білімге негізделген киберқауіпсіздікті қамтамасыз ету үшін үлкен деректерді өңдеуге бағытталған стратегиямен анықталатынын қорытындылайды. Бұл талдау техникалық мамандарға цифрлық қоғамдағы АТ және үлкен деректерді енгізуге арналған жаңа технологиялық шешімдерді негіздеу үшін қолданылуы мүмкін, сондай-ақ үлкен деректерді қолданбалы мақсаттарда жаппай өңдеу кезінде киберқауіпсіздікті жақсарту туралы ұсыныстардың негізділігін дәлелдеу үшін пайдаланылуы мүмкін.

Түйін сөздер: үлкен деректер, сандық ресурстар, дерекқорлар, библиометриялық талдау, ақпараттық технологиялар, деректерді интеллектуалды талдау.

Кіріспе.

Технологиялық инновациялар мен деректерді талдау саласындағы жетістіктер, соның

ішінде үлкен деректер, компаниялар үшін ең тиімді коммуникациялық стратегияларды әзірлеуде революциялық мүмкіндіктер ашады.

Бүгінгі таңда Big Data деректерді сақтау мен өңдеудің жаңа парадигмасын ұсынады. Үлкен деректер технологиялармен және ақпараттық жүйелермен қолдау көрсетілетін деректерге негізделген талдау маңызды деректерді алуға және оларды бизнес-инсайттарға айналдыруға мүмкіндік береді. Үлкен деректерді компаниялардың қалай пайдаланатынын және олардың сандық стратегияларды қалай анықтайтынын түсіну үлкен қызығушылық тудырады және осы зерттеудің өзектілігін негіздейді.

Ақпараттық жүйелерді қорғаудың классикалық архитектуралық әдістері жаңа технологияларға, соның ішінде үлкен деректер мен деректерді талдауға қолданылмайды. Үлкен көлемдегі ақпаратты жоғары жылдамдықпен талдау деректер ағындары мен есептеу инфрақұрылымы арасындағы үздіксіз өзара әрекеттесуді талап етеді.

Ақпараттық қауіпсіздік үшін үлкен деректерді зерттеу мен талдаудың мақсаты – үлкен көлемдегі ақпаратты нақты уақытта өңдеу. Алайда, бүгінгі таңда келесі маңызды міндеттер әлі де өзекті:

1. Үлкен деректердің көзі: аналитикада қолданылатын деректердің түпнұсқалығы мен тұтастығы, өйткені үлкен деректер көзі кеңеюде.

2. Үлкен деректер ағындарындағы ақпараттың құпиялылығы.

3. Ақпаратты талдау кезінде үлкен деректер жүйелерін түсінуді жеңілдететін құралдарды визуализациялау.

Ақпараттық технологиялар саласында орын алып жатқан үдерістерді үлкен деректер технологиясын қолдану негізінде талдау ақпараттық қауіпсіздікті қамтамасыз ету жағдайларын ғана емес, сондай-ақ коммуникациядағы көптеген жағымсыз әсерлерді болдырмауға мүмкіндік береді. Үлкен деректермен жұмыс істеуге арналған жаңа АТ-мамандарын даярлау қажеттілігі өздігінен қымбат және көп шығынды талап ететін процесс болып табылады.

Бұл зерттеудің мақсаты-үлкен деректерді өңдеу және сақтау қауіпсіздігін қоса алғанда, жаңа технологиялардың маңыздылығын ескере отырып, үлкен деректердің қауіпсіздігі мен түсіну шеңберін белгілеу.

Деректерді өңдеу, сақтау және беру қауіпсіздігі үлкен деректерді зерттеудің маңызды аспектісі болып табылады. Ақпараттың құны бар және оның ағып кетуі барлық салалардағы бизнес-процестерге айтарлықтай зиян келтіруі мүмкін; АТ-қызметтері жаңа технологияларды қолдау және сүйемелдеу үшін жеткілікті әлеуетке ие болмауы мүмкін; үлкен деректер қауіпсіздігі үлкен деректер қауіпсіздігінің маңызды аспектісі болып табылады.

Осылайша, үлкен деректер жаңа технологияларды зерттеуге және кәсіпорындарды дамыту үшін ақпараттық стратегияларды алуға мүмкіндік беретін заманауи цифрлық қоғамның дамуындағы маңызды құбылыс болып табылады.

Зерттеу мәселелерін шешу үшін таңдалған әдістеме 2013-2023 жылдар аралығында жарияланған мақалаларға негізделген әдеби шолуды және 2000-2023 жылдар аралығында мақалаларға библиометриялық талдауды қамтиды.

Төртінші өнеркәсіптік революцияның цифрлық технологиялары

Өнеркәсіптік революция (ӨР) әдетте өндіріс құралдары мен өндірісті ұйымдастырудағы революциялық өзгерістерді білдіреді, нәтижесінде доиндустриалды қоғамнан индустриалды қоғамға өту жүзеге асты.

Қазіргі цифрлық қоғамда «өнеркәсіптік революция» термині әдетте техникадағы, технологиялардағы және адамның еңбек құралдарымен байланысу тәсілдеріндегі төңкерістердің әсерінен қоғамда болып жатқан сапалы өзгерістерді білдіреді.

«Өнеркәсіптік революция» түсінігін «технологиялық революция» ұғымымен жиі шатастырады. Технологиялық революцияны технологиялық парадигманың өзгеруі, өндіріс негізінде жатқан негізгі технологиялардың жиынтығы ретінде анықтауға болады.

Технологиялық революция экономиканың әртүрлі салаларында және әлеуметтік салада еңбек өнімділігін түбегейлі немесе экспоненциалды түрде арттыратын технологиялық шешімдерді жаппай қолдануға негізделген экономикалық қызметті жүргізу тәсілін сапалы өзгертуді білдіреді.

Индустрия 4.0 дамуының қозғаушы күші ендірілген жүйелерден киберфизикалық жүйелерге (КФЖ) көшуі.

Енгізілген жүйелер – әртүрлі басқару объектілеріне орнатылған орталық басқару блоктары.

Киберфизикалық жүйелер – виртуалды және физикалық әлемдерді байланыстыратын, интеллектуалды объектілердің интернетті, желілерді және деректерді пайдалана отырып, бір-бірімен өзара әрекеттесуіне мүмкіндік беретін жаңа технологиялардың (enabling technologies) жиынтығы.

Industry 4.0 «Болашаққа арналған жоспар» құжаты сонымен қатар енгізілген жүйелерден киберфизикалық жүйелерге көшуді төртінші өнеркәсіптік революция деп атайды.

Төртінші өнеркәсіптік революцияның негізгі принциптері. Өндіріс цифрландыру жағдайында машиналардың, өндірістік жүйелер мен жабдықтардың өзара байланысы арқылы бизнес-процестерді өздігінен бақылау және басқару үшін интеллектуалды желілерді құруға мүмкіндік беретіні туралы идеяны қолдайды.

