

**С.К. Серикбаева<sup>1\*</sup>, Д.А. Тусупов<sup>1</sup>,  
М.А. Самбетбаева<sup>1,2</sup>, Г.С. Боранкулова<sup>3</sup>**

<sup>1</sup>Л.Н.Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан

<sup>2</sup>БҒМ ҒК Ақпараттық және есептеуіш технологиялары институты, Нұр-Сұлтан, Қазақстан

<sup>3</sup>М.Х.Дулати атындағы Тараз өңірлік университеті, Тараз, Қазақстан

**Авторлар жайлы ақпарат:**

Серикбаева Сандугаш Курманбековна – докторант, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан

<http://orcid.org/0000-0003-1729-6875>, e-mail: inf\_8585@mail.ru

Тусупов Джамалбек Алиаскарович – физика-математика ғылымдарының докторы, профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Нұр-Сұлтан, Қазақстан

<https://orcid.org/0000-0002-9179-0428>, e-mail: tussupov@mail.ru

Самбетбаева Мадина Аралбаевна – PhD, қауымдастырылған профессор, Л.Н.Гумилев атындағы Еуразия ұлттық университеті, БҒМ ҒК Ақпараттық және есептеуіш технологиялары институты, Нұр-Сұлтан, Қазақстан

<https://orcid.org/0000-0001-9358-1614>, e-mail: madina\_jgtu@mail.ru

Боранкулова Гаухар Сарсембаевна – техника ғылымдарының кандидаты, М.Х.Дулати атындағы Тараз өңірлік университеті, Тараз, Қазақстан

<https://orcid.org/0000-0001-5701-8074>, e-mail: b.gau@mail.ru

**ҒЫЛЫМИ-БІЛІМ БЕРУ ҚЫЗМЕТІН ҚОЛДАЙТЫН  
ИНТЕГРАЦИЯЛАНҒАН ТАРАТЫЛҒАН  
АҚПАРАТТЫҚ ЖҮЙЕЛЕРДІ ҚҰРУ**

**Андатпа.** Мақала ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйелердің ашық семантикалық кеңістікте интеграцияланатын ақпараттық жүйелерді құру үшін модельдер, әдістер жасау технологиясын әзірлеуге арналған. Жүйені құрудың мақсаты зерттеуші-ғалымдардың ақпараттық қажеттілігінің екі типін қанағаттандыру болып табылады: қажетті ғылыми ақпарат көздері туралы ақпаратқа қажеттілік және ең қажетті ғылыми ақпаратқа қажеттілік. Концептуальді модельді өңдеу DELOS DLRM және OAIS RM модельдеріне негізделген.

**Түйін сөздер:** таратылған ақпараттық жүйелер, концептуальды модель, интеграция, онтология, метадеректерді алу, мәліметтер қоры, Z39.50, LDAP, OAI-PMH.

**Кіріспе**

Ұйымның ақпараттық ресурстарын дамыту мақсатында таратылған ақпараттық жүйелерге қол жетімділікті қамтамасыз ететін бірыңғай ақпараттық жүйеге интеграциялауға алып келеді. Ақпараттық ресурстардың таралуына қосымша талаптар қойылады, атап айтатын болсақ:

- Ақпараттық ресурстардың таралуы және дербестігі жағдайында жүйелердің жұмыс жасау қабілеті;
- Интероперабельділікті қамтамасыз ету, әртүрлі қосымшадағы ақпараттық ресурстарды қайта қолдану;
- Жүйелердің өзара әрекеттесетін компоненттердің әрекеттесуін интеграциялау мүмкіндігі;

- Ескі жүйелердің өзара интеграциясын сақтай отырып, жаңа талаптарға және технологияларға жауап беретін жаңа жүйелерге көшуі;
- Жүйелердің өмірлік циклін қамтамасз ету.

Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйелер әртүрлі ақпараттар жүйесімен жұмыс жасай алады. Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйесін құрудың негізгі мақсаты – ғылыми ортада ақпарат алмасу қарқынын жылдамдату және сапасын арттыру болып табылады. Ақпараттық жүйенің бірыңғай интеграциялануы және ақпараттық ресурстарды жүйелеу жұмыстарын кәсіби салаларға бөлу өзекті мәселенің бірі. Бұл ғылыми мақалалар, ғылыми құжаттар, электронды коллекциялар, онтологиялық сипаттамалар, деректер жиынтығы, логикалық сипаттамалар және т.б. болуы мүмкін. Ақпараттық ресурстар арасындағы семантикалық байланыстар олардың құндылығын арттырады және ақпаратты іздеуге және сәйкестендіруге қосымша мүмкіндіктер береді. [2, 90 б.]

Ашық семантикалық кеңістікке интеграцияланған мәліметтер – бұл белгілі бір пәндік аймақ туралы білім жиынтығы. Сонымен бірге ресурстарды пайдалану ресурстардың нақты көшірмелерін алу және оларды белгілі бір зерттеу ортасының ерекшеліктеріне сәйкес түрлендіру ережелерін анықтау мәселелерімен қатар жүреді. Бұл мәселелер талап етілетін білімді қажет ететін ресурстарды таратудың тиісті механизмдерінің болмауынан, сонымен қатар ресурстарды файлдар, электрондық кестелер, мәліметтер қоры, электронды құжаттар, каталогтар және басқалар түрінде ұсыну мен сақтау тәсілдерінің біртектілігінің болмауынан туындайды. Көбінесе анықталған мәселелер әр жеке жағдайда жеке шешіледі.

