

¹С.К. Серикбаева*, ¹Д.А. Тусупов, ^{1,2}М.А. Самбетбаева

¹Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан

²БҒМ ҒК Ақпараттық және есептеуіш технологиялары институты, Алматы, Қазақстан

*e-mail: inf_8585@mail.ru

ТАРАТЫЛҒАН АҚПАРАТТЫҚ ЖҮЙЕЛЕРДЕ Z39.50 ХАТТАМАСЫН ҚОЛДАНУ

Аңдатпа. Бүгінгі таңда таратылған ақпараттық ресурстарға қол жеткізуді ұйымдастыру мәселелерін шешудің ең тиімді әдісі-олар туралы ақпаратты электрондық кітапханалар түрінде ашылған ақпараттық жүйелерге ұйымдастыру.

Бұл мақалада хаттаманың өзіне және стандарттардың тарихына функционалды шолу жасалады, оны қолданатын кейбір еуропалық бастамалар сипатталады, сонымен қатар оны болашақта пайдалану мен қабылдауға байланысты әртүрлі сұрақтар қойылады. Z39.50 болашақ таратылған ақпараттық жүйелердің маңызды құрылыс блогы болып табылады. Мақала тезаурустағы терминдерді іздеуді және Z39.50 протоколының Zthes профиліне сәйкес ұйымдастырылған тезаурусқа абстрактілі қол жетімділік негізінде терминдер ағашын құру алгоритмін сипаттауға арналған.

Негізгі сөздер: таратылған ақпараттық жүйелер, Z39.50, тезаурус, Zthes, ZSQL, RPN.

Кіріспе. Қазіргі қоғамда бірінші кезекте «үлкен мәліметтер», екіншіден, таратылған жүйелердегі мәліметтер өңдеу сапасы мен жылдамдығын жақсарту қажет [1]. Осыған байланысты, бұл мәселені шешу құралы ретінде үлестірілген сақтау жүйелерінің және деректерді өңдеудің маңыздылығы артып келеді. Кез-келген үлестірілген жүйенің негізгі міндеттерінің бірі - алынған мәліметтердің қасиеттерін талдау, оларды бірнеше себептер бойынша бір түйінде бағалау мүмкін емес. Осы мақсатқа жету және өңдеу уақытын жеделдету үшін бірінші кезеңде жүйенің үлестірілген түйіндеріне мәлімет жіберу керек, екіншісінде таратылған түйіндерден мәліметтер жинау және осы мәліметтерді жалпы ғаламдық көрініске жинақтау қажет. Бұл күрделі мәселелер, осы мәселелердің түрлерінде жиі кездесетін динамикаға байланысты, бұл бүкіл мәселенің жалпы ғаламдық қасиеттеріне әсер ететін жергілікті мәндерде өте жиі өзгеріс енгізеді. Тиімді және адаптивті үлестірілген жүйелерді құру деректерді өңдеу жылдамдығын едәуір жеделдете алады. Бұл мәселені қарастыру үшін біз таралған жүйелерді жобалау және пайдалану кезінде туындайтын мәселелерді талдаймыз.

Мәліметтер қорын құрудың, олардың мазмұнына қол жеткізуді ұйымдастырудың және мәліметтер қоры мен басқа ақпараттық ресурстардың арасындағы айқын және жасырын сілтемелерді жүзеге асырудың әртүрлі тәсілдері бар. Осы әдістердің көпшілігі қатаң онтологиялық модельдерге негізделген және практикалық іске асыру үшін ақпараттық жүйелер мен мәліметтер қорын ұйымдастыруға, аралық қоймаларға ақпараттардың толық шамадан тыс жүктелуіне дейін, олардың функционалдық қасиеттері берілген онтологиялық модельдер негізінде ақпараттық объектілер арасындағы барлық мағыналық қатынастарды анықтауға мүмкіндік беретін өте қатаң талаптар қояды. Мұндай тәсілдер әрекет етуге құқылы, дегенмен қолданыстағы таратылған ақпараттық ресурстарда және оларды мамандандырылған репозиторийлерге жүктеу мүмкін болмаған жағдайда мағыналық жағынан байланысты ақпаратты іздеуді қалай қосуға болады деген сұрақ туындайды.

Парадигма-бұл ғылыми қауымдастық қабылдаған және бөлісетін іргелі ғылыми білімнің, идеялар мен терминдердің жиынтығы. Бағдарламалау парадигмасы проблемаларды қоюдың бастапқы тұжырымдамалық схемасын және оны шешу жолдарын анықтайды.

Абстрактілі іздеу парадигмасын іске асыру бүгінде іздеу қызметтерін ұйымдастырудың бірнеше модельдері түрінде болады, мысалы Z39.50 моделі және қарапайым SRW/SRU моделі

[2]. Атрибуттық іздеу шарттарын тұжырымдау бойынша жетілдірудің жоғарғы жағы Z39.50 хаттамасына тиесілі.

Абстрактілі іздеу қызметтерін іс жүзінде жүзеге асыру тезаурустың айтарлықтай жаңа сапасын - олардың ресурстарын басқа жүйелермен статикалық WEB беттерінің сыртқы индекстеу деңгейінен жоғары деңгейде жаһандық іздеу жүйелеріне қосу мүмкіндігін береді [3].