Цифрлық технологиялар – төртінші өнеркәсіптік революцияның негізгі қозғаушы күші. Заттар интернеті, жасанды интеллект, машиналық оқыту және үлкен деректер сияқты концепциялар деректер көлемін айтарлықтай ұлғайтуға әкелетін цифрлық технологияның кейбір аспектілері болып табылады, бұл сақтау сыйымдылығын арттыруға және машиналық оқытудағы жетістіктерге әкеледі [1].

Рейнсель және т.б. 2018 жылы жасалған, жиналған немесе репликацияланған деректер көлемі 33 зеттабайттан 2025 жылға қарай 175 зеттабайтқа дейін өсетінін болжады. Data Age Report мәліметтері бойынша 2020 жылы адамзат шамамен 51 зеттабайт ақпарат қалыптастырды. 2025 жылға қарай бұл деректер көлемі жылына 175 зеттабайтқа дейін өседі. (1-сурет).



1 сурет – Қайталанатын Big Data көлемі

Индустрия 4.0 қолданылатын жаңалықтардың бірі - Big Data деп аталатын деректердің жинақталуы. Үлкен деректерді жинақтау – бұл сандық қоғам жағдайында бизнес - мақсаттарды шешуге мүмкіндік беретін коммерциялық міндет.

Зерттеу материалдары мен әдістері.

Үлкен деректер концепциясы жақында пайда болды: компаниялар алғашқы электронды есептеуіш машиналар пайда болған өткен ғасырдың ортасынан бастап

деректерді жинай бастады. Компьютерлер мен сақтау жүйелерінің қымбаттығына байланысты деректердің өсуі баяу болды.

Осылайша, "үлкен деректер" - бұл бастапқы сақтау ресурстары ретінде пайдаланылатын ақпараттық жүйелер мен дерекқорлар белгілі бір уақыт ішінде үлкен көлемдегі деректерді ғана емес, сонымен қатар нақты операцияларды да сақтай алатынын білдіретін жалпы термин.

Бүгінгі таңда қол жетімді құралдар үлкен деректерге тән барлық мәселелерді шеше алмайды. Дегенмен, олар аз уақыт ішінде әртүрлі деректерді өңдеуді біршама жеңілдетуге мүмкіндік береді.

Үлкен деректер тым үлкен және дәстүрлі әдістермен өңделмейтін деректер жиынын білдіреді, бұл өңдеудің жаңа технологияларын сақтау, басқару, талдау және визуализацияның озық және бірегей технологияларын талап етеді. Үлкен деректердің үш сипаттамасы: көлем, әртүрлілік және жылдамдық.

Деректер көлемі әр түрлі көздерден секунд сайын жасалатын деректер көлемін білдіреді. Үлкен деректердің көлемі деректердің уақыты мен түрі сияқты факторларға байланысты. Қазіргі уақытта үлкен деректер деп саналатын нәрсе болашақта шекке жетпеуі мүмкін, өйткені сақтау сыйымдылығы артады және одан да үлкен көлемдегі деректер жиналуы мүмкін.

Сонымен қатар, бірдей өлшемдегі екі деректер жиынтығы олардың түріне байланысты деректерді басқарудың әртүрлі әдістерін қажет етуі мүмкін, мысалы, мәтіндік деректер және бейне деректер.

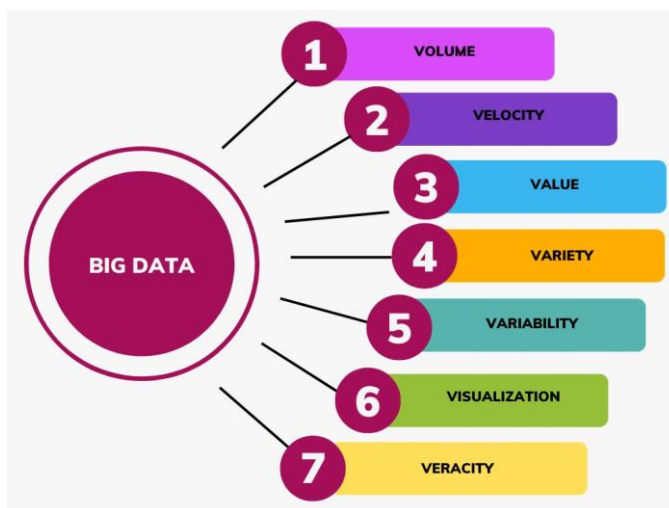
Жылдамдық көлеммен тығыз байланысты. Деректерді алу жылдамдығы неғұрлым жоғары болса, соғұрлым жасалған көлем үлкен болады. Бұл сипаттама ақпаратты ғаламтор арқылы беру және тарату жылдамдығына да, нақты уақыт режимінде деректерді талдауға қажетті жылдамдыққа да қатысты. Компаниялардың тұтынушылар деректерін алу, талдау және пайдалану жылдамдығы оларға бәсекелестерінен артықшылық береді. Ал нарықтағы ақпараттың көлемі мен әртүрлілігін ескере отырып, тәжірибе көрсеткендей, бәсекелестік артықшылық бірнеше минут ішінде жоғалуы мүмкін. Әртүрлілік деректер жиынының құрылымдық біркелкі еместігін білдіреді.

Датчиктер мен интеллектуалды құрылғылардың пайда болуымен деректердің күрделілігі артып, тек дәстүрлі деректерді ғана емес, сондай-ақ өңделмеген, жартылай құрылымдалған және құрылымдалмаған деректерді де қамтиды.

Дюкандж және т.б. бұл олқылықты жою үшін жаңа шешімдер қажет екенін мойындайды, өйткені дәстүрлі технологиялар құрылымдалмаған деректермен тиімді жұмыс істей алмайды.

Дегенмен, көлем, жылдамдық және әртүрлілік – бұл үлкен деректердің ерекше сипаттамаларының кейбіреулері ғана.

Тернер және т.б. көлем, әртүрлілік және жылдамдықты үлкен деректердің негізгі сипаттамалары деп санайды. (2-сурет).



2 сурет – Үлкен деректердің негізгі параметрлері

Сандық ортадағы ақпарат көздерінің көбеюі деректердің сенімділігі мен сенімділік деңгейін талдауды қажет етеді. Үлкен деректер дәуірінде компаниялар деректердің белгісіздігін мойындауы, оған бейімделуі және оны пайдалану жолдарын анықтауы керек [2].

Сенімділік алынған мәліметтердің сапасы мен сенімділігін білдіреді. Бұл аспект үлкен деректер мен қауіпсіздік әлемінде іргелі болып табылады, өйткені деректерге қол жеткізген кез келген адам қол жеткізетін деректердің тұтастығын тексеруге жауапты. Жалған деректер талдаудағы қателерге әкелуі мүмкін, бұл өз кезегінде компанияның стратегиясына қауіп төндіретін нашар шешімдерге әкелуі мүмкін.

