DOI 10.51885/1561-4212_2025_3_194
ҒТАХР 81.93.29

КОНФИДЕНЦИАЛДЫҒЫ САҚТАЛҒАН ДЕРЕКТЕР БАЗАСЫНА СҰРАНЫСТАР ЖӘНЕ ДЕРЕКТЕРДІ ТАЛДАУ

ЗАПРОСЫ К БАЗАМ ДАННЫХ С СОХРАНЕНИЕМ КОНФИДЕНЦИАЛЬНОСТИ И АНАЛИЗ ДАННЫХ

PRIVACY-PRESERVING DATABASE QUERIES AND DATA ANALYSIS

Ж.З. Жантасова [1*], С.А. Нургалиева [2], А.Р. Сыздыкпаева [1]

[1]С. Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен қ., Казақстан

[2]Astana IT University, Астана қ., Қазақстан

*Жауапты автор: Жантасова Женискуль Зейнешовна, e-mail: zheniskul_z@mail.ru

Түйінді сөздер:
сұраныстарды өңдеу, сұраныстар конфиденциалдығы, жүйелі шолу, дербестендірілген деректер, бұлттық қойма, шифрлеу, дифференциалды конфиденциалдылық.

ТҮЙІНДЕМЕ

Бұлтты қоймаларының деректерінің және үлестірілген сыртқы деректер көздерінің өсуі деректерге қол жеткізудің конфиденциалдығын қамтамасыз ету мәселесін күшейтеді. Бұл, ең алдымен, пайдаланушылар туралы дербес ақпарат өңдейтін мемлекеттік, қаржылық және медициналық ақпараттық жүйелерге қатысты. Демек, жүйелердегі конфиденциалдықтың бұзылуы елеулі құқықтық және экономикалық салдарға, сондай-ақ құпия деректердің таралуына әкелуі мүмкін.

Бұл мақалада дифференциалды конфиденциалдылыққа, k-ең жақын көршілес (kNN) сұраныстарды өңдеу алгоритмдері және криптографиялық хаттамаларға қатысты негізгі түсініктер талқыланады. Блокчейн технологияларының, онтологиялық модельдердің сұраныстар құпиялылығының мәселелерімен интеграциясы туралы зерттеулер анықталды. Конфиденциал-дылықты ескере отырып, сұраныстарды оңтайландырудың қолданыстағы тәсілдеріне талдау жүргізілді, негізгі тенденциялар мен перспективалық зерттеу бағыттары анықталды.

Жүйелі шолу әдістемесі PRISMA стандартына сәйкес келеді. Зерттеу мақсаттары мен ерекшелігін бейнелейтін релевантты кілт сөздерді пайдалана отырып, Web of Science және ResearchRabbit дерекқор-ларында жан-жақты іздеу жүргізілді. Алынған нәтижелер осы саладағы зерттеулердің ағымдағы жай-күйін кең түсінуге мүмкіндік береді және деректерді талдауда сұраныстарды өңдеудің тиімді және қауіпсіз механизмдерін әзірлеуде пайдалы болуы мүмкін.

Ключевые слова:
обработка запросов, конфиденциальность запросов, систематический обзор, персонализированные данные, облачное хранилище, шифрование, дифференциальная конфиденциальность.

АННОТАЦИЯ

Увеличение объема данных облачных хранилищ и распределенных внешних источников данных обостряет проблему обеспечения конфиденциальности доступа к данным. Прежде всего, это касается правительственных, финансовых и медицинских информационных систем, где происходит обработка персонализированных сведений о пользователях. Нарушение конфиденциальности в этих системах может привести к серьезным юридическим и экономическим последствиям, а также к утечкам чувствительных данных.

В работе рассматриваются ключевые концепции, затрагивающие дифференциальную конфиденциальность, алгоритмы обработки запросов kNN (k Nearest Neighbor) - k ближайших соседей и криптографические протоколы. Проведен анализ существующих подходов к оптимизации запросов с учетом конфиденциальности, выявлены основные тенденции и определены перспективные направления исследований.
Методология систематического обзора соответствует стандарту PRISMA. Всесторонний поиск проводился в базах данных Web of Science и ResearchRabbit с использованием релевантных ключевых слов, отражающих специфику и цели исследования. Полученные результаты позволяют лучше понимать текущее состояние исследований в данной области и могут быть полезны при разработке более эффективных и безопасных механизмов обработки запросов с последующим анализом данных.

Keywords:
query processing, query privacy, systematic review, personalized data, cloud storage, encryption, differential privacy.

ABSTRACT
The increasing volume of cloud storage data and distributed external data sources exacerbates the problem of ensuring confidentiality of access to data. First of all, this concerns government, financial and medical information systems, where personalized information about users is processed. Breaches of confidentiality in such systems can lead to serious legal and economic consequences, as well as the leakage of sensitive data.
The paper considers key concepts affecting differential privacy, kNN (k Nearest Neighbor) query processing algorithms and cryptographic protocols. Studies on the integration of blockchain technologies, ontology models with query confidentiality issues are identified. An analysis of existing approaches to query optimization taking into account confidentiality is carried out, the main trends are identified and promising research areas are determined.
The methodology of the systematic review complies with the PRISMA standard. A comprehensive search was conducted in the Web of Science and ResearchRabbit databases using relevant keywords reflecting the specifics and objectives of the study. The results obtained allow you better understand the current state of research in this area and can be useful in developing more efficient and secure mechanisms for processing queries with subsequent data analysis.