Электрондық кітапханаларға арналған ақпараттық жүйелердің көптеген технологиялық әзірлемелері бар. Ресейде ең танымал, мысалы, Euro CRIS (eurocris.org), eLibrary (elibrary.ru), ақпарат (informika.ru), MathNET (mathnet.ru). Қазақстанда Қазақстан Республикасы Жоғары оқу орындарының қауымдастығы және KAZRENA Қазақстанның Ғылыми-білім беру компьютерлік желісі мен ҚР БҒМ ҒК ақпараттық және есептеу технологиялары институты ұсынған Бірыңғай технология бойынша университеттік білім беру ресурстары қорында республикалық жоғары оқу орындары арасындағы электрондық кітапхана (PM ЭБ) жұмыс істейді. Ғылыми-білім беру ақпараттық жүйесін дамытудағы ізашарлардың бірі және соның салдарынан олардың әлемдік көшбасшыларының бірі - "Ресей ғылым академиясының Сібір бөлімінің электронды кітапханасы", ол бүгінде Ресей ғылым академиясының СБ порталы деп аталады (sbras.ru). бұл ресурс Webometrics мәліметтері бойынша Ресейде ғылыми ұйымдар арасында бірінші және әлемде 45-ші орын алады.

Z39.50 стандартының негіздері. Z39.50 протоколы клиент пен сервердің өзара әрекеттесу тәртібін, дерекқордан ақпаратты іздеу және алу процедураларын және осы ақпаратты ұсыну форматтарын анықтайды.

Z39.50 хаттамасы нақты ДБ деректерді сақтау форматтарын, оларды индекстеу тәсілдерін және әртүрлі ДҚБЖ жұмыс істеу рәсімдерін анықтамайды. Z39.50 протоколы пайдаланушы мен клиенттің өзара әрекеттесу интерфейстерін де анықтамайды.

Z39.50 стандарты (Information Retrieval (Z39.50): Application Service Definition and Protocol Specification) деректер қорынан ақпаратты іздеу және алу үшін қолданбалы қызмет пен Протокол сипаттамасын анықтайды.

Ол деректер қорына желілік қол жетімділікті біріктіруге арналған және ақпаратты іздеу, алу және ұсыну форматтарын анықтайды [4].

Z39.50 деп аталатын хаттаманың алғашқы нұсқасын ұйымның АҚШ — тың Ұлттық ақпараттық стандарттар жөніндегі комитеті дайындады-NISO (National Information Standards Organization) және 1988 жылы Z39.50-1988 стандартымен енгізілді. Бұл нұсқаның әрекеті тек библиографиялық ақпаратпен жұмыс істеуге қатысты болды.

1989 жылы АҚШ Конгресі зертханасының әкімшілік басқаруымен Z39.50 хаттамасын қолдау агенттігі (Maintenance Agency Z39.50) ұйымдастырылды, ал 1990 жылы орындаушылар тобы құрылды

Z39.50 (Z39.50 Implementors Group — ZIG). Оның мүшелері әртүрлі ақпарат түрлерін (библиографиялық, мәтіндік, графикалық, қаржылық, химиялық және т.б.) өндірушілер, сатушылар, таратушылар болды. Z39.50 қолдау агенттігі-бұл стандартты сүйемелдеумен және дамытумен айналысатын тұрақты орган. Агенттіктің желілік мекенжайы - <https://www.loc.gov/z3950/agency/> мұнда хаттама бойынша барлық ақпарат, жаңалықтар, құжаттама, объектілер тізілімі және т. б. бар.

1992 жылы аталған ұйымдар Стандарттың 2-нұсқасын әзірледі (Z39.50-1992), 1988 жылғы стандартты ауыстырды, Стандарттың 3-нұсқасы (Z39.50-1995) 1995 жылы әзірленді [5]. Z39.50-1995 стандарты 1992 нұсқасының дамуы болғандықтан және онымен үйлесімді болғандықтан, ол 2 және 3 нұсқалардың хаттамасын анықтайды. 1995 жылы Z39.50 хаттамасы Американдық ұлттық ANSI стандарты (ANSI/NISO Z39.50-1995), ал 1998 жылдың қарашасында — ISO-23950 халықаралық стандарты ретінде қабылданды. Қазіргі уақытта жұмыс істейтін ANSI/NISO Z39.50-2003 [6] стандарты 2002 жылдың қараша айында бекітілді. Бұл ANSI/NISO Z39.50-1995 стандартының техникалық өңдеуі және 2 және 3 нұсқаларын анықтайды, бірақ

сонымен қатар Z39.50 — ZIG орындаушылар тобы ұсынған әртүрлі түсіндірмелер, түзетулер мен келісімдерді қамтиды.

Алғашқы жылдары хаттама негізінен библиографиялық ресурстарға қол жеткізуді ұйымдастыру үшін пайдаланылды, бүгінгі күні оны қолдану саласы едәуір кеңейтілді. Ол ғылыми-техникалық, қаржылық ақпаратқа, геоақпараттық ресурстарға, метадеректердің Ғаламдық базаларына, тезаурустар мен айдарларға, сандық коллекцияларға және мұражай ақпараттарына қол жеткізу үшін қолданылады.

Z39.50 стандарты ақпаратты іздеу үшін клиент/сервер түріндегі қызмет пен хаттаманы анықтайды. Ол клиенттің сервер ұсынатын мәліметтер қорынан іздеу, мәліметтер қорынан жазбаларды алу және ақпаратты іздеуге байланысты басқа функцияларды орындау үшін процедуралар мен форматтарды анықтайды. Z39.50 протоколы тек клиенттік және серверлік ақпараттық-іздеу қосымшалары арасындағы өзара әрекеттесуді анықтайды, ол клиент пен соңғы пайдаланушы арасындағы өзара әрекеттесуді анықтамайды. Дәлірек айтқанда, Z39.50 протоколы нақты дерекқорларда деректерді сақтау форматтарын, оларды индекстеу әдістерін және әртүрлі ДҚБЖ жұмыс істеу процедураларын анықтамайды. Сондай-ақ, ол пайдаланушы мен клиенттің өзара әрекеттесу интерфейстерін анықтамайды.

Хаттаманың егжей-тегжейіне тоқталмай-ақ, Z39.50 стандарты әртүрлі мәліметтер қорына бірыңғай қол жеткізуге мүмкіндік беретін компьютерлердің өзара әрекеттесу ережелерін анықтайды деп айтуға болады. Осылайша, тек бір ғана клиенттік қосымшаны пайдаланушы құрылымы мен ақпараттық презентация форматы әр түрлі болатын қашықтан таратылатын мәліметтер қорынан ақпаратты іздей алады.