Үлкен деректер сенімді нәтижелерді алу үшін арнайы құралдар мен үлкен алгоритмдерді қажет етеді.

Үлкен деректердің көлемі өскен сайын оның ішкі мәнін және оны талдау компанияға қалай пайда әкелетінін ескеру қажет. Мәселе маңызды емес немесе маңызды емес деректерді жою және қалған деректердің кәсіпорын үшін пайдалы болуын қамтамасыз ету болып табылады.

Ең тиімді үлкен деректер шешімдері бизнес талаптарын анықтаудан басталатын, содан кейін бизнес мүмкіндігін қолдау үшін инфрақұрылымды, деректер көздерін, процестер мен мүмкіндіктерді бейімдейтін шешімдер болып табылады.

Үлкен деректер компанияда неғұрлым көп ақпарат болса, соғұрлым ол сенімдірек болады және жаңа түсініктерге ие болып, тұтынушылардың іс-әрекеті туралы болжам жасай алады деген принцип бойынша жұмыс істейді.

Үлкен деректердің сандық ресурстарын басқару стратегиясы

Үлкен деректер цифрландыру арқылы басқару стратегияларын өзгертеді және «Өнеркәсіптік революция 4.0» деп аталатын бизнестің бәсекеге қабілеттілігінің жаңа шекарасын білдіреді. Үлкен деректер технологиясы цифрлық ресурстарды басқарудың дәлірек, мақсатты және креативті стратегияларын қолдай алады. Тұжырымдамалық модель NBDRA (NIST Big Data Reference Architecture) - бұл үйлесімділік интерфейстерімен байланысты бес логикалық функционалды компоненттерден тұратын үлкен деректер жүйесі.

Бүкіл әлемде жаңа бизнес - модельдерді енгізу және бәсекелестік артықшылықтардың жаңа көздерін дамыту үшін үлкен деректер қолданыла бастады. Бәсекелестік артықшылықтың жаңа көзі - клиентке бағдарлану: кез келген басқа компанияға қарағанда оларға жақсырақ қызмет ету үшін олардың қажеттіліктерін терең түсіну.

Компаниялар өз ресурстарын шоғырландырып, қажетті құзыреттердің бірыңғай жүйесін құруы керек. Қажетті құзыреттердің бірыңғай жүйесі цифрлық активтерді басқару стратегиясында жаңа мүмкіндіктерді пайдалану және клиенттер мен мүдделі тараптар үшін пайда алу үшін берік негіз қалайды.

Бұл ретте компаниялар цифрлық деректердің қауіпсіздігі мен құпиялылығына қатысты ықтимал проблемаларды азайту үшін инновациялар мен пайда табуды тұтынушылардың күтулерімен және этикалық нормалармен біріктіруі керек.

Үлкен деректер технологияларын қорғаудың қолданыстағы тәсілдері, әдетте, бірыңғай қауіпсіздік тұжырымдамасы болмаған кезде әртүрлі құралдарды пайдалануға негізделген.

Бүгінгі таңда жинау, біріктіру, сақтау және талдау сияқты бірегей техникалық сипаттамалары бар үлкен деректерді, құрылымдалған және құрылымданбаған деректерді қорғаудың құрылымдық процедуралары мен әрекеттерін сипаттайтын нақты тұжырымдалған әдістеме жоқ. Жинау мен беруден бастап талдауға және қоймаға орналастыруға дейін өңдеудің барлық кезеңдерінде маңызды деректерді қорғауға бағытталған тәсіл қажет.

Мысалы, NIST бастапқыда деректерді өңдеудің барлық техникалық деңгейлерінде деректердің қауіпсіздігі мен құпиялылығына назар аударуды ұсынды, бұл деректермен өзара әрекеттесудің бес негізгі интерфейсіні қамтиды.

Үлкен деректер жүйелерінің қауіпсіздігіне жеткіліксіз көңіл бөлінеді. Көптеген жобалар ақпараттық қауіпсіздік мәселелерін есепке алмастан әзірленеді және жүзеге асырылады, бұл ерте ме, кеш пе, қауіпсіздік жүйелерін енгізудің уақыты мен құнын айтарлықтай арттырады, кейде бизнес үшін аса ауыр зардаптарға әкеледі. Үлкен деректер жобаларын жүзеге асыру кезінде қауіпсіздік мәселелері ең басынан бастап қарастырылуы керек. Әйтпесе, жоба бизнес мүмкіндігінен жаңа бизнес тәуекеліне айналуы мүмкін.

Нәтижелер және олардың талқылануы.

Әдебиетке жүйелі шолу жасау. Цифрлық деректерді (ресурстарды) басқару стратегияларында төртінші технологиялық революцияның жаңа технологияларын қолдануға деген аңғалдық кейбір ұйымдардың цифрлық стратегияларды дамытудағы осы технологиялардың мүмкіндіктері туралы жеткіліксіз хабардар болуын көрсетуі мүмкін.

Бұл жұмыстың мақсаты – компаниялардың үлкен деректерді пайдалану дәрежесін және олардың цифрлық стратегияларына әсерін зерттеу және талдау. Шолуды басшылыққа алу үшін келесі сұрақ қойылды: Үлкен деректер қауіпсіздігі мәселелерінің цифрлық активтерді басқару стратегияларына әсері қандай?

Цифрлық ресурстарды басқару стратегияларына қатысты үлкен деректер туралы ағымдағы түсінікті нығайту және осылайша қойылған сұрақтарға жауап беру үшін біз алдымен әдебиеттерге жүйелі шолу жасап, содан кейін библиометриялық шолу жасауды шештік.

Жүйелі шолулар, дәстүрлі шолулар сияқты, білімдегі олқылықтарды анықтауға көмектеседі. Алайда, әдебиеттерді жүйелі түрде шолуға қызығушылық - бұл процестің объективтілігі мен ашықтығын көрсететін бейтарап, ұтымды және стандартталған технологиялық процесс.