КІРІСПЕ

Бұлтты қоймалары және таратылған деректер көлемінің аса көбеюіне байланысты ақпараттың құпиялылығын қамтамасыз ету мәселесі маңызды бола түсуде. Бұл әсіресе пайдаланушының құпия ақпаратын өңдейтін қаржылық, медициналық және мемлекеттік ұйымдар үшін өте маңызды. Деректер базасына сұраныстарды орындау кезінде деректердің құпиялылығын қамтамасыз ету деректерді талдау саласындағы негізгі мәселелердің бірі. Шифрлау және жасыру сияқты дәстүрлі қорғау механизмдері қол жеткізу үлгілерін талдауға, сырт арналық шабуылдарға және бастапқы деректерді қайта құруға кедергі бола алмайды. Бұл конфиденциалдықты сақтай отырып, сұрауларды өңдеудің жаңа әдістерін әзірлеуді қажет етеді.

Осы шолу мақаланың негізгі мақсаты – конфиденциалдықты қамтамасыз етумен қатар деректер базасына сұраныстарды оңтайландыру бойынша бар зерттеулерді жүйелеу, негізгі үрдістерді айқындау және болашақ зерттеулердің перспективалық бағыттарын табу. Зерттеу мақсатына жету үшін авторлар келесі міндеттерді белгіледі:

– конфиденциалдықты қамтамасыз ете отырып, деректер базасына сұраныстарды өңдеудің негізгі терминдерін, үлгілерін және әдістерін нақтылау;
– қол жеткізілген нәтижелерді сипаттап зерттеу бағыттарын жүйелеу;
– таңдалған басылымдарға жүйелі шолу және библиометриялық талдау жасау;
– зерттеушілер ұсынған шешімдердің осалдылықтарын анықтау;
– сұраныстарды оңтайландырудың трендтерін болжау.

Соңғы жылдары сұраныстардың құпиялылығын қамтамасыз етудің көптеген әдістері ұсынылды. Мысалы, конфиденциалдық талаптарын ескере отырып, тұпнұсқалы деректерге жақын синтетикалық деректер базаларын (SD-Synthetic Datasets) жасау сұраныстардың өнімділік сипаттамаларын сақтауға мүмкіндік береді (Seeman, Sexton және т.б., 2024; Ghazarian et al., 2024; Xia et al., 2020). Дегенмен, SD пайдалану бірқатар мәселелермен байланысты: ақпараттың жайылып кетуі, деректерді қалпына келтіру әдістерімен шабуылдау, пайдалылық пен қауіпсіздік арасындағы тепе-теңдікті сақтау. Дифференциалды конфиденциалдық (Differential Privacy, DP) технологиялары сұраныс нәтижелеріне шу қосу арқылы деректердің жайылып кетуін азайтады, бірақ дәлдік пен қауіпсіздік арасында ымыраға келуді талап етеді (Przytarski және т.б., 2021; Ekaputra F.J. және т.б., 2024; Jang және т.б., 2015). Шифрланған деректер базасындағы қорғалған екі жақты есептеулерге негізделген конфиденциалдықты сақтайтын kNN сұраныстарды өңдеу алгоритмдері белсенді түрде зерттелуде (Ghazarian et al., 2024; Akgün, Pfeifer, Kohlbacher, 2022). Дегенмен, кейбір алгоритмдер шабуылдарға осал, ал басқалары сұраныстарды өңдеудің жоғары шығындарынан зардап шегеді.

Бұл саладағы нәтижелі өзгерістерге қарамастан, келесі мәселелердің шешімдерін табу керек:

– Масштабталатын (кеңейтілетін) жүйелерде қорғалған сұраныстарды орындау кезіндегі жоғары шығындар.
– Сұраныстар үлгілерін талдауда көптеген схемалардың осалдығы.
– Қорғау әдістерін пайдалану кезінде күрделі аналитикалық операцияларды жеткіліксіз қамтамасыз ету.
– Таратылған деректер базасында конфиденциалдықты қамтамасыз етудің бірыңғай стандарттарының болмауы.

ЗЕРТТЕУ МАТЕРИАЛДАРЫ МЕН ӘДІСТЕРІ

Web of Science және ResearchRabbit дерекқорларында зерттеудің ерекшеліктерін және осы шолудың мақсатына жетуге бағыттылығын көрсететін түйінді сөздерді пайдалану арқылы кешенді іздеу жүргізілді. Іздеу нәтижесінде 269 мақала табылды. Алдын ала іріктеу кезеңінде Web of Science деректер қорынан 220 жазба және ResearchRabbit дерекқорынан 49 жазба жүктелді. Microsoft Excel және Zotero көмегімен телнұсқалардың бар-жоғын тексергеннен кейін 9 жазба алынып тасталды, себебі олар әртүрлі көздерде қайталанды. Жазбалардың біртипті еместігіне байланысты 4 мақала қолмен алынып тасталды. Microsoft Excel бағдарламасында телнұсқалар 2 кезеңде жойылды: 1) «Author», «Article Title» критерийлерінің сәйкестігі бойынша; 2) « Author», «DOI».

Алғашқы іріктеу (screening) кезеңінде тақырыптар мен түйіндемелер талданды. Бұл талдау 250 релеванттық, зерттеуге қатысты жұмыстарды қалдыруға мүмкіндік берді. Содан кейін қолжетімділігі жабық мақалаларды алып тастау керек болды. Осы сүзгіден кейін 76 мақала қалды. Келесі кезекте критерийлерге сәйкестігін бағалау жүргізілді: жанама сәйкестік – 12 және басқа да себептер – 28: мысалы, қарастырылған қауіпсіздік мәселелері сұраныстарға қатысты емес, мақалалардың басылым мерзімі сәл ескірген (2003-2007 жж.) және т.б. Бұл кезеңде таңдалған дереккөздер саны 36 мақалалардан жинақталды.