Мұндай жағдайда ғылыми зерттеулер нәтижелерімен алмасу процесін біріздендіретін таратылған ақпараттық жүйелерді даярлау зерттеушілердің аймақтық және ұйымдық тұрғыдан бөлінген топтарының тиімді өзара іс-қимылына ықпал ететін болады, бұл осы жұмыстың тұтастай маңыздылығын көрсетеді.

Мәліметтерді ашық семантикалық кеңістікке интеграциялай отырып, ғылыми зерттеулерді қолдау үшін таратылған ақпараттық жүйелердің прототипін құру технологиясын жасау келесі мәселелерді шешуге қол жеткізіледі:

- Ақпараттық ресурстарды сақтау жүйесінің концептуальді моделін құру;
- Ғылыми-ақпараттық процессті мәліметтерді идентификациялап және байланыстыру алгоритмдерін қоса отырып, автоматтандыру алгоритмдерін жасау және принциптерін зерттеу, атап айтқанда: ақпараттық жүйелердің нысандар арасындағы қатынастардың сипаттамасы;
- Таратылған ақпараттық жүйелердің пәндік аймақтың онтологиялық моделін қолдану арқылы интеграциялау процесінде құжаттардан метадеректерді шығаруды өңдеу.

Дайындалатын таратылған ақпараттық жүйелердің ақпараттық моделі көп деңгейлі болуы керек және келесі компоненттерден тұруы керек: деректер қоры (репозиторий), метадеректер сервері, қосымшалар сервері (диспетчер), ресурстарды жүйелеуге арналған бақыланатын жіктеуіш сөздіктерді қолдау жүйесі.

Метадеректер серверінің ақпараттары бойынша жүйенің стандартты қызметін қолдауды қамтамасыздандыратын мәліметтер сақталатын мәліметтер қорының жүйесі мен мәліметтердің қызметтік қорын жүргізу жүйелерінің динамикалық генерациясы жүзеге асырылады.

Бағдарлама сервері пайдаланушы қызметтерін қолдауға және алгоритмдерді, мысалы, морфологиялық талдауды немесе кластерлеу және жіктеу процедураларын қолдауға жауап береді. Ресурстарды басқару жүйесі таратылған ақпараттық жүйелер конфигурация және нақты мазмұн жөнінде бірыңғай анықтамалық жүйені жасауды талап етеді. Осындай анықтамалық жүйенің болуы таратылған ақпараттық жүйелері болып табылады [2].

Ғылыми-білім беру қызметін қолдауға бағытталған ақпараттық жүйеде ақпараттық ресурстардың барлық өмірлік циклінің сипаттамасын сақтау және кез-келген уақытта ресурстың жағдайын қалпына келтіру мүмкіндігі маңызды. Ұзақ уақыт қол жетімді болуы мүмкін ақпараттық ресурстар бар. Оларға мысалы, ғылыми мұраны құрайтын құжаттар мен жариялымдар, ұзақ мерзімді заңды күші бар құжаттар, патенттер, кез-келген уақыт аралығында талап етілуі мүмкін тарихи оқиғалар жөнінде мультимедиялық ақпараттар жатады. Сонымен қатар, институттардың ғылыми есептері, ғалымдардың сөйлеген сөздері, хаттар мен қызметтік жазбалар уақыт өте келе құнды бола отырып, үлкен тарихи маңызға ие болуы мүмкін. Сондықтан ақпараттық жүйе оларды қалпына келтіру мүмкіндігі бар ұзақ және сенімді ақпараттық ресурстарды сақтау мүмкіндігін қолдауы қажет.

Әрбір ресурстың эталондық модельге сәйкес идентификаторы бар:

- Ресурстың сипаттамасына сәйкес ұйымдастырылған. Ресурс күрделі және құрылымды болуы мүмкін, өйткені ол өз кезегінде аз ресурстардан тұруы мүмкін және басқа ресурстармен байланысы болуы мүмкін. Ұйымдастырушылық тұрғыдан алғанда, ресурстарды біртұтас субъект ретінде қарастырылатын ресурстар жиынтығына топтастыруға болады;

- Оның өмірлік циклін басқаратын функциялармен реттелуі мүмкін, оның қасиеттері мен басқа ресурстармен байланысын сипаттайтын өзіне тән атрибуттар мен әдістер жиынтығымен сипатталады;

- Ақпараттық нысан арқылы білдіріледі;

- Метадеректермен сипатталуы қажет, сонымен қатар суреттелуі немесе қосымша метадеректер мен аңдатпалармен толықтырулары мүмкін.

Ақпараттық жүйеде бастапқы ресурстың сипаттамасын қамтитын дәстүрлі қайталама ақпараттық нысан, яғни ақпараттық нысан – бұл ақпараттық жүйелер нысандары (физикалық нысандар, ресурстар, ақпараттық нысандар) туралы ақпаратты сақтайтын нысан әрбір Ақпараттық нысанға сәйкес келеді.

Ақпараттық нысан ресурс болғандықтан, ол ресурстың жоғарыда аталған барлық қасиеттерін мұра етеді, яғни идентификаторы бар.

Ақпараттық нысан – ақпараттық жүйеде ақпараттың еркін бірлігін білдіретін жүйеде ең жалпы түсінік.

Ақпараттық нысандар, сонымен бірге күрделі нысан болуы мүмкін және ақпараттық нысандар жинағына топтасуы мүмкін. Жинақтар, өз кезегінде,

ақпараттық нысандар болып табылады, олар ақпараттық нысандарды модельдеудің барлық аспектілерін және оларға қызмет көрсету құралдарын мұра етеді, мысалы, олар аннотациялануы мүмкін.

Ақпараттық нысандар ғылыми ақпараттық кеңістіктің жарияланым, құжат, тұлға, сөздік мақала, функция және пайдаланушы сияқты негізгі нысандарын, сондай-ақ олардың арасындағы байланыстарды сипаттайды.