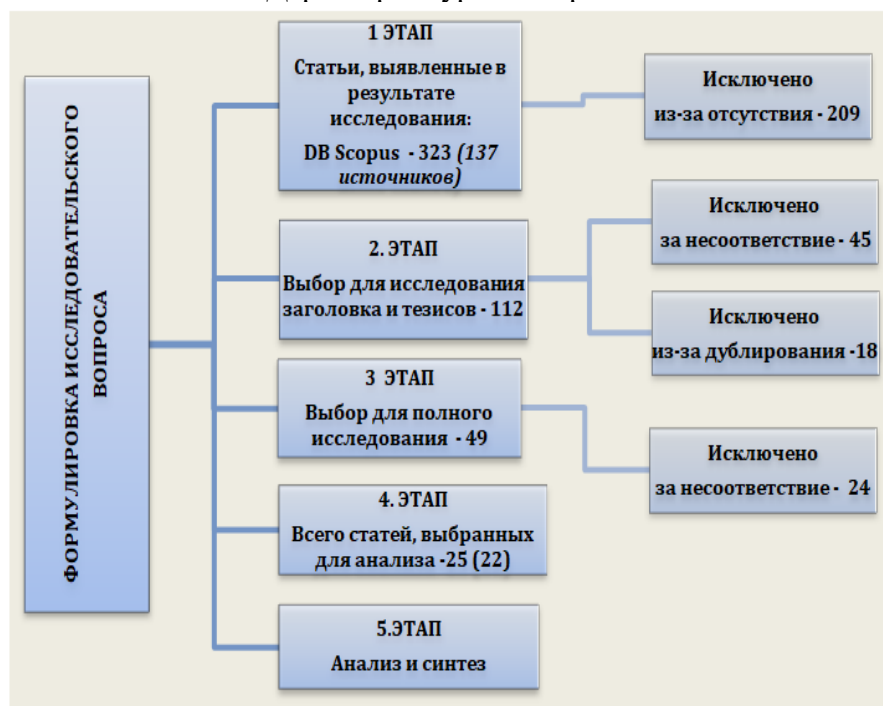
Тиімді шолу білімді ілгерілету үшін күшті негіз береді, теорияның дамуын жеңілдетеді, көп зерттеулер бар аймақтарды жабады және зерттеу қажет аймақтарды ашады. Осы себепті ғылыми қатаңдықты қамтамасыз ету және біржақтылықты азайту үшін жүйелі шолу процесін қайталауға болады.

Әдебиеттерді жүйелі шолу процесі медицина ғылымынан бастау алады және жылдар өткен сайын басқа пәндер бойынша білім құруға, әсіресе гуманитарлық зерттеулерге инвестиция салынды, бұл әдіс жүргізілетін зерттеулерді кластерлеу үшін маңызды бола бастады [3].

Зерттеу сұрағының тұжырымы. Зерттеулерді іріктеу процесі 2023 жылдың қаңтарында өтті және Scopus мәліметтер базасында әдебиеттерді іздеу жүргізілді, оның барысында 137 дереккөзде 323 құжат тіркелді. Scopus дерекқорында TOPIC: ("үлкен деректер") және TOPIC: ("криптография") қолданғаннан кейін 2013 және 2023 жылдар аралығында 323 мақала алынды.

Қосу критерийлері: 2013 және 2023 жылдар аралығындағы мақалалар, аталған дерекқорларда толық оқуға қол жетімді және цифрлық стратегиялардағы үлкен деректерге баса назар аударылды. Ерекшелік критерийлері: қайталанатын мақалалар, іріктеу критерийлерімен байланысты емес зерттеулер, жүйелі шолулар мен диссертациялар, сондай-ақ қорытындысы жоқ құжаттар.

6-суретке сәйкес авторлар зерттеудің нақты бағытын анықтаудың маңыздылығын мойындайды деген қорытынды жасау керек. Сондықтан бірінші кезең зерттеу сұрақтарын тұжырымдауға арналған. Бұл сұрақ қандай зерттеулерді қосу керектігін, тиісті бастапқы зерттеулерді анықтау үшін қандай іздеу стратегиясын қолдану керектігін және әрбір зерттеу үшін қандай деректерді алу керектігін насихаттайтын шолуды бағыттайды. Әдебиетке жүйелі шолу бес кезеңге бөлінеді. Деректер 3-суретте көрсетілген.



3 сурет – Зерттеу сұрағының тұжырымы: жүйелі шолу кезеңдері

Зерттеудің орналасуы. Әдебиеттерге шолулар жасау нақты сұрақтарына сәйкес келетін зерттеулерді мүмкіндігінше табуға, таңдауға және бағалауға тырысады. Зерттеулерді жан-жақты іздеу шолу нәтижелері барлық қолда бар дәлелдерді ескеретініне және жоғары сапалы материалдарға негізделгеніне көз жеткізуге мүмкіндік береді.

Зерттеуді таңдау және бағалау. Қарастырылатын зерттеулердің тек шолу сұрағына жауап беруге қатысы бар зерттеулердің болуын қамтамасыз ету үшін іріктеу критерийлері қолданылады. Шешімдер ақпарат көздерін қосу және алып тастау себептерін дәл көрсете отырып жазылады.

Талдау және синтез. Тиісті дереккөздердің жинағын алғаннан кейін, ақпаратты талдап, қорытындылайтын уақыт келді. Бұл қадамның мақсаты-әртүрлі зерттеулерді талдау және олардың өзара байланысын сипаттау.

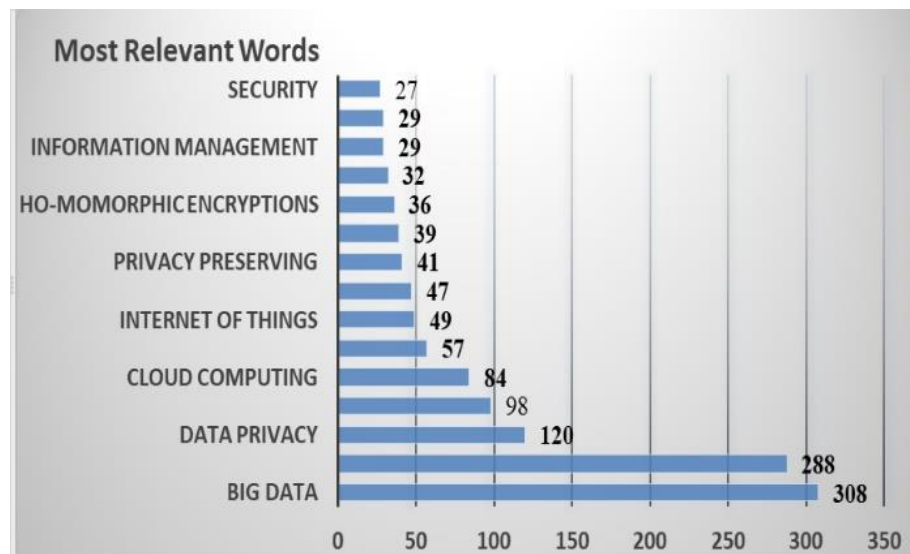
Жүйелі шолудың соңында зерттеу саласының толық жиынтық көрінісін қамтамасыз ететін барлық енгізілген зерттеулердің ақпараты көрсетіледі.

Нәтижелерді жариялау. Соңғы кезеңде алынған нәтижелер ұсынылады және талқыланады.

Осы тақырып бойынша ғылыми әдебиеттерді зерттеу кезінде үлкен деректер саласындағы зерттеушілердің қызығушылығының артуы үлкен деректердің қауіпсіздігін және олардың қолданылуын қамтамасыз ету тұжырымдамаларын ұсынатын зерттеулер мен мақалалардың көбеюі байқалды.