Әрбір басылымды таңдау PRISMA әдістемесіне сәйкес жүргізілді (Yulianti және т.б., 2024). Ғылыми қатаңдығы және дәлелді базасы жоғары жұмыстар сенімді деп танылды. Іріктелген зерттеулердің шынайлығын қамтамасыз ету үшін, осы жұмыстың авторлары зерттелетін тақырыптың мазмұнына нақты талаптарға сәйкес қол жеткізуге кепілдік беру сұрақтарын қосымша талқылады. Нәтижесінде, қосымша критерийлер тұжырымдалып келесі факторлар назарға алынды: зерттеу мақсатына жетуде дәйекті ілгерілеуді нақты қадағалау, іріктеменің репрезентаттылығы, зерттеу кезеңдерін логикалық дәйекті түрде ұсыну, шынайы статистикалық мәліметтердің болуы, нәтижелердің қайта пайдалану мүмкіндігі. Авторлар арасында айырмашылықтар болған жағдайда даулы мәселелер бойынша пікірталас және конструктивті келісім ұйымдастырылды. Орындалған жұмыстарға осы сапаны бақылау процедурасын енгізу 22 басылымды алып тастап егжей-тегжейлі талдау жүргізуге мүмкіндік берді. Әрі қарай қарастыру үшін алып тастау критерийлерінің қатарында ұсынылған шешімдердің жергілікті (шектеулі) қолданылуы, қол жеткізілген нәтижелерді сыни тұрғыдан пайымдаудың жеткіліксіздігі және зерттелетін тақырыптың даму болашағына әлсіз көзқарас жатады. Қалған 14 ғылыми жарияланымдар зерттелетін мәселелік критерийлеріне сәйкес келеді, өзектілігі болып жатқан ақпараттық процестерге (бұлтты технологиялардың өсуі, ақпараттың жайылып кету мәселелерінің шиеленісуі, деректер көлемінің өсуі және т.б.) байланысты анық, зерттелетін мәселелердің негізгі бағыттары қадағаланады. Бұл мақалалар зерттеу тақырыбына жүйелі шолу жасау үшін талданған.

1-сурет. Prisma диаграммасы бойынша мақалаларды іріктеп алу процедурасы

Ескерту: авторлар (Yulianti және т.б.., 2024) деректері негізінде құрастырған

НӘТИЖЕЛЕР ЖӘНЕ ОЛАРДЫ ТАЛҚЫЛАУ

Бұл бөлімде жүйелі шолу нәтижелері түсіндірілген және болашақ зерттеулер мен практикалық қолдану үшін қорытындылар тұжырымдалған. Авторлар өз зерттеулерін сипаттау үшін пайдаланған негізгі терминдердің анықтамалары (1-кесте) қазіргі уақытта нақты ақпараттық жүйелерде жүріп жатқан сұраныстарды өңдеу процестерін көрсетті. Деректерді бұлттық қоймада немесе таратылған жүйелерде сақтау және өңдеу қосымша қауіпсіздік шараларын қажет етеді. Мақалалар авторларына мақсатталған нәтижелерге қол жеткізуге мүмкіндік беретін ең жиі аталған әдістер мен технологиялардың ішінде дифференциалды конфиденциалдық әдісін, аутентификация және сұраныстарды өңдеу үшін шифрлау алгоритмдерін пайдалану (Secure Multi-Party Computation SMPC), k-ең жақын көршілестерді өңдеудің kNN алгоритмдерін атауға болады. Ең осал деректер көздері медициналық және жекелендірілген қызметтер болып табылады (Przytarski және т.б., 2021; Cunningham, Cormode және Ferhatosmanoglu, 2021; de Souza, de Lassus Cammarota, 2024). Біріктірілген сұраныс құрылымдарына негізделген конфиденциалдықты сақтайтын модельді келесі авторлар (Riyana, Sasujit, Homdoung, 2023; Ekaputra F.J. et al., 2024; Yang et al., 2022; Zheng, 2022) ұсынады. Ұсынылған шешімдердің ерекшелігі конфиденциалдықты жоғалту мен деректердің пайдалылығы арасындағы қол жеткізілген ымыраға келу болып табылады. Жеке меншік деректерді қорғау әдісі ретінде дифференциалды конфиденциалдық (ДК) ақпараттың жайылып кетуін шектеуді қамтамасыз етеді (Przytarski және т.б., 2021; Yang, Kelarev және Yi, 2024; Yang, 2022). Келесі авторлардың (Xia және т.б., 2020; Jang және т.б., 2015) жұмыстарында кеңістіктік деректердің сыни аймағында дифференциалдық конфиденциалдықпен синтетикалық деректерді генерациялау шешімі және олардың талдауы ұсынылады.

Бұлттық есептеулердің негізгі қызметтерінің бірі пайдаланушылардың үлкен көлем деректерін тасымалдауға және сақтауға мүмкіндік беретін аутсорсинг болып табылады. Дегенмен, деректер бұлтта шифрланбаған түрде сақталса, конфиденциалдық ақпараттың жайылып кету қаупі жоғары. Бұл әсіресе бизнес деректері мен медициналық жазбалар үшін өте маңызды. Көптеген бұлттық провайдерлер пайдаланушылар деректерін қарамайтынына немесе өзгертпейтініне кепілдік бермейді. Сонымен қатар, бұлттық серверлерге қаскүнемдер шабуыл жасауы мүмкін немесе провайдерлер ресурстарына жүктемені азайту үшін сұраныс нәтижелерін әдейі жасыруы немесе ауыстыруы мүмкін (Zheng және т.б., 2022; Jang және т.б., 2015). Шешім ретінде авторлар деректерді бұлтқа тасымалдау кезінде іздеуге қабілетті симметриялық шифрлауды SSE ((Searchable Symmetric Encryption)) қарастырады. Ұсынылған шифрлау схемасы B+ ағашына және Блум сүзгісіне негізделген. Сонымен бірге бұлттық серверден шабуылдардың алдын алу үшін уақыт белгілер механизмі ұсынылады. Белгілі іздеу үлгілерін кеңейтетін практикалық криптографиялық модельді келесі авторлар (Crescenzo және т.б., 2016) ұсынады.