Интеграцияланған ақпараттық жүйелерді жасау кезінде келесі фактіні ақылға қонымды түсіну: құжат мәтініндегі және метадеректеріндегі нысандар арасындағы қатынастар жиынтығы ақпараттық жүйенің онтологиясында сипатталған. Атап айтқанда, келесі маңызды ескертпе шығады: бұл жүйенің ішіндегі факт ретінде қарастырылатын нәрсені анықтайтын интеграцияланған ақпараттық жүйелер онтологиясы, себебі ол ақпараттық жүйенің мәндері арасындағы қатынастардың сипаттамасын белгілейді.

Ақпараттық жүйенің негізін немесе ядросын ақпараттық жүйелер жүйелейтін базалық нысандар құрайды. Әрбір базалық нысанда өзінің атрибуттар жиынтығы бар. Негізгі нысандардың даналары бір субъект ішінде де, қалған барлық нысандармен де аталған қатынастармен байланысты болуы мүмкін.

Электронды кітапханалар – бұл электронды құжаттарды желілер арқылы деректерді беру, жинақтау, сақтауға, өңдеуге, таратуға, талдауға, сондай-ақ іздестіру жұмыстарын ұйымдастыруға мүмкіндік беретін таратылған ақпараттық жүйелер каталогы.

Электрондық коллекциялар – бұл ақпаратты сақтау мен алмасудың жаңа түрі болып табылады.

Коллекция дегеніміз – семантикалық негізде біріктірілген және құрылымы бірдей құжаттар жиынтығы. Жинақ – мазмұны бірдей тақырыптық бағытқа ие, белгіленген құрылымы бар көптеген құжаттар. Құжаттармен жұмысты біріздендіру тұрғысынан біз ақпараттық жүйені коллекциялар жиынтығы түрінде ұсынамыз.

### **Материалдар мен әдістер**

Ғылыми-ақпараттық үрдісті автоматтандыру алгоритмдерін жасау және зерттеу қағидалары. Болжамды зерттеу ғылыми зерттеулер нәтижелерімен – ғылымды қажет ететін ресурстармен алмасу процесін біріздендіретін ақпараттық жүйе (АЖ) моделін дайындауға бағытталған. Зерттеу нәтижесі зерттеу нәтижелерімен алмасу процесін біріздендіретін деректерді байланыстыруды ескере отырып, ғылыми қызметті қолдаудың таратылған интеграцияланған ақпараттық жүйелердің (ИАЖ) прототипін жасауға болады.

Таратылған ақпараттық жүйелер моделі келесі нысандармен жұмыс істейді: Құзыреттер, оқу курстары, ұйымдар, деректер қоймасы, конференциялар, тұлғалар, жарияланымдар, ғылыми жобалар, гранттар. Қазіргі кездегі негізгі нысандар Құжаттар мен мәліметтер болып табылады.

Модель келесі қызметтерді қамтамасыз етуі қажет:

Тіркеу, атау, аннотациялау рәсімдерін және форматты анықтауды қамтитын ресурстарды жариялау.

Ресурстарды аналитикалық өңдеу.

Динамикалық құрылымдау функциясын қоса жарияланған ресурстарға қол жеткізу.

Автоматтандырлығын функциялау үшін ресурстарды бақылау және олардың метасипаттамаларын жаңарту функциясы, пайдаланушыларға жаңа ресурстардың пайда болуы және барларын жаңарту туралы хабарлау функциялары, диспетчерлеу функциясы қажет.

Қазіргі таңда таратылған ақпараттық жүйелерді құрудың негізгі үш технологиялық «киттер»: бұл WWW технологиялары, LDAP каталог моделі және Z39.50 протокол моделі [4] болып табылады. Бұл технологиялық компоненттер таратылған ақпараттық жүйенің негізгі функцияларын орындайды. Z39.50 стандарты бұл үшін барлық қажетті компоненттерден тұратынын белгілейміз. Басқа бірде-бір стандарт (бұл халықаралық стандарт, ал ұсыныс емес) мәліметтер қорына желілік қол жетімділікті реттейтін және ақпаратты іздеу үшін де, оны сақтау орындарынан алу және пайдаланушыға қажетті форматта беру үшін нақты сызбанұсқалар мен мәліметтер құрылымдарынан абстракциялауды реттейтін қажетті ережелердің толық жиынтығын қамтымайды. Басқару қызметін жүзеге асыру үшін (таратылған ақпаратқа қол жеткізуді басқару және таратылған ақпаратты басқару) LDAP хаттамасының моделіне негізделген технологиялар ең қолайлы, себебі қазіргі таңда LDAP негізінде таратылған ақпараттық жүйелердегі пайдаланушыларды сәйкестендіруді, аутентификациялауды және авторизациялауды қоса алғанда, ақпараттық ресурстарға «бірыңғай кіру нүктесі» технологиясын ұйымдастыру оңай. Бұл ретте LDAP каталогтардың барлық элементтерін жаһандық сәйкестендіру аясында ақпаратты (каталог ағаштарын) үлестірілген сақтау идеологиясына негізделгені, оның ішінде әртүрлі серверлер арасында мәліметтерді репликациялау тетіктері мен рәсімдерінің айқындамалары қамтылғандығы және қолданбалы және жүйелік бағдарламалық қамтылым дайындаушылары өте жақсы қолдайтыны маңызды жағдай болып табылады. Соңғысы жергілікті жүйеден таратылған ақпараттық жүйелерді басқаруға және олардың ресурстарына қол жетімділікті басқаруға ауысуға мүмкіндік береді [5].