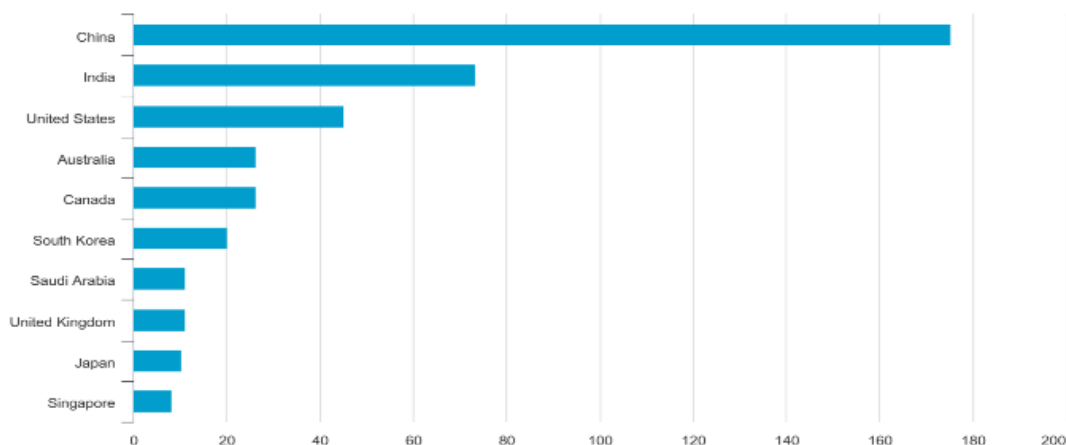
Осы тұрғыдан алғанда, бұл зерттеу цифрлық ресурстардың қауіпсіздігін басқару стратегияларында үлкен деректердің қалай қолданылатынын анықтау, талдау және сыни талқылау арқылы әдебиеттерге жалпы шолу жасауға бағытталған.

Іздеудің тиімді болуы үшін және мүмкіндігінше тиісті зерттеулерге әкелуі үшін іздеу терминдері ретінде "үлкен деректер" және "криптография" анықталды. Бұл критерийлер Зерттеудің мақсаты мен зерттеу мәселелерін ескере отырып анықталды (4-сурет).



4 сурет – SCOPUS дерекқорындағы зерттеулерді іріктеу процесі

Bibliometrix бағдарламалық құралын пайдалану бірте-бірте ғылыми зерттеулердің барлық пәндеріне таралады және эмпирикалық үлеске баса назар аудару цифрлық ресурстардың үлкен, бөлшектелген және қарама-қайшы ағындарына әкелетін уақытта (5-Сурет) осы зерттеуді картаға түсіру үшін қолайлы. Желілік матрицаны құру үшін R Studio, Bibliometrix - <http://www.bibliometrix.org> бағдарламалық жасақтамасы қолданылды.



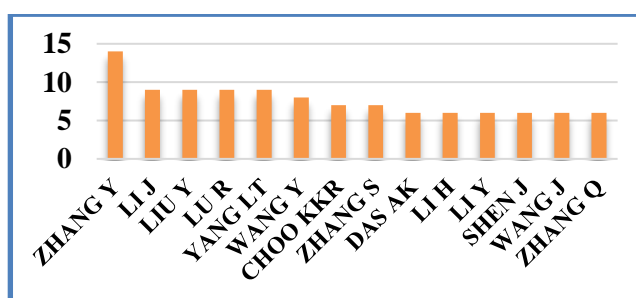
7 сурет – Елдер бойынша бірлескен авторлар желісі

Тақырыптық сөз бейнесі – бұл деректердегі сөздердің жиілігі мен маңыздылығын бірден түсінуге мүмкіндік беретін визуалды ұсыну түрі – бұл жағдайда мақалалардағы кілт сөздер. Мақалалардағы терминдердің (кілт сөздердің) санына пропорционалды қаріп өлшемі өзгеріп отырады, бұл олардың ең маңыздысын жеңіл қабылдауға мүмкіндік береді.

Бұл опцияның мақсаты мүмкіндігінше кең ауқымды сөздерді қамту және әрбір бірлескен сілтемеге бірдей салмақ беретін толық санау әдісін пайдалану болды. Ең жиі қолданылатын келесі кілт сөздерді бөліп көрсетуге болады: «үлкен деректер», «криптография», «деректердің құпиялылығы», «сандық сақтау», «желілік қауіпсіздік», «интернет заттары», «кіруді басқару», «құпиялылық» және «аутентификация».

«Кілт сөздерді сәйкестендіру желісі (KCN) әдебиеттегі кілт сөздер арасындағы байланыстарды зерттеу арқылы ғылыми/техникалық саладағы білім компоненттері мен білім құрылымын түсінуге бағытталған» [4].

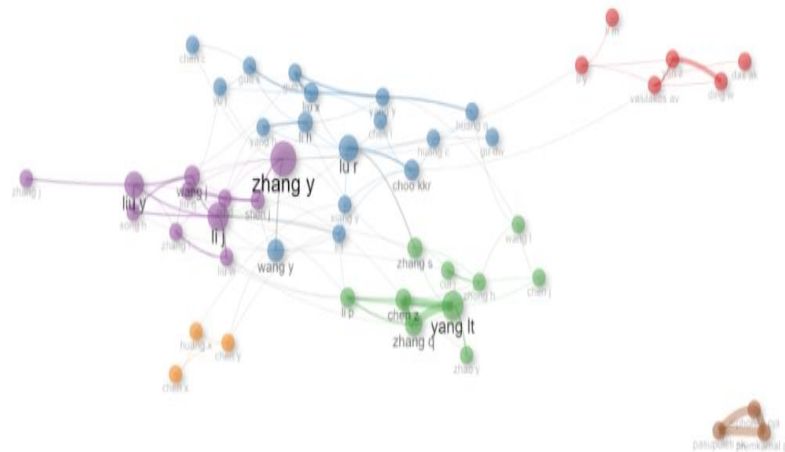
Зерттеу барысында дерекқорда кемінде бес құжаты және өз мақалаларының кемінде үш дәйексөзі бар авторлар ғана ескерілді. Бұл шектеу 17 топқа топтастырылған 35 элементі (авторлары) бар желіні құрды, бұл осы саладағы зерттеулерге қатысты дисперсияның белгілі бір дәрежесін көрсетеді (8-сурет).



8 сурет – Бірлескен авторлық желі

Бірлескен авторлықты білдіретін элементтері көп топтар-1 (қызыл) және 2 (жасыл) кластерлер. 1-кластердің барлық авторларында мақалалар саны бірдей және сілтемелер саны бірдей.