Деректерді бұлтқа жібермес бұрын, авторлар (Ghazarian, Zheng, Rakovski, 2024) k-ең жақын көршілердің сұраныстарын өңдеу kNN алгоритмін пайдалануды ұсынады. Конфиденциалдық пен тұтастықты қолдау үшін сұраныс аутентификациясы үшін бұл алгоритмді келесі авторлар да (Akgün, Pfeifer, Kohlbacher, 2022; Elmehdwi және т.б., 2013) ұсынады. Аутентификация схемасы биттік карталар негізінде жасалған, хэштеу индексі сұранысты биттік форматына түрлендіреді және сәйкес тірек нүктелері туралы ақпаратты тиімді шығаруға мүмкіндік береді. Блокчейнді пайдалану және оның сұраныстарды өңдеуге қойылатын талаптары сұраныс тиімділігін арттыру контекстінде талқыланады (Kim және т.б., 2022). Конфиденциалдықты сақтай отырып, тексерілетін деректерді талдауға қол жеткізу қиын есеп болып табылады, өйткені бар әдістер мен

құралдар аудит журналдарын жүргізу, деректерді пайдалану саясаттарын автоматты түрде тексеру немесе деректерді анонимизациялау (жасыру) сияқты жартылай шешімдерді ғана ұсынады. Келесі авторлар (Seeman, Sexton және т.б., 2024) жұмысында аудитті қамтамасыз ету үшін автоматтандырылған режимде бір уақытта жеткілікті ақпаратты жинау кезінде конфиденциалдықты сақтайтын деректерді талдауға мүмкіндік беретін семантикалық қолдау көрсететін WellFort архитектурасын ұсынады. Алынған нәтижелер сұраныс жасалған деректердің конфиденциалдығын қамтамасыз ете отырып, әртүрлі деректер базасы сұраныстарын интеграциялау сұрақтарына ғалымдардың шоғырланған назарын көрсетеді.

Зерттеудің негізгі терминдері. Басылымдарды қарау, негізгі түсініктерді белгілеу, зерттеушілер қарастырған есептердің контекстін салыстыру жиі кездесетін түсініктерді және олардың анықтамаларын іріктеп алуға мүмкіндік берді (1-кесте). Негізгі терминдер қатарына объекттер, субъекттермен қатар үрдістер, әдістер және қауіпсіздік саясат атаулары (Kazimova және т.б., 2025) енгізілген.

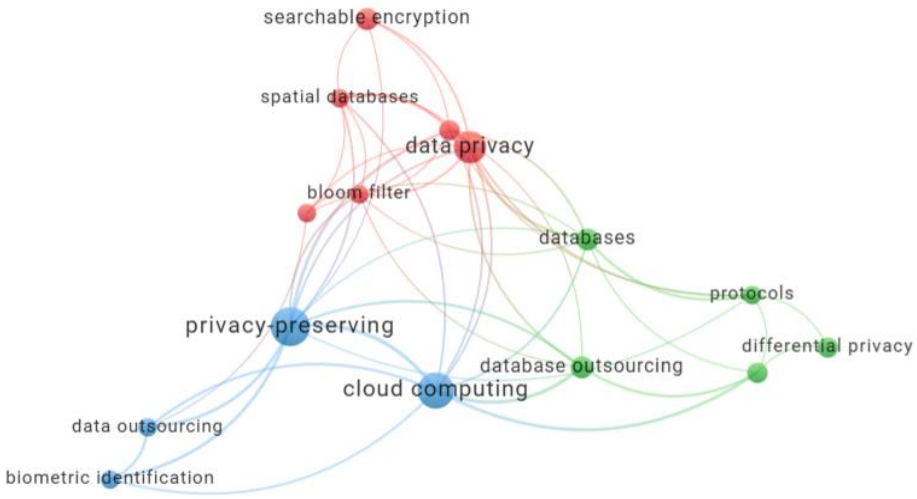
1-кесте. Зерттеу тақырыбының негізгі түсініктері

Негізгі термин	Анықтама
Сұраныстардың агрегатталған (біріктірілген) құрылымдары (Aggregate query frameworks)	Деректердің конфиденциалдығын қорғау үшін қолданылатын құрылымдар (Riyana, Sasujit, Homdoung, 2023; Yang, Kelarev,Yi, 2024; Yang және т.б., 2022).
Верификациялануы (Auditability)	Деректерді өңдеу үрдістерін тексеру және қадағалау. Верификациялану қасиеті аналитикалық жүйелерге және нормативті талаптардың орындалуына сенімді жоғарылатады (Seeman,Sexton және т.б., 2024).
Сұраныстарды өңдеу блокчейні (Blockchain for Query Processing)	Сұраныстарды қорғау және аутентификациялауға қолданылатын блокчейн. Сұраныстарды өңдеуде деректердің өзгермеуін және ашықтығын қамтамасыз етеді. (Kim және т.б., 2022; Ghazarian, Zheng, Rakovski, 2024; Papadopoulos, 2020).
Келісіммен бақару (Consent Management)	Пайдаланушылардың деректерін өңдеуде бекітілген тәртіптерге (мысалы GDPR регламенті) сәйкес рұқсатын басқару, әсіресе медициналық және жекешелендірілген сервистерде. (Seeman,Sexton және т.б., 2024).
Деректер базасының аутсорсингісі (Database Outsourcing)	Аутсорсинг – компания жұмысының бір бөлігінің өзге компанияға тапсыру. Сыртқы серверлерде немесе бұлттық платформаларда деректер базасын сақтау және басқару тәжірибесі (Elmehdwi,2013; Chaudhari, 2024).
Дифференциалдық конфиденциалдық (Differential Privacy)	Деректердің конфиденциалдығын сұраныстардың нәтижесіне шу қосу арқылы қорғау. Әдіс нәтижесінде жазбаларды анықтауға тосқауыл қойылады. (Przytarski және т.б., 2021; (Seeman,Sexton және т.б., 2024; Ekaputra, 2024; Ghazarian, Zheng, Rakovski, 2024; Xia, 2020; Zhang және т.б., 2024; Scaglione, 2022).