Екі негізгі ерекшелік Z39.50 протоколын басқа хаттамалардан ерекшелендіреді. Біріншіден, бұл ақпарат ұсынудың абстрактілі моделі. Z39.50 идеологиясында бірдей мәліметтер схемасы шеңберінде, пайдаланылған ДҚБЖ, өрістер мен сұраныстар синтаксисіндегі физикалық айырмашылықтарға қарамастан, барлық мәліметтер базалары бірдей. Басқаша айтқанда, хаттама клиент пен сервердің өзара әрекеттесуінің әр кезеңінде ақпаратты ұсынудың абстрактілі моделін ұсынады. Z39.50-де клиент әрдайым бір сұраныс жүйесімен жұмыс істейді және мәліметтерді бірдей форматта алады. Екінші ерекшелігі-Z39.50 протоколы клиент пен сервердің сессиялық өзара әрекеттесуін толығымен қамтамасыз етеді. Бұл мүмкіндік протоколдың өзінде бар және серверлік жүйе немесе клиенттік бағдарлама болсын, оның барлық қосымшаларында жүзеге асырылады.

Z39.50 технологиясындағы тезаурустар. Халықаралық стандарттау ұйымының (ISO) анықтамасына сәйкес, тезаурус-бұл индекстеу тілімен басқарылатын сөздік, тұжырымдамалар арасында айқын априорлық қатынастар орнату үшін ресми түрде ұйымдастырылған [6]. Бұл анықтама тезаурусты құрайтын элементтер ретінде осы бірліктер арасындағы лексикалық бірліктер мен семантикалық қатынастарды орнатады. Тезаурустық қатынастар (тұқым — түр, бөлік — тұтас және т.б.) таксономия құрылымына қолданылады, яғни пәндік аймақтың негізгі таксономиялары анықталады.

Тарихи тұрғыдан, тезаурустар құжаттарды қолмен индекстеу үшін жасалды және оларды жасау кезінде автоматты индекстеуге қатысты мәселелер ескерілмеді.

Индекстелген ақпараттың барлық тақырыптық әртүрлілігіне сәйкес келетін тезаурусты құрудың қиындығы оның қазіргі ақпараттық жүйелерде танымал болмауының басты себебі болып табылады. Дегенмен, ғылыми-білім беру қызметін қолдау үшін ақпараттық-іздеу жүйелерінің тиімділігі мамандандырылған тезаурусты пайдалануға тікелей байланысты.

Ақпараттық-іздістіру тезаурусы (стандарттардың анықтамаларына сәйкес) — бұл құжаттар мен іздеу сұраныстарының мазмұнын сипаттауға арналған терминдер арасындағы нақты көрсетілген семантикалық байланысы бар табиғи тілдегі негізгі терминдердің нормативтік (бақыланатын) сөздігі [7]. Тезаурус белгілі бір пәндік аймақты сипаттауға арналған, оның әр термині берілген пәндік аймақтың кез-келген тұжырымдамасын білдіреді немесе сипаттайды.

Z39.50 жалпы идеологиясына сүйене отырып, кез-келген дерекқорға қол жеткізу барлық жеке құрылымдар дұрыс көрсетілетін бірыңғай стандартты деректер схемасы арқылы жүзеге асырылуы керек. Схема Zthes деп аталды.

Zthes Z39.50 протоколын қолдана отырып жұмыс істеуге арналған. Бұл схема терминдер арасындағы қатынастар түрлерінің өте шектеулі жиынтығын қолдануды қамтитынын ескеріңіз. Бұл үлкен үйлесімділік үшін әдейі жасалады.

Терминдер арасында, стандарттың ұсыныстарына сәйкес келесі байланыс түрлері орнатылады:

* BT-ата-аналық терминмен, яғни кең мағынадағы терминмен байланыс;

* NT-бала терминімен, яғни тар мағынадағы терминмен байланыс. BT ↔ NT байланысы өзара кері;

* USE — оның орнына қолданылатын терминмен байланыс;

* UF-өзара кері байланыс USE, Use u UF;

* Байланысты терминді анықтайтын RT байланысы;

* LE-лингвистикалық эквивалентті терминдер арасындағы байланыс.

BT және NT сілтемелері, сонымен қатар USE және UF өзара кері болып табылады. RT және LE байланыстары симметриялы.

Сонымен қатар, терминнің түрі стандарттың ұсыныстарына сәйкес анықталады.

Zthes схемасында терминдердің келесі түрлері ажыратылады:

TT - бұл жоғары деңгейдегі термин, яғни. неғұрлым кең кластың терминдері жоқ термин (BT сілтемесі бар терминдер);

NT - бұл жоғары деңгейлі термин емес, яғни BT типті сілтемелері бар дескриптор;

ND - негізгі емес термин;

NL - бұл жалған термин, яғни құжаттарды индекстеу үшін пайдаланылмайтын, бірақ класс бөлімінің логикалық негізін көрсету үшін иерархияға енгізілген термин.

Сұраныстарды түрлендіру және іздеу. Z39.50 протоколы бойынша жұмыс істейтін кез келген жүйе іздеу сұрауында терминдерді көрсету тәсілін айқындайтын Bib-1 атрибуттар жиынтығын міндетті түрде қолдауы тиіс. Bib - 1 құрамына атрибуттардың алты тобы кіреді:

Use жиынтығы - тікелей өрістер немесе олар "кіру нүктелері" деп аталады";

Relation жиынтығы - қатынас атрибуттары (мысалы, >, <, = және т. б.);

Position жиынтығы - іздеу жүйесінің орналасқан жерін анықтайтын атрибуттар

Өрістегі терминдер (өрістегі бірінші сөз, ішкі өрісте, кез-келген позиция)

Structure жиынтығы - іздеу терминінің құрылымын анықтайтын атрибуттар (сөз, сөйлем, кілт және т. б.)