Онтологияның негізі ұсынылған ақпараттың толықтығын ескере отырып, басқарылатын сөздіктерді (тезаурус) құру болып табылады [6], яғни онтология элем туралы білімді тіл туралы біліммен байланыстыратын пәндік аймақ пен тезаурус моделінің жалпы бөлігі ретінде қарастырылады, онтологиясыз толық тезаурус мүмкін емес, өйткені ол қарапайым түрде болса да, кез-келген мәліметтер жүйесінің және/немесе білімнің қаңқасы болып табылады. Тезаурусты осы тақырыптағы құжаттардағы семантикалық ақпараттың қоры, яғни білім құрылымының сипаттамасы ретінде түсіндіруге болады.

Ресей Ғылым академиясының есептеуіш технологиялары институтының Сібір филиалының Z39.50 хаттамасына және Zthes модификацияланған деректер схемасына сәйкес информатика бойынша көптілді тезаурус, атап айтқанда, ғылыми-білім беру қызметінде ақпараттық жүйені қолдау үшін лингвистикалық қолдау жасалды.

Тезаурусты жүзеге асыратын жасалған компоненттер web-формалар арқылы жүйеге тезаурус терминдерін қарауға, өңдеуге және қосуға, сондай-ақ, жазбаларды SKOS, MARC, MODS [7] мәліметтер схемаларына түрлендіруге болатын XML, RDF, DTD файлдары түрінде импорттауға және экспорттауға мүмкіндік береді.

Біз ақпараттық технологиялармен байланысты таңдалған, өте тар пәндік салада ақпаратты жіктеу мәселесін қарастырдық.

Айта кететін жай, ғылыми-білім беру ақпараттық жүйелерін құру барысында құжаттардың библиографиялық белгілері әдеттегіден әлдеқайда маңызды емес, фасет мәндерін құрайтын библиографиялық метадеректердің көптеген мағыналарының жиынтығы, әдетте, кеңірек болады. Сонымен, ғылыми-білім беру жүйелері тұрғысынан бір құжаттың әртүрлі қайта басылуына (стереотиптік қайта басып шығаруға) сілтемелерді балама деп санаған жөн.

Құжаттарды сипаттайтын негізгі терминдер, сонымен қатар аталған тақырып бойынша мәтіндерді жіктеу үшін олардың маңыздылығын көрсететін салмақтарды салыстыруға болады. Мысалы, ИТ-технологиялар бойынша «информатика» сөзінің маңыздылығы, себебі ол осы тақырып бойынша барлық дерлік құжаттарда кездеседі. Негізгі терминдердің маңыздылығын анықтау үшін аталған тақырып мәтінінде терминдердің маңыздылығын бағалайтын мысалы, TF-IDF өлшемін қолдануға болады.

Егер құжатты сипаттайтын терминдер көптеген синонимдерге немесе әртүрлі ассоциативті байланыстарға ие болса, құжат мәтініне ғана назар аударатырып, негізгі терминдерді таңдау дұрыс шешім болмауы мүмкін. Ғылыми терминдер бір-бірімен күрделі қарым-қатынаста болуы, неғұрлым тар немесе кеңірек ұғымдарды білдіруі, ұқсастық, қарама-қайшылық немесе басқа ассоциациялармен байланысты болуы мүмкін. Осыны ескеру үшін әр тілде арнайы тұжырымдамалық анықтамалықтар (тезаурус) жасау қажет [8].

### **Нәтижелер және талқылау**

Ақпараттық қоймасы ретінде DSpace жүйесі таңдалды. Жүйе пайдаланушылар туралы ақпаратты сақтау, авторизацияны қолдайды және осы ақпаратқа топтар, желілік мекен-жайлар бойынша және LDAP хаттамасының негізінде қол жеткізуді ажыратады, бұл АЖ құру кезінде қолданушыларды аутентификациялаудың бұрыннан бар жүйесін пайдалануға (өз бетінше әзірлемеуге) мүмкіндік береді және метадеректерге еркін қол жеткізуді қалдырып, жария және қызметтік ресурстарды оңай саралауға мүмкіндік береді. Мәліметтердің провайдерлердің болуы OAI-PMH, Z39.50 және SRW/SRU протоколдары үшін DSpace негізінде құрылған мәліметтер қоймасымен әртүрлі ақпараттық жүйе өзара әрекеттесу үшін бағдарламалық интерфейсті жасауға мүмкіндік береді.

Linked Open Data кеңістігімен интеграция аталған тапсырманы әмбебап шешімдерінің ішінде біреуі болып табылады. Linked Open Data пәндік аймақта өз ішінде неғұрлым көбірек ақпарат біріктіру үшін жасалған. Осылай осы кеңістікте мәліметтер жариялау арқылы біз бір жағынан барлық ақпараттарға қол жеткіземіз, ал екінші жағынан өз ақпараттарымызға қол жеткізе аламыз.

Жасалған жүйе модульді болуы, қолданыстағы автоматтандырылған жүйелермен масштабталатын және үйлесімді қосымша функционалдылықты арттыру мүмкіндігін қолдау керек. Жүргізілген салыстырудың нәтижесінде бағдарламалық қосымшаларды дайындаудың келесі технологиялары, тілдері мен құралдары ең жақсы болады деген қорытынды жасалды:

XSLT (XML/XSL-технология). Ерекшелігі – мәліметтерді олардың визуалды көрінісінен бөлу мүмкіндігі. Мәліметтерді ұсыну үшін XML файлындағы құжатты ұсынудың XSLT шаблонна сілтеме жасау жеткілікті, ол каскадты стиль кестелерінен (CSS) айырмашылығы, сүзгілерді, сұрыптауды, шарттарды және циклдерді қоса алғанда, мәліметтерді ұсынудың бай функционалдығын қамтамасыз етеді.