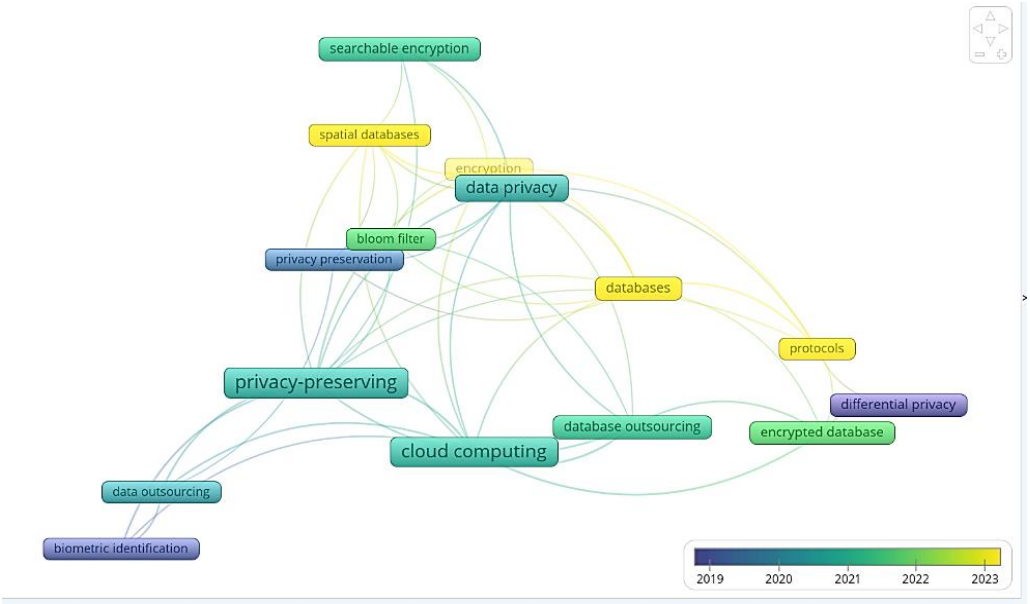
1-кестенің соңы

Негізгі термин	Анықтама
Шифрланған сұраныстарды өңдеу (Encrypted Query Processing)	Шифрланған деректердің үстінен сұраныстарды өңдеу бұлттағы деректердің мазмұнын үшінші жаққа ашпай оларды іздеуге және талдауға мүмкіндік береді (Kim және т.б., 2022; Shi, 2020; Cunningham, Cormode, Ferhatosmanoglu, 2021; Akgün, Pfeifer, Kohlbacher, 2022; Zheng, 2022; Chaudhari, 2024).
Деректерді қорғаудың жалпы ережелері GDPR (General Data Protection Regulation)	Интернет желісінде Еуроодақтың пайдаланушыларының деректерін жинақтап өңдейтін компанияларға біріктірілген жарлығы (Seeman,Sexton және т.б., 2024).
Конфиденциалдықты сақтайтын деректердің талдауы (Privacy-Preserving Data Analysis)	Пайдаланушылардың жеке ақпаратын қорғауды қамтамасыз ету әдістерін қамтитын конфиденциалдықты сақтайтын деректерді талдау (Riyana, Sasujit, Homdoung, 2023; Przytarski және т.б., 2021; Cunha, Mendes,Vilela, 2021; Welten және т.б., 2022; Scaglione, 2022).
Конфиденциалдықты сақтайтын сұраныстар (Privacy-Preserving Queries)	Бұлтта немесе таратылған жүйелерде сұраныстар өңделген кезде пайдаланушылар деректерінің қорғалуын қамтамасыз ететін сұраныстар (Ghazarian, Zheng, Rakovski, 2024; Crescenzo және т.б., 2016; Xia және т.б., 2020; Jang және т.б., 2015; Akgün, Pfeifer, Kohlbacher, 2022).
Конфиденциалдықты сақтайтын kNN сұраныстарды өңдеу (Privacy-Preserving kNN Query Processing)	Сұраныстарды ширлау негізінде конфиденциалдықты қамтамасыз ететін және қатынау паттерндер арқылы ақпараттың жайылып кетуін алдын алатын k-жақын көршілерді өңдейтін алгоритмдер (kNN) (Ghazarian, Zheng, Rakovski, 2024).
Сұраныстың аутентификациясы және тұтастықтың верификациясы (Query Authentication and Integrity Verification)	Деректер базасы және бұлттық қоймаларды қорғау үшін қолданылатын сұраныстар нәтижелерін бұрмалауын алдын алуға бағытталған сұраныстарды аутентификациялау және деректердің тұтастығын тексеру әдістері (Shi және т.б., 2020; Akgün, Pfeifer, Kohlbacher, 2022; Elmehdwi және т.б., 2013).
Қауіпсіз көптарапты есептеулер (Secure Multi-Party Computation - SMPC)	Бірнеше тараптарға өз деректерін бір біріне ашпай бірігіп функциялар есептеуге мүмкіндік беретін криптографиялық әдіс (Seeman, Sexton және т.б., 2024; Shi және т.б., 2020; Ghazarian, Zheng, Rakovski, 2024; Cunningham және т.б., 2021; Jang және т.б., 2015).
Ескерту – Авторлар әдебиеттер көздерін талдау негізінде құрастырған	