Truncation жиынтығы - терминнің қысқару мүмкіндігін анықтайтын атрибуттар;

Completeness жиынтық - өрістің немесе ішкі өрістің толықтығын анықтайтын атрибуттар.

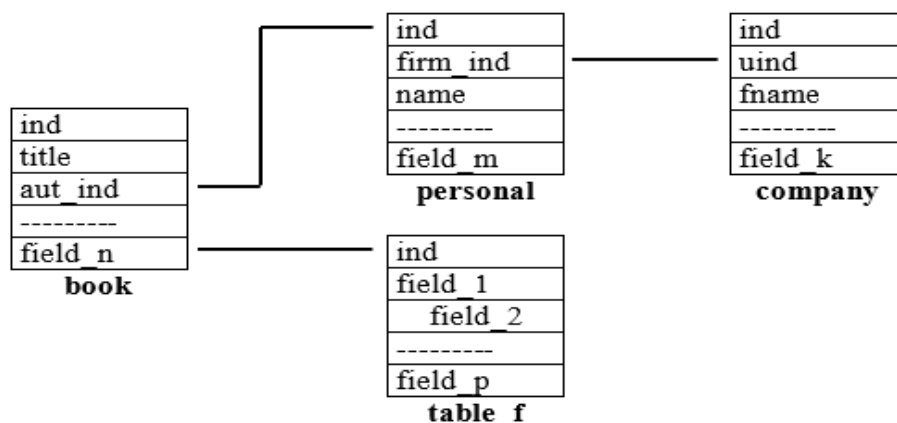
Бір жағынан, іздеу терминдерінің қасиеттерін анықтаудағы мұндай егжей-тегжей сұраныстарды дайындауда айтарлықтай икемділікті білдіреді, бірақ екінші жағынан, бір уақытта бірнеше серверге жіберілетін бір сұрау бойынша іздеуге кедергі келтіреді. Бұл әр түрлі серверлерде дерекқорды басқарудың әртүрлі құралдарын қолдануға байланысты стандарт ұсынатын барлық атрибуттар жиынтығын жүзеге асыру іс жүзінде мүмкін емес екендігіне байланысты. Көбінесе бір серверде іздеу үшін кез-келген атрибутты көрсету сәтті іздеудің қажетті шарты болып табылады, ал басқа серверде бұл атрибут немесе тіпті атрибуттардың бүкіл тобы іске асырылмайды. Бұл жағдайда деректер қорына жүгіну үшін бірыңғай сұрау салуды пайдалану мүмкін болмайды.

Бұл мәселе іздеу кезінде, ресурстың нақты орналасқан жері белгісіз болған кезде және іздеу бірнеше Z серверлерінде жүргізілуі керек. Ең алдымен, сұранысты құру кезінде іздеу өрісіне қосылатын іздеу өрістерін (кіру нүктелерін) таңдау мәселесі туындайды. Әр түрлі Z серверлері қолдайтын кіру нүктелерінің жиынтығы жиі ерекшеленеді. Автор, тақырып немесе

жариялау күні сияқты өрістер барлық дерлік мәліметтер қорында сақталса да, бұл көбінесе жеткіліксіз. Мысалы, ISSN сияқты өрістер серверге қолдау көрсетпеуі мүмкін. Мұның бәрі іздеу табысының күтпегендігіне әкеледі.

Z39. 50 протоколына негізделген әрбір жүйе, кем дегенде, абстрактілі идентификаторлар (oid сандық индекстері) жиынтығы арқылы көрсетілетін RPN (кері поляк белгісі) форматындағы сұраныстарды өңдеуді қолдауы керек [8]. Z39. 50 сервері осындай сұранысты алғаннан кейін оны белгілі бір дерекқорға қолайлы пішінге түрлендіруі керек. Z39. 50 сервері zsql түрінде сұраныстарды өңдей алады, олар абстрактілі деректер құрылымына да сүйенуі мүмкін (мысалы, кесте өрісінің атауының орнына оның сандық мәнін тиісті атрибуттар жиынтығынан пайдалану керек, ал FROM бөлімінде деректер схемасының идентификаторын көрсету керек және т.б.). ZSQL сұраулары "таныс" SQL түрінде де көрсетілуі мүмкін, бірақ бұл Z39.50 идеологиясына қайшы келеді, дегенмен оны ODBC, JDBC және т.б. технологияларға балама ретінде қолдануға болады, деректерді ортақ пайдалану үшін бірыңғай желілік интерфейсті ұсынады. Реляциялық ДҚБЖ-ны пайдалану кезінде сұрауға қандай кестелер қатысатынын, олардың арасындағы байланыс түрі қандай болатынын анықтау керек, сұраулардың еркін ұя салу мүмкіндігін ескеру қажет. - Сур.1-де Bib-1 іздеу атрибуттары мен белгілердің (tagset-G) OID көмегімен RPN және ZSQL сұрауының мысалы көрсетілген. Сауалдар туралы толық ақпаратты хаттаманың спецификацияларының сипаттамасынан алуға болады.