- SOAP. Аталған технология қолданыстағы веб-протоколдардың үстіне таратылған қосымшалар арасында құрылымдалған ақпарат алмасуға арналған. Ол кез келген платформада және кез келген тілде жазылған кез келген түрдегі нысандарға өзара әрекеттесу үшін мүмкіндік береді».

- WSDL. веб-сервисті, оның мекен жайын, хаттамасын, қолданылатын порттың нөмірін, сұраныс форматын және басқа көптеген нақты сипаттауға арналған технология.

- UDDI. веб-сервисті индекстейтін стандарт. Провайдерлер веб-сервисті өзі тіркей алатын, ал дайындаушылар – оны таба алатын бизнес-реестрді анықтау.

- Z39.50 хаттаманың ерекшелігі «жүйенің күйін сақтау және әрбір жағдайға тиісті сәйкестендіргіш беру мүмкіндігі» болып табылады. Бұл «уақытты үнемдеуге» мүмкіндік береді, яғни кез-келген уақытта сізге бұрын дайындалған белгілі бір іздеу нүктесіне оралуға болады.

- HTML5. Онда іздеу жүйелерімен өзара әрекеттесішу сайттарды жеңілдетін жаңа элементтер мен атрибуттар енгізілген.

Аjax – сервермен асинохронды өзара қарым-қатынас технологиясы. Осы технологияның арқасында серверге сұрақ жіберіп және парақты фондық режимде қайта жүктемей одан жауаптар алуға, қарқынды интерактивті веб-қосымшалар құруға болады.

- PHP + ZendFramework, бағдарламалық архитектура «Модель-Вид-Контроллер» (Model-View-Controller). Өңдеу және қолдау процесі қосымшаны олардың әрқайсысын модификациялау басқаларына минималды әсер ететіндей етіп жоғарыдағы үш компонентке бөлуден тұрады.

Библиографиялық ақпараттар көмегімен алмасуға болатын әртүрлі протоколдар бар. Ең таралған және қолданыстағы барлық дерлік АБИС Z39.50 және ОАИ-РМН (Open Archives Initiative Protocol for Metadata Harvesting) протоколдары болып табылады.

Жоғарыда көрсетілген және көптеген басқада технология қолдана отырып өңделген ZooSPACE технологиялық платоформасы, Есептеу технологиялары институты Сібір бөлімшесі Ресей Ғылым академиясы сатысында жетілдіріліп жатыр.

ZooSPACE платформасы осы ресурстарға қол жетімділікті ұйымдастырудың бірыңғай саясатына негізделген әр түрлі ДҚБЖ басқаратын гетерогенді ақпараттық ресурстарды біріктіретін таратылатын ақпараттық жүйелерді құруға бағытталған.

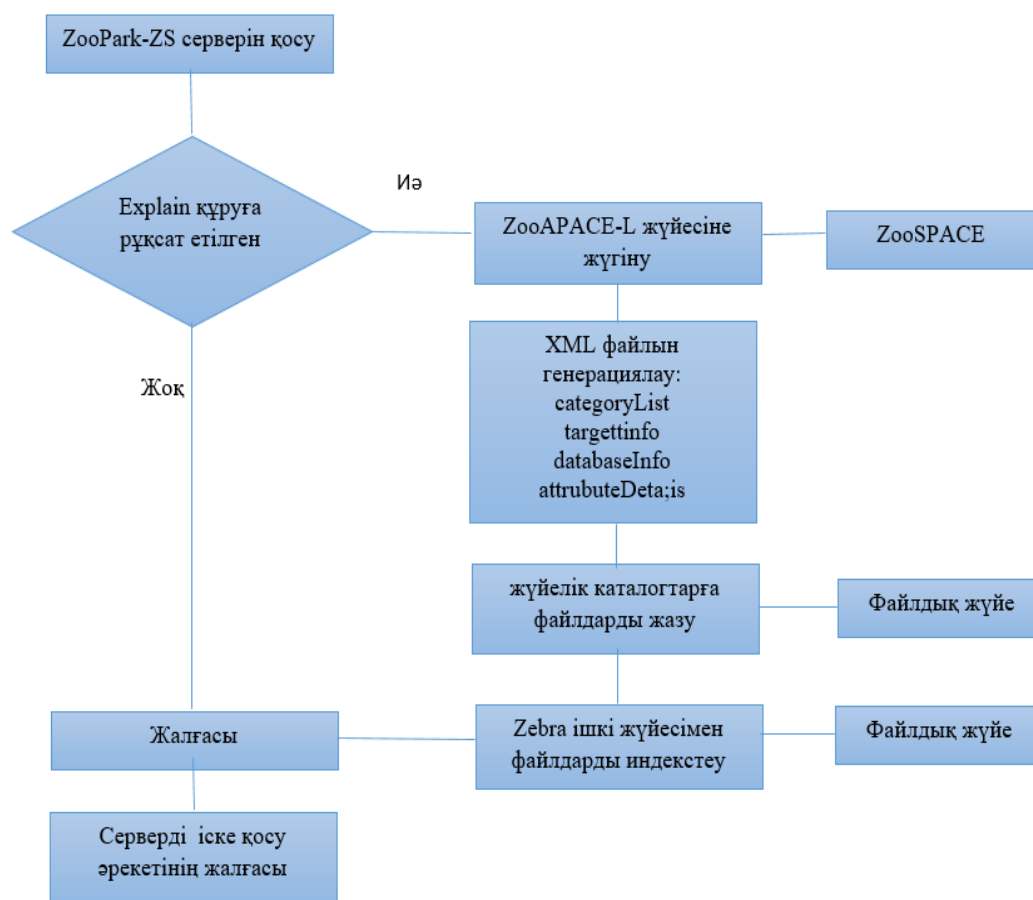
ZooSPACE бағдарламалық компоненттер серверлердің әртүрлі таратылған жүйе торабында шоғырланған бағдарламалы-аппараттық платформаларында қызмет етеді.