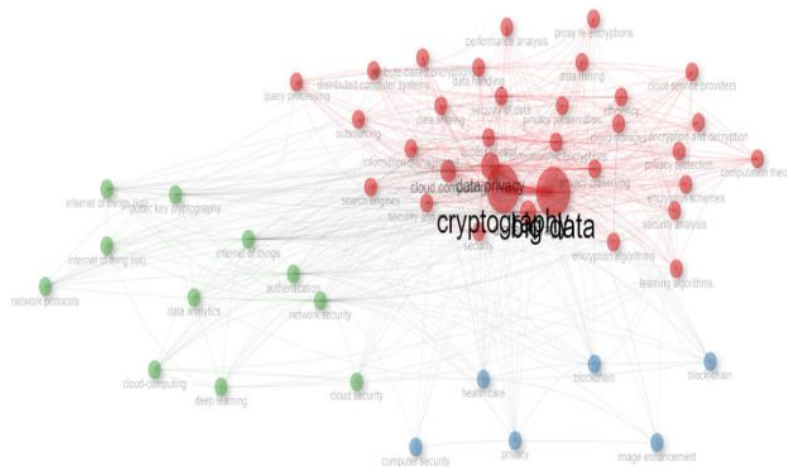
2-кластерде Zhang Y. үлгідегі ең көп мақалалары бар автор [5] және бірлескен авторлыққа ең көп сілтемелері бар авторлар. Сол кластерде көптеген құжаттары бар басқа авторларды табуға болады, мысалы Li J., Liu Y., Lu R., және Yang L. [5,6] Т. Қалған кластерлер келесідей жіктеледі: үш элементі бар бір кластер, екі элементі бар үш кластер және он екі кластерде тек бір элемент бар (9-сурет).



9 сурет – Авторлардың бірлескен авторлық кластері

10-сурет біздің зерттеуіміздің KCN-ін ұсынады. Мақалаларды талдау үш кластерге әкелді, атап айтқанда: 1-кластерде (қызыл) 34 термин, 2-кластерде (жасыл) 11 термин және 3-кластерде (көк) 6 термин бар [7]. Бұл кластерді бұрын жасалған мақалаларды талдаумен салыстырған кезде, қолданылатын терминдердегі сәйкестікті көруге болады.

1-кластер: үлкен деректерді талдау: деректер оларға өз аудиториясын білуге және осылайша олардың үміттеріне сәйкес ұсыныстар беруге мүмкіндік береді. Деректер көлемі орасан зор және кез келген субъектінің оңай қолы жететін ортада тұтынушылар өздерінің жеке деректерін рұқсат етілмеген мақсаттарда заңсыз пайдалану туралы көбірек алаңдайды.



10 сурет – Кілт сөздерді сәйкестендіру желісі

Үлкен деректерді талдау тәртібін анықтайтын негізгі элемент ұйымдарға ақпараттың сипаты мен мақсатына тұтынушылардың қауіпсіздігі мен сенімділігін қамтамасыз етуге ықпал ететін деректерді қорғаудың ашық тәжірибелері мен саясаттарын енгізу қажет [8]. Негізгі есеп беру, ашықтық және ашықтық цифрлық ресурстардың қауіпсіздік стратегиясын басқару, олардың өзара әрекеттесуі және шешім қабылдау үшін үлкен маңызға ие.

2-кластер: Үлкен деректердің сандық активтер қауіпсіздігін басқару стратегияларына әсері Активтердің қауіпсіздігін басқарудың сандық бизнес стратегиялары бизнестің бәсекеге қабілеттілігі мен өзектілігін сақтау үшін жаңа ресурстарды қарастыруды қажет етеді.

Заманауи сценарийде деректерді өндіру және талдау басқару процесінің тиімділігі мен нәтижесін арттыруға ықпал етеді. Кәсіпорындар нақты уақыт режимінде талдауға болатын және болашақ стратегияларды байытатын деректердің керемет көлеміне қол жеткізе алады.

Сандық активтердің қауіпсіздігін басқару стратегияларына арналған талдау құралдары өнімділік пен қаржылық кірістің жақсаруына әкелетін неғұрлым негізделген және сенімді шешімдер қабылдауға мүмкіндік береді.

3-кластер: Технологиялық өзгерістерге ілесу үшін компаниялар енгізуі керек жүйелердің түрлері [9]. Нақты уақытта деректерді талдау мүмкіндігі кез келген компания үшін тағы бір артықшылық болып табылады. Бұл, әсіресе, жоғары жылдамдықты нарықтарда жұмыс істейтін компанияларға қатысты, мұнда олар сандық активтерді кепілді басқару стратегияларын жүзеге асыру үшін қиындықтар мен мүмкіндіктерді тезірек анықтау керек. Бұл оларға тұтынушылардың талаптарын қанағаттандыратын тиімдірек және креативті ұсыныстарын жасауға мүмкіндік береді [10].

Компаниялар автоматтандырылған жасанды интеллект агенттерімен жабдықталған үлкен деректер қоймаларын енгізеді, олар машиналық оқыту әдістеріне және Индустрия 4.0 технологияларына, әсіресе үлкен деректерге және онымен байланысты технологияларға сүйенеді [11].

Қорытынды.

Сандық активтерді басқарудағы Big Data немесе Үлкен Деректер рөлі соңғы бірнеше жылда артты. Сондықтан цифрлық қоғамда Big Data ақпараттық технологияларын енгізу тақырыбы бойынша осы зерттеудің негізгі үлесі қазіргі уақытта осы білім саласында не жарияланғаны туралы көрнекі түсінік болып табылады. Біз осы тақырып бойынша ең маңызды және өзекті жарияланымдарды анықтадық, ең үлкен ғылыми қызығушылық тудырған тақырыптарды әзірледік және цифрлық контексте технологияны пайдалану туралы толық ақпарат ұсындық. Зерттеу сұрақтарына жауап беру үшін RStudio бағдарламалық құралының көмегімен жүйелі әдебиеттерге шолу және библиометриялық талдау әдістемесі таңдалды.

Үнемі өзгеріп отыратын АТ тенденцияларына байланысты жаңа оқиғаларды болмай тұрып болжау ұйымның заманауи цифрлық активтерді басқарудағы табысының маңызды бөлігі болып табылады. Егер цифрлық активтерді қорғау жөніндегі стратегтер шешім қабылдаудың негізгі мотивтерін түсінсе, олар белгілі бір сенімділік дәрежесімен қарсыластардың болашақ іс-әрекетін болжай алады.