Vosviewer бағдарламасы библиометриялық желілерді визуализациялау үшін кеңінен қолданылды. VOSviewer бағдарламасының көмегімен мәліметтерді талдаудың негізгі әдісі ретінде «co-occurrence» әдісі таңдалды. Ол кілт сөздерді олардың бір жұмыста қаншалықты жиі кездесетіндігі бойынша кластерлейді. Осылайша, түйінді сөздер тақырыптық кластерлерді құрайды (2-сурет). «Overlay visualization» визуализация құралы тақырыптық кластерлер бойынша жарияланымдар хронологиясын көрсетті (3-сурет).



2-сурет. Зерттеудің түйінді терминдері
Ескерту – Авторлар Vosviewer бағдарламасы көмегімен тұрғызды



3-сурет. Зерттеудің түйінді терминдері
Ескерту – Авторлар Vosviewer бағдарламасында – Overlay visualization құралы арқылы тұрғызды

Басылымдар авторларының жүргізген зерттеулері, негізгі алынған нәтижелері, қолданылған үлгілері мен құралдары зерттеу тақырыбындағы негізгі бағыттарды, қолданылатын тәсілдерді анықтады. Назарда болған зерттеудің көрсеткіштерінің ақиқаттық дәрежесін растау үшін анықталған негізгі бағыттар бойынша мақала авторлары қосымша басылымдарды қарап шықты. Бұл жұмыстың нәтижелері де талқылауда және берілген ұсыныстарда өз орнын тапты.

ҚОРЫТЫНДЫ

Бұл мақалада деректер базасына конфиденциалдық сұрақтарын назарға алып сұраныстарды тиімділеу шешімдері қарастырылған басылымдарды талдау әрекеттері жасалды. Қарастырылған мәселе бойынша іріктеу, жалпы даму тенденцияларды табу

жолында өзекті аспектілер, қолданылатын әдіс-тәсілдер мен құралдар айқындалды. Сонымен қатар, ұқсас зерттеулерді топтастырып ірілендірілген шешімдерді іздестіру бағыттарды қалыптастыру сәтті болды. Қойылған есептің спецификасына байланысты әдістің жүзеге асырылуындағы айырмашылықтар көрсетілді. Осы мақала авторлары қол жеткізген нәтижелері қорғалған деректер базасы аймағында жүргізілген зерттеулермен расталған және келісілген. Сондай-ақ, алынған нәтижелердің құндылығы дәйекті басқарушылық шешімдер қабылдау мақсатында деректердің интеллектуалдық талдау потенциалы болуында. Орындалатын сұраныстардың конфиденциалдығын қамтамасыз етуде зерттеушілік жобалар тақырыптары бола алатын үш негізгі бағыттарын атап өтуге болады. Ең алдымен, деректерді бұлтта сақтауда олардың конфиденциалдығын, тұтастығын және өзектілігін қамтамасыз ету. Екінші негізгі даму бағыты деректер базасына жасалатын сұраныстардың қауіпсіздігіне кепілдік бере алатын криптографиялық үлгілер мен хаттамалар байланысты болып келеді. Дифференциалдық конфиденциалдық механизмдері арқылы орналасқан жері туралы синтетикалық деректерді генерациялау әдістерін алу үшінші бағыттың негізі болып келеді. Сонымен қатар, авторлардың пікірінше, гибридтік жүйелерді қарастыру қызық та орынды болып көрінді. Олар блокчейн мен деректер базасы функцияларын интеграциялау арқылы деректердің тұтастығын қамтамасыз ете алады және мүлдем жаңа сақтау құрылымдарына әкеледі. Деректердің онтологиялық көрсетілімін беретін семантикалық торларды қарастыру жаңа шешімдерге әкеледі деген қызығушылық тудырды. Мақала авторларының қол жеткізген нәтижелері ғылыми қоғамдастықтың жинақталған тәжірибесінің жүйеленген форматы болып келеді. Әсіресе, медициналық деректерге қатысты қауіпсіздік мәселелерді шешуде қолданылатын әдіс-тәсілдердің байланыстары, жүзеге асырылу айырмашылықтары туралы деректер. Деректер базасында күннен күнге ақпарат көлемінің көбеюуі, талдауға және шешім қабылдауға қажетті түрлі деректерді жылдам шығарып алу қажеттілігі, тасымалданатын ақпараттық ағымдардың тұтатсығын қамтамасыз ету қолданылатын қорғау процедураларының маңыздылығын күшейте түседі.

Алынған деректердің маңыздылығына қарамастан, басылымдар авторлары қол жеткізген нәтижелерге қатысты белгілі бір шектеулерді атап өткендері ескертіледі. Ұсынылған шешімдердің барлығы дерлік оңтайлы шешімге кедергі болатын мәселелерді сыни тұрғыдан түсінуді көрсеткен. Көптеген шешімдер бұлттық серверлер «адал (честный)», бірақ «қызығушылық (любопытный)» таныта алады шартымен берілген. Конфиденциалдық пен тиімділікті сақтай отырып, бұл нәтижелерді зиянды қаскүнемдер сияқты күшті қарсылас модельдерге кеңейту бүгінгі күнде қиын есеп болып келеді. Масштабтау сұрақтары да орын алған. Ұсынылған схемалардың көпкөлемді және әртүрлі деректер жиындарына масштабталатындығы толық анықталмаған. Көп өлшемді деректермен жұмыс істеуде енгізу/шығару кідірістерінің мәселелері болғандығы және де деректер өлшемін азайту әдістері қолданылғаны туралы деректер келтірілген. Есептеулердің жоғары бағасы да ұсынылған шешімдердің қолдану аймағын кеңейтуді шектейді.