- а)
@and @attr 1=4 @attr 5=3 {система} @attr 1=1005 {ОИПТМ} @attr 1=1003 {Серикбаева}
- б)
**select [(2,1)]
 from [1.2.840.10003.13.2]
 where [(2,4)] = (select min([(2,4)]) from [1.2.840.10003.13.2])**
- в)



Сурет 1. Сұраныстарды түрлендіру: а) RPN түрінде сұрау; б) ZSQL-abstract түрінде сұраныс; в) қажетті ақпаратты қамтитын реляциялық деректер қорындағы кестелер мен олардың арасындағы қатынастардың мысалы

Әрбір дерекқорға ғана тән болуы мүмкін бұл файлдарда OID элементтерінің SQL сұрау элементтеріне сәйкестігі туралы ақпарат болуы керек. Егер сіз тек іздеу сұрауларымен (select) шектелсеңіз, онда конфигурация файлдарында кесте өрістерінің іздеу атрибуттары мен деректер схемаларының элементтеріне сәйкестігі, өрістердің түрлері, кестелер арасындағы қатынастардың сипаты туралы ақпарат болуы керек. Жалпы жағдайда, егер бірыңғай сұрау түрлендіргіштері жасалса, онда белгілі бір ДҚБЖ SQL диалектілері туралы ақпарат қажет

болады, өйткені SQL стандарттарының болуына қарамастан, сұрау тілдерінде кейбір айырмашылықтар бар [9].

Жалпы жағдайда конфигурация файлының құрылымы өте күрделі. Бірақ бірқатар қосымшалар үшін мұндай файлдарды айтарлықтай жеңілдетуге болады. Шынында да, ашық Ақпараттық жүйелер ұсынатын ақпараттың көлемі мен сипаты көптеген байланысты кестелерге күрделі сұраныстарды қолдануды білдірмейді. Z39.50 серверіне жүктеудің орнына сақталған процедуралар мен жалған кестелерді (view) қолданған дұрыс.

RPN (PQF)-сұраныстарды құру үшін келесі үлгіні пайдалану керек:

```
query ::= top-set query-struct
top-set ::= [ '@attrset' attrsetname ]
attrsetname :: 'Bib-1' | 'XD-1' | 'util' | 'Zthes'
query-struct ::= simple | complex
complex ::= operator query-struct query-struct.
operator ::= '@and' | '@or' | '@not'
simple ::= attr-spec term.
attr-spec ::= '@attr' [ attrsetname ] typ '=' value [ attr-spec ]
typ :: '1' | ... | '12'
value :: numeric or string
term ::= string.
```

Қарапайым RPN-сұраныстың мысалы ретінде тақырыпта кездесетін «Информационная система» терминін іздеу сұрауын келтіруге болады (termName):

```
@attr XD-1 1=1 Информационная система
мұнда
```

```
@attr XD-1 1=1 – «Атау термині» өрісіне сәйкес келеді;
Информационная система – іздеу термині.
```

Тақырыптарда кездесетін «Информационная система» немесе "Информатика" терминдерін іздеудің неғұрлым күрделі сұранысының мысалы:

```
@or @attr XD-1 1=1 Информационная система @attr XD-1 1=1 Информатика
```

Атау өрісінде «Информационная система» өрісінде termName және «ru» өрісінде termLanguage немесе «Информатика» өрісінде termName и «ru» өрісінде termLanguage бар жазбаларды табатын сұраныстың мысалы:

```
@or @and @attr XD-1 1=1 Информационная система @attr util 1=3 ru
@and @attr XD-1 1=1 Информатика @attr util 1=3 ru
```

Іздеу термині ретінде сөздегі таңбалар тізбегін орнатуға мүмкіндік беретін қосымша атрибутты қолдана отырып сұраныстың мысалы:

```
@attr XD-1 1=1 @attr util 5=4 @attr util 9=2 {ақпаратты іздеу}
```

Мұндай сұраудың нәтижесі "ақпаратты іздеу" және "ақпаратты іздеу"терминдері болуы мүмкін.

Сөйлем іздеу термині ретінде берілген сұраныс мысалы. Фраза-бұл бос орынмен бөлінген сөздер жиынтығы. Сұраныс нәтижесі сұраудағы сөздердің реттілігіне байланысты. Бұл жағдайда жауапта «ақпаратты іздеу» деген тіркес болады.

```
@attr XD-1 1=1 @attr util 5=4 {ақпаратты іздеу}
```

2019 жылғы барлық жазбаларды табатын сұрау мысалы, іздеу termCreatedDate өрісінде жүзеге асырылады:

```
@attr util 1=1 @attr util 12=2 {2019}
```

Іздеу нәтижелері

Деректер қорында іздеу нәтижесінде клиент серверден келесі ақпаратты ала алады: қате туралы хабарлама, табылған жазбалар саны немесе табылған жазбалар. Бірінші жауап кез-келген қатеге байланысты, екінші және үшінші нұсқалар сәтті іздеуге сәйкес келеді. Клиент

алатын олардың қайсысы серверге іздеу сұрауымен бірге берілетін параметрлерге байланысты болады.

Ол үшін шағын, орта және үлкен жиынтық ұғымдары енгізіледі. Мұнда жиын дегеніміз-ақырына дейін нөмірленген табылған жазбалар жиынтығы. Шағын жиынтықтағы барлық жазбалар әрдайым қайтарылады, үлкен жиынтықтағы барлық жазбалар ешқашан қайтарылмайды, ал кейбір жазбалар орта жиынтықтан қайтарылады. Келесі параметрлер орнатылады [10].

- Кіші жиынның жоғарғы шегі, яғни бірінші жазбадан басталатын кіші жиындағы максималды жазба нөмірі.