Мысалы, тұтынушылардың мінез-құлқы мен әдеттерін болжау арқылы кәсіпорынның цифрлық активтерінің қауіпсіздік менеджерлері бұзушы тұтынушыларды анықтап, сегменттей алады. Шынында да, мақсатты аудиторияны және әрбір сегментке ең үлкен әсер ететін хабарларды түсіну барлық салаларда бизнес-процестерді дәлірек ете алады. Цифрлық даму моделі экономика мен өнеркәсіпті толық цифрландыруды көздейді. Бұл өнімдердің, жабдықтың және өндірістің «ақылды моделін» немесе «ақылды цифрлық егізін» пайдалану. Таяу болашақта ғылымды дамыту, жоғары білікті кадрлар даярлау, озық өндіріс технологияларын енгізу және әлемдік деңгейдегі бірегей құзыреттерді дамыту да басты басымдықтар болмақ. Дегенмен, ақпараттың қол жетімділігі оның қауіпсіздігіне кепілдік бермейді және компанияның бизнес мәселелерінің нақты шешімін қамтамасыз етпейді. Деректерді талдау шешімдері жай ғана құрал емес, нарық пен компанияның өзі жасаған құнды ақпарат негізінде ақпарат алу және шешім қабылдау құралы болып табылады.

Бизнестің бәсекеге қабілеттілігі артқан сайын, ұйымның стратегиясы бизнестің бәсекеге қабілеттілігі мен өзектілігін сақтаудың жаңа ресурстары мен жолдарын көбірек есепке алуы керек. Бұл нақты және сенімді ақпаратқа, сондай-ақ басқарушылық интуицияға негізделген шешім қабылдауды талап етеді. Ал бұл бизнес мәселелерін шешуге мүлдем басқа көзқарас болып табылады. Үлкен деректердің аналитикалық мүмкіндіктері мақсатты

аудиторияның мінез-құлқы туралы маңызды түсініктер бере алады және үлкен деректер ресурстарының цифрлық қауіпсіздігін басқару стратегиясының арқасында компанияны нарықта неғұрлым өткір және тиімді орналастыруға көмектеседі.

Әдебиеттер тізімі

1. Четвертая промышленная революция. URL: <https://www.tadviser.ru/index.php>
2. Проблемы безопасности больших данных. (2017). URL: <https://www.osp.ru/os/2017/04/13053380>
3. Туароб, С., Бхатиа, С., Митра, П., & Джайлз, К. Л. (2016) «AlgorithmSeer: система для извлечения и поиска алгоритмов в научных больших данных», IEEE Trans. Большие данные, том. 2, нет. 1, стр. 3–17.
4. Ся, Ф., Ван, В., Бекеле, Т. М. & Лю, Х. (2017). «Большие научные данные: обзор», IEEE Trans. Большие данные, том. 3, нет. 1, стр. 18–35.
5. Li, J., Zhang, Y., Chen, X. & Xiang, Y. (2018). «Безопасный обмен данными на основе атрибутов для пользователей с ограниченными ресурсами в облачных вычислениях», Comput. Безопасность, том. 72, стр. 1-12. <https://doi.org/10.1016/j.cose.2017.08.007>
6. Ли, П., Ли, Дж., Хуанг, З., Гао, К.З., Чен, В.Б. & Чен, К. (2017). «Аутсорсинговая классификация с сохранением конфиденциальности в облачных вычислениях», Cluster Comput., стр. 1-10.
7. Сухак, М., Гани, А., Хан, М.К. & Буйя, Р. (2017). «Динамический удаленный аудит данных для обеспечения безопасности хранения больших данных в облачных вычислениях», Inf. науч., том. 380, стр. 101-116.
8. Chen, X., Li, J., Huang, X., Ma, J. & Lou, W. (2015). «Новые публично проверяемые базы данных с эффективными обновлениями», IEEE Trans. Зависеть. Безопасный компьютер., том. 12, нет. 5, стр. 546-556.
9. Оренчик, К., Сельчук, А., Саваш, Э. & Кантарчиоглу, М. (2016). «Поиск по нескольким ключевым словам в зашифрованных данных с оценкой и запутыванием шаблона поиска», Int. Дж. Инф. Безопасность, том. 15, нет. 3, стр. 251-269.
10. Шен, Дж., Чжоу, Т., Чен, Х., Ли, Дж. & Сусило, В. (2018). «Обмен анонимными и отслеживаемыми групповыми данными в облачных вычислениях», IEEE Trans. Судебно-медицинская экспертиза, том. 13, нет. 4, стр. 912-925.
11. Цзян, Т., Чен, Х., & Ма, Дж. (2016). «Аудит общедоступной целостности для общих динамических облачных данных с отзывом группы пользователей», IEEE Trans. вычисл., том. 65, нет. 8, стр. 2363-2373.

References

1. Chxetvertaya promyshlennaya revolyuciya. URL: <https://www.tadviser.ru/index.php>
2. Problemy bezopasnosti bol'shih dannyh (2017). URL: <https://www.osp.ru/os/2017/04/13053380>
3. Tuarob, S., Bhatia, S., Mitra, P. & Dzhajlz, K. L. (2016). «AlgorithmSeer: sistema dlya izvlecheniya i poiska algoritmov v nauchnyh bol'shih dannyh», IEEE Trans. Bol'shie dannye, tom. 2, net. 1, str. 3–17.
4. Sya, F., Van, V., Bekele, T. M. & Lyu, H. (2017). «Bol'shie nauchnye dannye: obzor», IEEE Trans. Bol'shie dannye, tom. 3, net. 1, str. 18–35.
5. Li, J., Zhang, Y., Chen, X. & Xiang, Y. «Secure attribute-based data sharing for resource-limited users in cloud computing», Comput. Bezopasnost', tom. 72, str. 1-12. <https://doi.org/10.1016/j.cose.2017.08.007>.
6. Li, P., Li, Dzh., Huang, Z., Gao, K.Z., Chen, V.B. & CHen, K. (2017). «Autsorsingovaya klassifikaciya s sohraneniem konfidencial'nosti v oblachnyh vychisleniyah», Cluster Comput., str. 1-10.

7. Suhak, M., Gani, A., Han, M.K. & Bujya, R. (2017). «Dinamicheskij udalennyj audit dannyh dlya obespecheniya bezopasnosti hraneniya bol'shikh dannyh v oblachnyh vychisleniyah», Inf. nauch., tom. 380, str. 101-116.

8. Chen, X., Li, J., Huang, X., Ma, J. & Lou, W. (2015). «Novye publichno proveryaemye bazy dannyh s effektivnymi obnovleniyami», IEEE Trans. Zaviset'. Bezopasnyj komp'yuter., tom. 12, net. 5, str. 546-556.

9. Orenchik, K., Sel'chuk, A., Savash, E. & Kantarchioglu, M. (2016). «Poisk po neskolk'im klyuchevym slovam v zashifrovannyh dannyh s ocenкой i zaputyvaniem shablona poiska», Int. Dzh. Inf. Bezopasnost', tom. 15, net. 3, str. 251-269.

10. SHen, Dzh., CHzhou, T., CHen, X., Li, Dzh. & Susilo, V. (2018). «Obmen anonimnymi i otslezhivaemymi gruppovymi dannymi v oblachnyh vychisleniyah», IEEE Trans. Inf. Sudebno-medicinskaya ekspertiza., tom. 13, net. 4, str. 912-925.