[9] жұмыста іздеудің адаптивті моделі өңделген және жүзеге асырылған және Z39.50 и SRU (1,2 сурет) протоколы үшін ZooSPACE-W кіші жүйесінде ақпараттар ұсынылған. Адаптация WEB-сервер жағында Explain ақпараты негізінде ақпараттық көздердің тізімін таңдағаннан кейін болады.

ZooSPACE-W кіші жүйесі ZooSPACE таратылған жүйесіндегі ақпаратқа қол жеткізу үшін әкімшілік және қолданушы веб-интерфейстерді ұсынуға арналған

Web-сервердің келесі қызметтері HTTP-Z39.50/SRW/SRU шлюзіне талаптарына сәйкес құрылымдалған:

- ZooSPACE мәліметтеріне қол жеткізу үшін қолданатын графикалық және әкімшілік Web-интерфейстерді ұсыну [10];
- LDAP арқылы қолданушылардың аутентификациясы;
- ZooSPACE өзекті серверларымен өзара әрекеттесу.

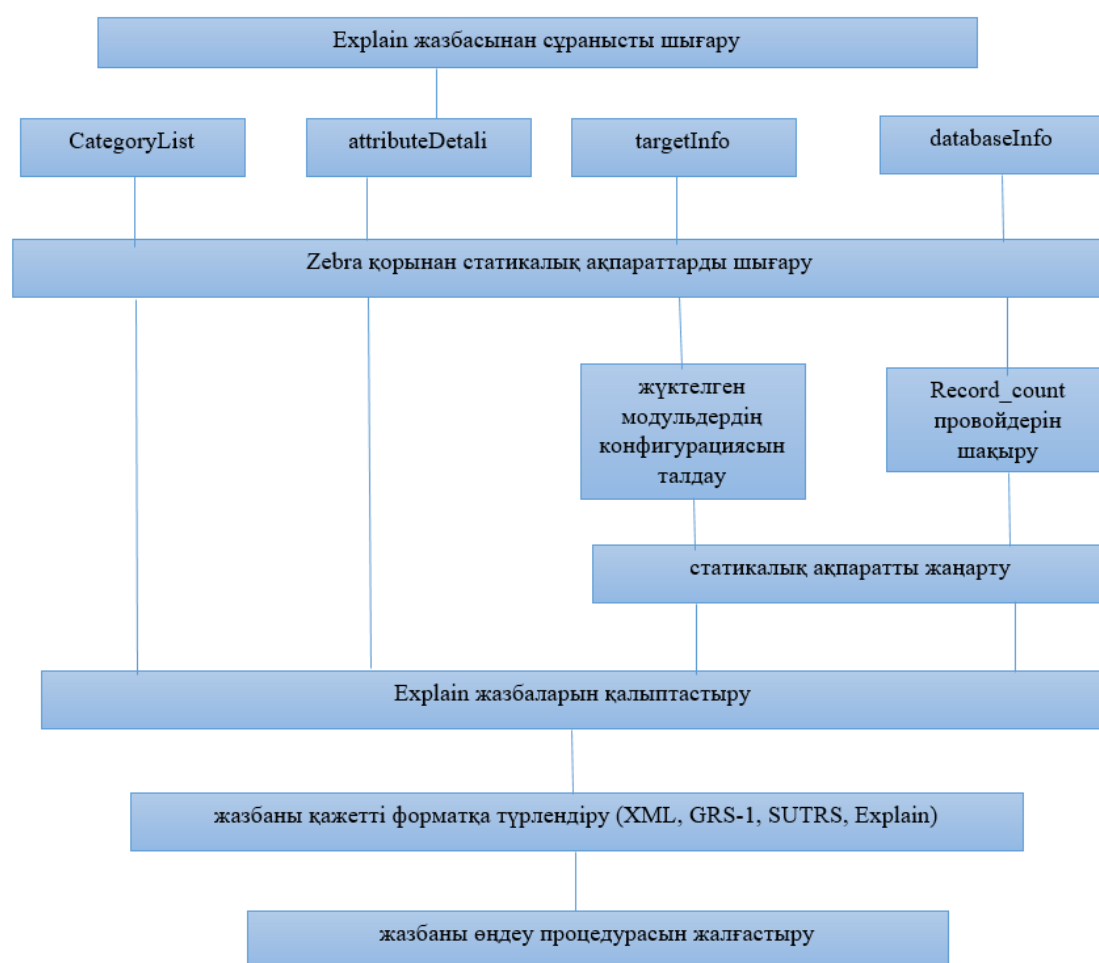


1-Сурет – Explain статистикалық ақпаратының генерациясы [автор материалы]



Web-интерфейстерді құрылымдау және мәліметтерді ішкі өңдеу үшін ZooSPACE-W кіші жүйесінің әрбір сервері қажетті функционалдылықты іске асыратын бағдарламалық қамтылымның бірдей жиынтығын сақтайды. ZooSPACE-W кіші жүйесінде бірнеше серверлердің болуы серверлердің қол жетімділік деңгейін жоғарылатады және әртүрлі тораптар арасындағы трафикті азайтады. Осының әсерінен клиент кез-келген серверге функционалдығын жоғалтпай қол жеткізе алады [11].

Explain жазбаларын алуға сұрау салуды қалыптастыру кезінде мәліметтер қорындағы жазбалардың ағымдағы саны және оларға қол жеткізу тәсілдері жөнінде ақпарат мүлдем жоқ. Оны пайдалану үшін сұрау қалыптастыру кезінде 1-суретте көрсетілген алгоритмге сәйкес осы ақпаратқа тиісті деректер провайдері шақырылады.