Зерттеу тақырыбына сәйкес қарастырылған еңбектер көлемі мақала авторларының қорытындыларының толық болуына әсер етуін ескеру керек. Жабық қолжетімдіктегі басылымдардың атаулары мен түйіндемелері қарастырылғаны мен шектелді. Және де ертерек уақытта басылымнан жарық көрген ғылыми енбектердің қарастырылмауы заманауи зертеулердің назарынан маңызды нәтижелердің тыс қалу ықтималы болуы мүмкін.

Әрі қарайғы зерттеулер қазіргі деректер қоймаларына тән құрылымдалмаған деректерге бағытталуы керек екендігі аталып отыр. Сонымен қатар, деректерді интеллектуалды талдау, предиктивтік аналитика және деректерді визуализациялау

заманауи қажеттіліктері тиімді шешімдерді алуда жасанды интеллекттің өсіп келе жатқан рөлін көрсетеді. Іргелі зерттеулерді жетілдіру тұрғысынан математикалық модельдеудің классикалық әдістерін белгілі бір мәселенің қойылымына сәйкес қарастыру өзекті болып табылатыны сөзсіз. Қажетті қауіпсіздік пен пайдаланылған деректердің нақты пайдалылығы арасындағы компромисстік шешімдер күтілетін нәтижелерге өзгертулерде жасай алады.

МҮДДЕЛЕР ҚАЙШЫЛЫҒЫ: Авторлар мүдделер қайшылығы жоқ екенін мәлімдейді.

ӘДЕБИЕТТЕР ТІЗІМІ

Akgün M., Pfeifer N., Kohlbacher O. (2022). Efficient privacy-preserving whole-genome variant queries., Bioinformatics, vol. 38, no. 8, pp. 2202–2210, https://doi.org/10.1093/bioinformatics/btac070.

Chaudhari P., Chin J.-J., Mohamad S.M. (2024). An In-Depth Analysis on Efficiency and Vulnerabilities on a Cloud-Based Searchable Symmetric Encryption Solution., J. Inform. Web Eng., https://doi.org/10.33093/jiwe.2024.3.1.19.

Crescenzo G.Di., Cook D.L., McIntosh A., Panagos E. (2016). Practical and privacy-preserving information retrieval from a database table, J. Comput. Secur., i. 4, vol. 24, pp. 479–506, https://doi.org/10.3233/JCS-160550.

Cunha M., Mendes R., Vilela J. P. (2021). A survey of privacy-preserving mechanisms for heterogeneous data types, Comput. Sci. Rev., vol. 41, p. 100403, https://doi.org/10.1016/j.cosrev.2021.100403

Cunningham T., Cormode G., Ferhatosmanoglu H. (2021). Privacy-Preserving Synthetic Location Data in the Real World. 17th International Symposium on Spatial and Temporal Databases, pp. 23–33. https//doi.org/10.1145/3469830.3470893.

De Souza, F.D.M., de Lassus, H. , Cammarota, R. (2024). Private detection of relatives in forensic genomics using homomorphic encryption. BMC Med Genomics 17, 273. https://doi.org/10.1186/s12920-024-02037-9

Ekaputra F.J. et al. (2024). Semantic-enabled architecture for auditable privacy-preserving data analysis, Semantic Web, т. 15, vol. 3, pp. 675–708, https//doi.org/10.3233/SW-212883.

Elmehdwi Y., Samanthula B.K., Jiang W. (2013). Secure k-Nearest Neighbor Query over Encrypted Data in Outsourced Environments., https://doi.org/10.48550/arXiv.1307.4824.

Ghazarian A., Zheng J., Rakovski C. (2024). Privacy-Preserving ECG Data Analysis with Differential Privacy: A Literature Review and A Case Study. https://doi.org/10.48550/arXiv.2406.13880.

Jang M., M. Yoon M., Y. Song Y., и J.-W. Chang J.-W. (2015). A Signature-based Data Authentication Method with Bitmap-based Transformed Data in Database Outsourcing, 6th International Conference On Ambient Systems, Networks And Technologies (ANT-2015), The 5th International Conference On Sustainable Energy Information Technology (SEIT-2015), i. 52. в Procedia Computer Science, vol. 52. https://doi.org/10.1016/j.procs.2015.05.073.

Jang M., Yoon M., Song Y., Chang J.-W. (2015). A New Query Integrity Verification Method for Encrypted Data in Database Outsourcing., Advanced Multimedia and Ubiquitous Engineering: Future Information Technology, i. 352. в Lecture Notes in Electrical Engineering, vol. 352. pp. 67–72, https://doi.org/10.1007/978-3-662-47487-7_10.

Kim H.-J., Lee H., Kim Y.-K., Chang J.-W. (2022). Privacy-preserving kNN query processing algorithms via secure two-party computation over encrypted database in cloud computing., J. Supercomput., т. 78, вып. 7, сс. 9245–9284, https://doi.org/ 10.1007/s11227-021-04286-2.

Papadopoulos P., Pitropakis N., Buchanan W. J., Lo O., Katsikas S., (2020) Privacy-Preserving Passive DNS, Computers, vol. 9, iss. 3., https://doi.org/10.3390/computers9030064.

Przytarski D., Stach C., Gritti C., B. Mitschang B. (2021). Query Processing in Blockchain Systems: Current State and Future Challenges. Future Internet, т. 14, вып. 1, с. 1, https://doi.org/10.3390/fi14010001.