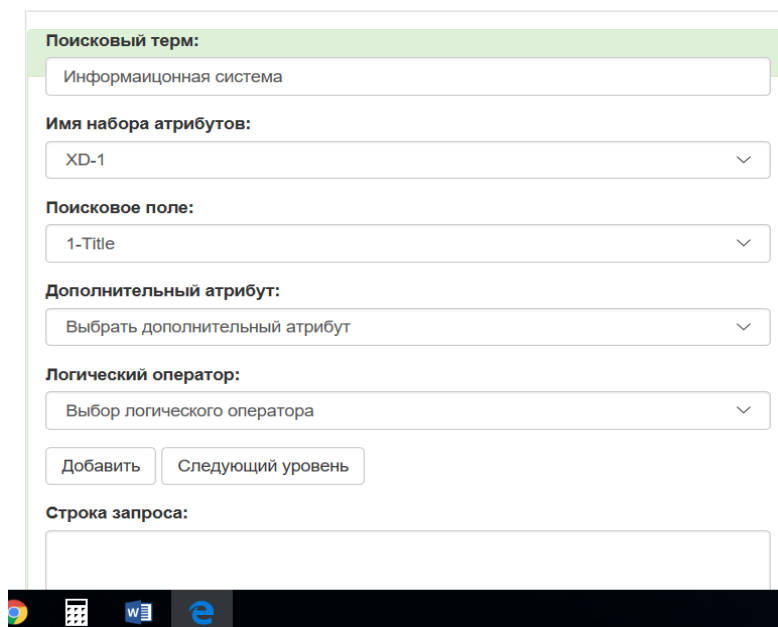
- Үлкен жиынның төменгі шекарасы, яғни жазбалар үлкен жиынға енетін Сан. Нөмірлері кіші жиынның жоғарғы шегінен үлкен, бірақ үлкен жиынның төменгі шегінен кіші барлық жазбалар орта жиыннан алынған жазбалар болып саналады.

- Орташа жиынтықтан қайтарылған жазбалар саны.

Осы үш параметрді өзгерту арқылы сіз жазбалардың кез-келген санын, соның ішінде біреуін де қайтара аласыз.

Соңында, жазбаны іздеген кезде табылған барлық сервер кейінірек пайдалану үшін сеанс блогында сақталуы керек екенін атап өткен жөн. Егер сервер іздеу нәтижесіне ат қоюға мүмкіндік берсе, бұл сақталған жиынтыққа атау берілуі мүмкін, егер жоқ болса, популяция аталмаған күйде сақталады және келесі іздеу кезінде қайта жазылады. Серверде сақталған нәтижелер жиынтығын келесі RPN сұраныстарында қолдануға болады, онда олар АРТ блоктары сияқты операндалар ретінде әрекет етеді.

Тезаурус дерекқорына қол жеткізуді орындау үшін пайдаланушы дерексіз сұрау салуы керек. Ол үшін тезаурус дерекқорына дерексіз сұраулар жасайтын пайдаланушы WEB қосымшасы жасалды. Дерексіз сұрау алу үшін пайдаланушы нысанды енгізу өрістерін келесі іздеу параметрлерімен толтыруы керек: іздеу терминінің атауы, атрибуттар жиынтығының атауы, Use типті іздеу атрибуты (Access Point). Пайдаланушы қосымша іздеу атрибуттары мен логикалық мәлімдемелерді қолдана отырып, дәлірек ақпаратты іздеу үшін сұраулар жасай алады.



Поисковый терм:
Информационная система

Имя набора атрибутов:
XD-1

Поисковое поле:
1-Title

Дополнительный атрибут:
Выбрать дополнительный атрибут

Логический оператор:
Выбор логического оператора

Добавить Следующий уровень

Строка запроса:

Сурет 2. Сұраныстың абстрактты параметрлерін таңдау

Деректерді толтырғаннан кейін «Добавить» батырмасын басу керек. Таңдалған іздеу параметрлері төмендегі өрісте көрсетіледі.

The screenshot shows a web interface for configuring a search level. At the top, there is a browser address bar with the URL 'tova/pages/wide.html'. Below it, the interface is organized into sections:

- Дополнительный атрибут:** A dropdown menu with the text 'Выбрать дополнительный атрибут'.
- Логический оператор:** A dropdown menu with the text 'Выбор логического оператора'.
- Two buttons: 'Добавить' and 'Следующий уровень'.
- Строка запроса:** A text input field containing the query '@attr XD-1 1=1 Информационная система'.
- At the bottom, two buttons: 'Очистить' and 'Сформировать запрос'.

Сурет 3. Іздеу параметрлерін көрсету

Неғұрлым күрделі іздеу құрылымдарын құру үшін пайдаланушы «Следующий уровень» батырмасын қолдануы керек. Осылайша, жаңадан таңдалған іздеу параметрлері дерексіз сұраудың алдыңғы фрагментіне қосылады.

The screenshot shows a more complex search configuration interface. It includes the following elements:

- Поисковое поле:** A dropdown menu with '1-Title' selected.
- Дополнительный атрибут:** A dropdown menu with 'Усечение справа' selected.
- Логический оператор:** A dropdown menu with 'or' selected.
- Buttons: 'Добавить' and 'Следующий уровень'.
- Строка запроса:** A text input field containing a complex query: '@or @attr XD-1 1=1 @attr 5=1 Информационная система @or @attr XD-1 1=1 @attr 5=1 Информатика'.
- Buttons: 'Очистить' and 'Сформировать запрос'.

Сурет 4. Күрделі сұраныстар құрудың мысалы

Абстракттілі сұранысты нақты SQL сұранысына тезаурус дерекқорына түрлендіру үшін келесі алгоритм бойынша жұмыс істейтін ДҚБЖ-ге ендірілген функция жасалды.