2-Сурет – Explain динамикалық ақпаратының генерациясы [автор материалы]

Бұл жүйе сандық контентті басқарудың үлкен мүмкіндіктеріне ие және жоғарыда аталған талаптарды толығымен қанағаттандырады: ашық код, әр түрлі ресурстармен жұмыс, LDAP интеграциясы, басқа жүйелермен мәліметтер алмасу үшін ОАІ протоколдарын қолдау, метадеректер мен толық мәтіндерді іздеу, кең таралуы және еркін таралуы [12].

## Қорытынды

Сандық репозиторияның (мәліметтер қоры) ақпараттық жүйеге интеграциялау процесі ақпараттық ресурстарды сақтауға арналған көптеген жүйелер қолдайтын OAI-PMH хаттамасына сәйкес метадеректерді біріктіру және тарату моделіне негізделген. Бағдарламалық қамтамасыз етудің салыстырмалы сараптамасы репозиторийлерді әзірлеу үшін Dublin Core схемасына және оның кеңейтімдеріне негізделген метадеректер саясаты қол жетімді DSpace бағдарламалық жасақтамасын қолданудың артықшылықтарын көрсетті. DSpace кітапханалық салалары үшін стандартты OAI-PMH, OpenURL және SWORD хаттамаларымен жұмыс істейді.

Құжаттарды рубрикациялау (классификациялау) міндеті, яғни құжатты бір немесе бірнеше тақырыпқа жатқызу қолжетімді толық мәтінді ақпарат көлемінің өсуіне байланысты өте өзекті болып табылады.

Жарияланымды айдарлар бойынша бөлу міндеті жақындықтың белгілі бір өлшеміне сүйене отырып, кластерлердің біріне сыныптың ең жақын данасын анықтау міндетіне дейін азаяды. Жұмыста тезаурусты қолданып, құжаттарды автоматты түрде айдарлау технологиясы сипатталады.

Алгоритмдерді тестілеу Ресей Ғылым академиясының Сібір бөлімінің Есептеу технологиялары институтында дайындалған «электрондық кітапханаларды басқару жүйесі» электрондық мәліметтер қорында өткізілді, онда жарияланымдар мен тақырыптық кластары жинағы бар.

## Әдебиеттер:

1. Арский Ю.М., Гиляревский Р.С., Туров И.С., Черный А.И. *Инфосфера: Информационные структуры, системы и процессы в науке и обществе*. М.: ВИНТИ. 1996, 489 с.
2. Fedotov A.M., Tusupov J.A., Sambetbayeva M.A., Fedotova O.A., Sagnayeva S.K., Baranov A.A., Tazhibaeva S.Z. *Classification model and morphological analysis in multilingual scientific and educational information systems. Journal of Theoretical and Applied Information Technology*. 2016, 86(1), 96-111.
3. Резниченко В.А., Проскудина Г.Ю., Кудим К.А. *Концептуальная модель электронной библиотеки. Труды XI Всерос. науч. конф. «RCDL'2009»*. Петрозаводск, Россия. 2009, 23–31.
4. Федотов А.М., Федотова О.А. *Модель информационной системы для поддержки научно-образовательной деятельности. Материалы Международной конференции «Вычислительные и информационные технологии в науке, технике и образовании ВИТ-2013»*. 2013, 2, 249-265.
5. Жижимов О.Л., Федотов А.М., Федотова О.А. *Построение типовой модели информационной системы для работы с документами по научному наследию. Вестник Новосибирского государственного университета. Серия: Информационные технологии*. 2012, 10(2), 5-14.
6. Fedotov A.M., Tusupov J.A., Sambetbayeva M.A., Sagnayeva S.K., Baranov A.A., Nurgulzhanova A.N., Yerimbetova A.S. *Using the thesaurus to develop it inquiry systems. Journal of Theoretical and Applied Information Technology*. 2016, 86 (1), 44-61.
7. Нариньяни А.С. *Кентавр по имени ТЕОН: тезаурус+онтология. Междунар. семинар по компьютерной лингвистике и ее приложениям. Аксаково*. 2001, 1, 184-188.

8. *ISO 5964:1985. Guidelines for the establishment and development of multilingual thesauri. – Geneva: International Organization for Standardization, 1985.*
9. Шокин Ю.И., Федотов А.М., Жижимов О.Л., Федотова О.А. Эволюция информационных систем: от Web-сайтов до систем управления информационными ресурсами. Вестник Новосибир. гос. ун-та. Серия: Информационные технологии. 2015, 13 (1), 117-134.
10. Федотов А.М. Методологии построения распределенных систем. Вычислительные технологии. 2006, 11, 3 – 16.
11. Федотов А.М., Жижимов О.Л., Федотова О.А., Барахнин В.Б. Модель информационной системы для поддержки научно-педагогической деятельности. Вестник Новосибир. гос. ун-та. Сер. Информ. технологии. 2014, 12 (1), 89 – 101.
12. Федотов А. М., Барахнин В. Б., Жижимов О. Л., Федотова О. А. Технология создания корпоративных информационных систем учета трудов научных работников. Вестник Новосибир. гос. ун-та. Сер. Информ. технологии. 2011, 9 (2), 31–41.