Riyana S., Sasujit K., Homdoung N. (2023). Aggregate Query Frameworks for Preserving Privacy Data in Big Data Analytics, In Review, https://doi.org/10.21203/rs.3.rs-2812564/v1.

Seeman J., Sexton W., Pujol D., Machanavajjhala A. (2024). Privately Answering Queries on Skewed Data via Per Record Differential Privacy, Proc. VLDB Endow., т. 17, вып. 11, сс. 3138–3150, https://doi.org/ 10.14778/3681954.3681989.

Shi Z., Fu X., Li X., Zhu K. (2020). ESVSSE: Enabling Efficient, Secure, Verifiable Searchable Symmetric Encryption, IEEE Trans. Knowl. Data Eng., сс. 1–1, https://doi.org/10.1109/TKDE.2020.3025348.

Scaglione A. (2022). The Use of Differential Privacy for Energy Data, CPSS'22: Proceedings Of The 8th ACM Cyber-Physical System Security Workshop. ASSOC Computing Machinery, https://doi.org/10.1145/3494107.3522780.

Welten S. et al., (2022). A Privacy-Preserving Distributed Analytics Platform for Health Care Data., Methods Inf. Med., vol. 61, iss. S 01, pp. e1–e11, https://doi.org/10.1055/s-0041-1740564.

Xia H., Xiong Y., Huang W., Meng Z., Miao F. (2020). Limiting Privacy Breaches in Average-Distance Query., Security And Communication Networks, V. 2020. https://doi.org/10.1155/2020/8895281.

Yang X., Kelarev A., Yi X., (2024). Privacy-enhancing data aggregation and data analytics in wireless networks for a large class of distributed queries, Wirel. Netw., vol. 30, iss. 6, pp. 4749–4759, https://doi.org/10.1007/s11276-022-03108-4.

Yang X., Yi X., Kelarev A., Rylands L., Lin Y., Ryan J. (2022). Protecting Private Information for Two Classes of Aggregated Database Queries., Informatics, vol. 9, iss. 3, p. 66, https://doi.org/10.3390/informatics9030066.

Yue R., Shaowei W., Xiangyu W. (2024). Emerging Trends in Sports and Artificial Intelligence: A Scientometric Analysis in Citespace, J. Electr. Syst., vol. 20, iss. 2, pp. 1897–1910, https://doi.org/10.52783/jes.1637

Yulianti E., Suwono H., Abd Rahman N. F., Phang F. A. (2024). State-of-the-Art of STEAM Education in Science Classrooms: A Systematic Literature Review, Open Educ. Stud., vol. 6, iss. 1, с. 20240032, 2024, https://doi.org/ 10.1515/edu-2024-0032.

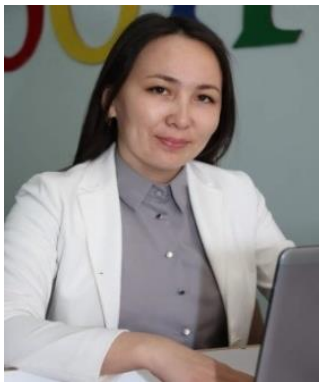
Zhang M., Wei E., Berry R., Huang J. (2024). Age-Dependent Differential Privacy, IEEE Transactions On Information Theory, vol. 70, iss. 2, https://doi.org/10.1109/TIT.2023.3340147.

Zheng Y., Wang W., Wang S., Jia X., Huang H., Wang C. (2022). SecSkyline: Fast Privacy-Preserving Skyline Queries over Encrypted Cloud Databases, https://doi.org/10.48550/arXiv.2209.07064.

Kazimova D., Tazhigulova G., Shraimanova G., Zatyneyko A., Sharzadin A. (2025). Transforming University Education with AI: A Systematic Review of Technologies, Applications, and Implications, Int. J. Eng. Pedagogy IJEP, vol. 15, iss. 1, pp. 4–24, 2025, https://doi.org/10.3991/ijep.v15i1.50773.

Авторлар туралы мәліметтер
Информация об авторах
Information about authors

Жантасова Женискуль Зейнешовна – техникалық ғылымдар кандидаты, С.Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен, Қазақстан
Жантасова Женискуль Зейнешовна – кандидат технических наук, Восточно-Казахстанский университет им. С. Аманжолова, г. Усть-Каменогорск, Казахстан
Zhantassova Zheniskul Zeineshovna – Candidate of Technical Sciences, East Kazakhstan University named after S. Amanzholov, Ust-Kamenogorsk, Kazakhstan,
e-mail: zheniskul_z@mail.ru,
ORCID: https://orcid.org/0000-0001-5550-7587



Нургалиева Сымбат Алтыбаевна – философия докторы (Phd), Astana IT University, Астана қаласы, Қазақстан
Нургалиева Сымбат Алтыбаевна – доктор философии (Phd), Astana IT University, Астана, Казахстан
Nurgalieva Symbat Altybaevna – PhD, Astana IT University, Astana, Kazakhstan
e-mail: symbat2030@gmail.com
ORCID: https://orcid.org/0009-0009-5880-6166

Сыздыкпаева Айгуль Рамазановна – техникалық ғылымдар кандидаты, С. Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен қ., Қазақстан
Сыздыкпаева Айгуль Рамазановна – кандидат технических наук, Восточно-Казахстанский университет им. С. Аманжолова, г.Усть-Каменогорск, Казахстан
Syzdykpayeva Aigul Ramazanovna – Candidate of technical science, S. Amanzholov East Kazakhstan University, Ust-Kamenogorsk, Kazakhstan
e-mail: aigul_uk@list.ru
ORCID: https://orcid.org/0000-0003-4965-0453