Жаңадан бастаушылар үшін пайдаланушы абстракттілі сұрау салуы керек, мысалы, қосымшаның көмегімен және оның дұрыстығын қосымшаның көмегімен тексеру керек:

@or @attr XD-1 1=1 Информационная система @attr XD-1 1=1 @attr 5=1 Информатика

Бұл жол түрінде сұраныс енгізу параметрі ретінде функцияға беріледі. Әрі қарай, сұраныс талданады:

@or – логикалық оператор «немесе»

@attr XD-1 1=1 – «Атау термині» бойынша іздеу

Информационная система – іздеу термині

@attr XD-1 1=1 – «Атау термині» бойынша іздеу

@attr 5=1 – оң жақта кесу

Информатика – іздеу термині

Алынған нәтиже SQL сұранысының фрагменттеріне ауыстырылады:

@or -> or

@attr XD-1 1=1 -> title

Информационная система -> 'Информационная система'

@attr XD-1 1=1 -> поиск по title

@attr 5=1 -> LIKE

Информатика -> 'Информатика%'

Әрі қарай логикалық операторлар мен жақшалар орналасады:

((title = 'Информационная система') or (title LIKE 'Информатика%'))

Осыдан кейін SQL сұранысы тезаурус дерекқорына орындалады:

SELECT * FROM zthes_cat WHERE ((title = 'Информационная система') or (title LIKE 'Информатика%'))

Алынған нәтиже пайдаланушы қосымшасындағы кестеде көрсетіледі:

| № | title | link_id | term_qualifier | term_vocabulary | description | document_language | term_category |
|---|------------------------|----------|----------------|-----------------|--|-------------------|---------------|
| 7 | Информатика | 54F38E0C | publ555 | | <i>Информатика — это наука, которая занимается вычислением, хранением и обработкой информации. Она развивается вместе с компьютерами и сетью интернет, а потому базируется на компьютерной технике и невозможна без нее.</i> | ru | |
| 2 | Информационная система | BFC88BB8 | abacus_ru | - | Информационная система — это взаимосвязанная совокупность средств, методов и персонала, используемых для хранения, обработки и выдачи информации для достижения цели управления. | ru | - |

Сурет 5. Іздеу нәтижесі

Қорытынды. Мақалада таратылған ақпараттық жүйенің негізгі түсініктері Z39.50 хаттамасында құрудың жолдары қарастырылып, Z39.50 хаттаманың дерексіз RPN (PQF) сұрауларымен (сұрау салудан бастап оны орындауға дейін) толық жұмыс циклы берілген, сонымен қатар SQL сұрау тіліне негізделген балама тәсіл қарастырылған. Сұраулар қарапайым және күрделі болуы мүмкін. Жазылған WEB қосымшасы, ол интерактивті түрде RPN сұранысын жасайды, оның дұрыстығын тексереді және орындайды.

Z39.50-бұл клиент-сервер архитектурасының стандарты, онда іздеу жүйесі мен интерфейс тәуелсіз бөліктерге бөлінеді. Әр түрлі жергілікті жүйелердегі кең таралған мәліметтер қорын бірдей жергілікті тұтынушы немесе интерфейс арқылы табуға болады. Бұл интерфейс қалай көрінуі керек немесе ол қалай әрекет етуі керек деген мәселені шешпейді, бұл интерфейс ті таңдау үшін пайдаланушыға байланысты. Кітапхана жүйелерінің интернетпен байланысы және Z39.50 хаттамасының дамуы жергілікті автоматтандырылған жүйе арқылы библиографиялық дерекқорлар мен толық мәтінді дерекқорлардың үнемі өсіп келе жатқан массивіне қол жеткізу мүмкіншілігін арттырады. Пайдаланушыларды әртүрлі есептеу платформаларын ұсынатын ресурстармен тікелей байланыстыру қабілеті Z39.50 хаттаманың институционалды жүйелерді байланыстыратын кітапханалар үшін артықшылығын арттырды. Осы хаттаманы пайдалану нәтижесінде әртүрлі ұйымдардың мәліметтер қорын қамтитын таратылған ақпараттық жүйелерді құруға болады.

ЭДЕБИЕТТЕР

[1] Мазов Н.А., Жижимов О.Л. Применение протокола Z39.50 для работы с тезаурусами и классификационными схемами // VII Международная конференция «Библиотеки и ассоциации в меняющемся мире: новые технологии и новые формы сотрудничества» - Крым-2000 (Судак, Украина, 03.06 - 13.06.2000): Труды конференции. - М.: ГПНТБ России, 2000. - Т.1. - С.402-405.

[2] ANSI/NISO Z39.50-2003. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification // NISO Press, Bethesda, Maryland, U.S.A. — ISBN 1-880124-55-6. — 267 p.

[3] ГОСТ Р ИСО/МЭК 7498-1-99 Информационная технология (ИТ). Взаимосвязь открытых систем. Базовая эталонная модель. Часть 1. Базовая модель // Москва, ИПК, Издательство стандартов. - 1999. - С. 62.

[4] Жижимов О.Л., Мазов Н.А. Принципы построения распределенных информационных систем на основе протокола Z39.50. - ОИГТМ СО РАН, Новосибирск: ИВТ СО РАН. - 2004. - ISBN 5-9554-0017-6. - 361с.

[5] Шокин Ю.И., Федотов А.М. Поддержка и развитие распределенных информационно-вычислительных ресурсов СО РАН // Вычислительные технологии. (Совместный выпуск). Вестн. КазНУ им. Аль-Фараби. Серия: Математика, механика, информатика. Ч. 4. — 2004. — Т.42. — Ч.4. — № 3. — С.324-334.

[6] Z39.50 Utility Attribute Set. Library of Congress. - March 2003 - <https://www.loc.gov/z3950/agency/attrarch/util.html>

[7] ANSI/NISO. Z39.19:2005 Guidelines for the construction, format and management of monolingual controlled vocabularies. NISO Press: Bethesda, MD, 2005. ISBN:1-880124-65-3

[8] Федотов А.М., Идрисова И.А., Самбетбаева М.А., Федотова О.А. Использование тезауруса в научно-образовательной информационной системе // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2015. - Т.13. - № 2. - С.86-102. - ISSN 1818-7900. - EISSN 2410-0420.