11. Czyan, T., Chen, X. & Ma, Dzh. (2016). «Audit obshchedostupnoj celostnosti dlya obshchih dinamicheskikh oblachnyh dannyh s otzyvom grupy pol'zovatelej», IEEE Trans. vychisl., tom. 65, net. 8, str. 2363-2373.

ТЕХНОЛОГИИ БОЛЬШИХ ДАННЫХ: ОБЗОР

Аннотация. Тенденции в сфере ИТ постоянно меняются, и прогнозирование новых событий является ключевым фактором успешного управления цифровыми ресурсами организации, включая большие данные. Цель данного исследования — изучение и анализ больших данных для обеспечения безопасности в режиме реального времени.

В результате библиометрического анализа было получено 323 научные статьи из 137 источников информации, содержащих ключевые слова, связанные с исследованием. В исследование вошли журнальные статьи, главы книг и патенты. Научные работы были отобраны в базах данных Scopus и WoS по основным и вторичным ключевым словам. В исследовании применялись следующие методы: (а) систематический обзор литературы на основе публикаций за период 2013–2023 гг.; (б) библиометрический анализ статей, опубликованных в 2000–2023 гг., с использованием программ RStudio и Bibliometrix.

На основе проведенного анализа авторы приходят к выводу, что в ближайшие десятилетия деловая среда цифрового общества будет определяться стратегией обработки больших данных, направленной на обеспечение кибербезопасности, основанной на глубоком понимании человеческого поведения и искусственного интеллекта. Данный анализ может быть использован техническими специалистами для обоснования новых технологических решений по внедрению ИТ и больших данных в цифровом обществе, а также для подтверждения обоснованности предложений по повышению кибербезопасности при массовой обработке больших данных в прикладных целях.

Ключевые слова: большие данные, цифровые ресурсы, базы данных, библиометрический анализ, информационные технологии, интеллектуальный анализ данных.

BIG DATA TECHNOLOGIES: AN OVERVIEW

Abstract. Trends in the IT sector are constantly evolving, and the ability to anticipate new developments is a key factor in an organization's success in managing digital resources, including big data. The aim of this study is to explore and analyze big data in order to ensure security in real time.

As a result of the bibliometric analysis, 323 scientific articles were obtained from 137 information sources containing the keywords relevant to the study. The research included journal articles, book chapters, and patents. Scientific publications were retrieved from the Scopus and WoS databases using primary and secondary keywords. The study employed the following methods: (a) a systematic literature review based on publications from 2013 to 2023, and (b) a bibliometric analysis of articles published from 2000 to 2023 using the RStudio and Bibliometrix

software.

Based on the analysis, the authors conclude that in the coming decades, the business landscape of the digital society will be shaped by a strategy focused on processing big data to ensure cybersecurity, grounded in a deep understanding of human behavior and artificial intelligence. This analysis can be used by technical professionals to justify new technological solutions for implementing IT and big data in a digital society, as well as to substantiate recommendations for improving cybersecurity during large-scale processing of big data for applied purposes.

Keywords: big data, digital resources, databases, bibliometric analysis, information technology, data mining.

Авторлар туралы мәліметтер

Кубигенова Акку Тулегеновна	Техника ғылымдарының магистрі, С.Сейфуллин атындағы Қазақ агротехникалық зерттеу университетінің докторанты, Астана қ., Қазақстан E-mail: akku_kubigenova@mail.ru *
Актаева Алимбуби Умирбековна	PhD, А.Мырзахметов атындағы Көкшетау университетінің «Ақпараттық жүйелер» кафедрасының қауымдастырылған профессоры, доцент, Көкшетау қ., Қазақстан, E-mail: aaktaewa@list.ru
Шәріпбай Алтынбек Әмірұлы	Техника ғылымдарының докторы, Л.Н. Гумилев атындағы Еуразиялық ұлттық университетінің «Жасанды интеллект технологиясы» кафедрасының профессоры, Астана қ., Қазақстан, E-mail: sharalt@mail.ru
Сухомлин Владимир Александрович	Техника ғылымдарының докторы, М.В.Ломоносов атындағы Мәскеу мемлекеттік университетінің «Есептеу математикасы және кибернетика» факультетінің «Ашық ақпараттық технологиялар» зертханасының меңгерушісі, профессор, Мәскеу қ., Ресей Федерациясы E-mail: sukhomlin@mail.ru
Молдашева Раушан Нуркожаевна	PhD, Х. Досмұхамедов атындағы Атырау университеті, «Бағдарламалық инженерия» кафедрасының қауымдастырылған профессоры, Атырау қ., Қазақстан, E-mail: raushan85_07@mail.ru

Сведение об авторах

Кубигенова Акку Тулегеновна	Магистр технических наук, докторант Казахского агротехнического исследовательского университета имени С.Сейфуллина, г. Астана, Казахстан E-mail: akku_kubigenova@mail.ru
Актаева Алимбуби Умирбековна	PhD, Ассоциированный профессор кафедры «Информационные системы» Кокшетауского университета им. А.Мырзахметова, г. Кокшетау, Казахстан E-mail: aaktaewa@list.ru
Шәріпбай Алтынбек Әмірұлы	Доктор технических наук, профессор кафедры «Технология искусственного интеллекта» Евразийского Национального Университета им. Л.Н. Гумилева. E-mail: sharalt@mail.ru
Сухомлин Владимир Александрович	Доктор технических наук, зав. лабораторией «Открытые информационные технологии» факультета ВМК МГУ им.М.В.Ломоносова, профессор, г.Москва, Российская Федерация, E-mail: sukhomlin@mail.ru
Молдашева Раушан Нуркожаевна	PhD, Ассоциированный профессор кафедры "Программная инженерия" Атырауского государственного университета им. Х. Досмұхамедова, г. Атырау, Казахстан E-mail: raushan85_07@mail.ru

Information about authors

Kubigenova Akku	Master of Technical Sciences, Doctoral student of Kazakh Agrotechnical Research University named after S. Seifullin, Astana, Kazakhstan, E-mail: akku_kubigenova@mail.ru
Aktaeva Alimbubi	PhD, Associate Professor of the Department of Information Systems, Kokshetau University named after A. Myrzakhmetov, Kokshetau, Kazakhstan, E-mail: aaktaewa@list.ru
Sharipbay Altynbek	Doctor of Technical Sciences, Professor, Department of Artificial Intelligence Technology, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan, E-mail: sharalt@mail.ru
Sukhomlin Vladimir	Doctor of Technical Sciences, Professor, Head of the Laboratory of Open Information Technologies, Faculty of Mechanical Engineering, Lomonosov Moscow State University, Moscow, Russia, E-mail: sukhomlin@mail.ru
Moldasheva Raushan	PhD, Associate Professor of the Department of Software Engineering, Kh. Dosmukhamedov Atyrau University, Atyrau, Kazakhstan, E-mail: raushan85_07@mail.ru