[9] Жижимов О.Л., Федотов А.М., Шокин Ю.И. Платформа ZooSPACE - организация доступа к разнородным распределенным ресурсам // Электронные библиотеки. - 2014. - Т.17. - № 2. - ISSN 1562-5419. <http://www.elbib.ru/index.php?page=elbib/rus/journal/2014/part2/ZFS>

[10] Santeyeva S.A., Zhizhimov O.L. User interfaces for working with thesauri and rubricators in distributed heterogeneous information systems on the example platform ZooSPACE // Совместный выпуск по материалам международной научной конференции «Вычислительные и информационные технологии в науке, технике и образовании» (СITech-2015) (24-27 сентября 2015 года): Вычислительные технологии, т.20; Вестник КазНУ им. Аль-Фараби, Серия математика, механика и информатика № 3 (86) / КазНУ им. аль-Фараби. - 2015. - Part I. - Алматы - Новосибирск. - P.224-230.

REFERENCES

[1] Mazov N.A., Zhizhimov O.L. Primenenie protokola Z39.50 dlya raboty s tezaurusami i klassifikatsionnymi skhemami // VII Mezhdunarodnaya konferentsiya «Biblioteki i assotsiatsii v menyayushchemsya mire: novye tekhnologii i novye formy sotrudnichestva» - Krym-2000 (Sudak, Ukraina, 03.06 - 13.06.2000): Trudy konferentsii. - M.: GPNTB Rossii, 2000. - T.1. - S.402-405.

[2] ANSI/NISO Z39.50-2003. Information Retrieval (Z39.50): Application Service Definition and Protocol Specification // NISO Press, Bethesda, Maryland, U.S.A. — ISBN 1-880124-55-6. — 267 p.

[3] GOST R ISO/MEK 7498-1-99 Informatsionnaya tekhnologiya (IT). Vzaimosvyaz' otkrytykh sistem. Bazovaya etalonnaya model'. Chast' 1. Bazovaya model' // Moskva, IPK, Izdatel'stvo standartov. - 1999. - S. 62.

[4] Zhizhimov O.L., Mazov N.A. Printsipy postroeniya raspredelennykh informatsionnykh sistem na osnove protokola Z39.50. - OIGGM SO RAN, Novosibirsk: IVT SO RAN. - 2004. - ISBN 5-9554-0017-6. - 361s.

[5] Shokin Yu.I., Fedotov A.M. Podderzhka i razvitie raspredelennykh informatsionno-vychislitel'nykh resursov SO RAN // Vychislitel'nye tekhnologii. (Sovmestnyi vypusk). Vestn. KazNU im. Al'-Farabi. Seriya: Matematika, mekhanika, informatika. Ch. 4. — 2004. — Т.42. — Ч.4. — № 3. — S.324-334.

[6] Z39.50 Utility Attribute Set. Library of Congress. - March 2003 - <https://www.loc.gov/z3950/agency/attrarch/util.html>

[7] ANSI/NISO. Z39.19:2005 Guidelines for the construction, format and management of monolingual controlled vocabularies. NISO Press: Bethesda, MD, 2005. ISBN:1-880124-65-3

[8] Fedotov A.M., Idrisova I.A., Sambetbaeva M.A., Fedotova O.A. Ispol'zovanie tezaurusa v nauchno-obrazovatel'noi informatsionnoi sisteme // Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Informatsionnye tekhnologii. - 2015. - T.13. - № 2. - S.86-102. - ISSN 1818-7900. - EISSN 2410-0420.

[9] Zhizhimov O.L., Fedotov A.M., Shokin Yu.I. Platforma ZooSPACE - organizatsiya dostupa k raznorodnym raspredelennym resursam // Elektronnye biblioteki. - 2014. - T.17. - № 2. - ISSN 1562-5419. <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2014/par2/ZFS>

[10] Santeyeva S.A., Zhizhimov O.L. User interfaces for working with thesauri and rubricators in distributed heterogeneous information systems on the example platform ZooSPACE // Sovmestnyi vypusk po materialam mezhdunarodnoi nauchnoi konferentsii «Vychislitel'nye i informatsionnye tekhnologii v nauke, tekhnike i obrazovanii» (CITech-2015) (24-27 sentyabrya 2015 goda): Vychislitel'nye tekhnologii, t.20; Vestnik KazNU im. Al'-Farabi, Seriya matematika, mekhanika i informatia № 3 (86) / KazNU im. al'-Farabi. - 2015. - Part I. - Almaty - Novosibirsk. - P.224-230.

¹S. K.Serikbayeva*, ¹J.A. Tussupov, ^{1,2}M.A. Sambetbayeva

¹L. N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan

²Institute of Information and Computational Technologies CS MES RK, Almaty, Kazakhstan

*e-mail: inf_8585@mail.ru

APPLICATION OF THE Z39.50 PROTOCOL IN DISTRIBUTED INFORMATION SYSTEMS

Abstract. To date, the most effective way to solve the problems of organizing access to distributed information resources is to organize information about them in information systems, denounced in the form of electronic libraries.

The article describes the search for terms in the thesaurus and the algorithm for constructing a term tree based on abstract access to the thesaurus, organized in accordance with the Zthes profile of the Z39.50 protocol

Keywords: distributed information systems, Z39. 50, thesaurus, Zthes, SQL, RPN.

¹С.К. Серикбаева*, ¹Д.А. Тусупов, ^{1,2}М.А. Самбетбаева

¹Евразийский национальный университет им. Л.Н.Гумилева, Нур-Султан, Казахстан

²Институт информационных и вычислительных технологий МОН РК, Алматы, Казахстан

*e-mail: inf_8585@mail.ru

ПРИМЕНЕНИЕ ПРОТОКОЛА Z39.50 В РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ

Аннотация. На сегодняшний день наиболее эффективным способом решения проблем организации доступа к распределенным информационным ресурсам является организация информации о них в информационные системы, обличенных в форму электронных библиотек.

Статья посвящена описанию поиска терминов в тезаурусе и алгоритма построения дерева терминов на основе абстрактного доступа к тезаурусу, организованного в соответствии с профилем Zthes протокола Z39.50

Ключевые слова: распределенных информационных системы, Z39.50, тезаурус, Zthes, ZSQL, RPN.