### References:

1. Arsky YuM, Gilyarevsky RS, Turov IS, Cherny AI (1996) *Infosfera: Informatsionnyie strukturyi, sistemyi i protsessyi v nauke i obschestve [Infosphere: Information structures, systems and processes in science and society]. M.: VINITI. Moscow, Russia. 489. (in Russ.)*
2. Fedotov A.M., Tusupov J.A., Sambetbayeva M.A., Fedotova O.A., Sagnayeva S.K., Bapanov A.A., Tazhibaeva S.Z. *Classification model and morphological analysis in multilingual scientific and educational information systems. Journal of Theoretical and Applied Information Technology. 2016, 86(1), 96-111. (in Eng.)*
3. Reznichenko V.A., Proskudina G.Yu., Kudim K.A. *Kontseptualnaya model elektronnoy biblioteki [Conceptual model of the electronic library]. Trudyi XI Vseros. nauch. konf. «RCDL'2009» = Proceedings of the XI All-Russian Scientific Conference «RCDL'2009». Petrozovodsk, Russia. 2009, 23-31. (in Russ.)*
4. Fedotov A.M., Fedotova O. A. *Model informatsionnoy sistemyi dlya podderzhki nauchno-obrazovatelnoy deyatel'nosti [Model of the information system for supporting scientific and educational activities]. Materialy Mezhdunarodnoy konferentsii «Vyichislitelnyie i informatsionnyie tehnologii v nauke, tehnike i obrazovanii VIT-2013» = Proceedings of the International Conference “Computing and information technologies in science, technology and education VIT-2013”: 2013, 2, 249-265. (in Russ.)*
5. Zhizhimov O. L., Fedotov A.M., Fedotova O. A. *Postroenie tipovoy modeli informatsionnoy sistemyi dlya raboty s dokumentami po nauchnomu naslediyu [Building a typical model of an information system for working with documents on scientific heritage]. Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Informatsionnyie tehnologii = Bulletin of the Novosibirsk State University. Series: Information Technologies. 2012, 10 (2), 5-14. (in Russ.)*
6. Fedotov A.M., Tusupov J.A., Sambetbayeva M.A., Sagnayeva S.K., Bapanov A.A., Nurgulzhanova A.N., Yerimbetova A. S. *Using the thesaurus to develop it inquiry systems. Journal of Theoretical and Applied Information Technology. 2016, 86 (1), 44-61. (in Eng.)*
7. Narinjani, A.S. *Kentavr po imeni TEON: tezaurus ontologiya [Centaur named THEON: Thesaurus+Ontology]. Mezhdunar. seminar on computational linguistics and its applications= International. Seminar on computational linguistics and its applications. Aksakovo, 2001, 1, 184-188. (in Eng.)*
8. *ISO 5964:1985. Guidelines for the establishment and development of multilingual thesauri. – Geneva: International Organization for Standardization. 1985. (in Russ.)*
9. Shokin Yu.I., Fedotov A.M., Zhizhimov O.L., Fedotova O.A. *Evolyutsiya informatsionnyih sistem: ot Web-saytov do sistem upravleniya informatsionnyimi resursami [Evolution of information systems: from Web sites to information resource management systems]. Vestnik Novosib. gos. un-ta. Seriya: Informatsionnyie tehnologii= Bulletin of the Novosibirsk State University. Series: Information Technology. 2015, 13 (1), 117-134. (in Russ.)*

10. Fedotov A.M. *Metodologii postroeniya raspredelennykh sistem [Methodologies for constructing distributed systems]. Vyichislitelnyye tehnologii=Computational Technologies. 2006, 11, 3–16. (in Russ.)*
11. Fedotov A.M., Zhizhimov O.L., Fedotova O.A., Barakhnin V.B. *Model informatsionnoy sistemyi dlya podderzhki nauchno-pedagogicheskoy deyatelnosti [Model of information system for support of scientific and pedagogical activity] Vestnik Novosib. gos. un-ta. Seriya: Informatsionnyye tehnologii= Bulletin of the Novosibirsk State University. Series: Information Technology. 2014, 12 (1), 89-101. (in Russ.)*
12. Fedotov A.M., Barakhnin V. B., Zhizhimov O. L., Fedotova O. A. *Tehnologiya sozdaniya korporativnykh informatsionnykh sistem ucheta trudov nauchnykh rabotnikov [Technology for creating corporate information systems for accounting for scientific workers works]. Vestnik Novosib. gos. un-ta. Seriya: Informatsionnyye tehnologii= Bulletin of the Novosibirsk State University. Series: Information Technology. 2011, 9 (2), 31-41. (in Russ.)*

**С.К. Серикбаева<sup>1\*</sup>, Д.А. Тусупов<sup>1</sup>,  
М.А. Самбетбаева<sup>1,2</sup>, Г.С. Боранкулова<sup>3</sup>**

<sup>1</sup>Евразийский национальный университет им. Л. Н. Гумилева, Нур-Султан, Казахстан

<sup>2</sup>Институт информационных и вычислительных технологий КН МОН РК,  
Нур-Султан, Казахстан

<sup>3</sup>Таразский региональный университет им. М. Х. Дулати, Тараз, Казахстан

## **ПОСТРОЕНИЕ ИНТЕГРИРОВАННЫХ РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ, ПОДДЕРЖИВАЮЩИХ НАУЧНО-ОБРАЗОВАТЕЛЬНУЮ ДЕЯТЕЛЬНОСТЬ**

### **Информация об авторах:**

Серикбаева Сандугаш Курманбековна – докторант, Евразийский национальный университет им. Л.Н. Гумилева, Нур-Султан, Казахстан

<http://orcid.org/0000-0003-1729-6875>, e-mail: [inf\\_8585@mail.ru](mailto:inf_8585@mail.ru)

Тусупов Джамалбек Алиаскарович – доктор физико-математических наук, профессор, Евразийский национальный университет им. Л. Н. Гумилева, Нур-Султан, Казахстан

<https://orcid.org/0000-0002-9179-0428>, e-mail: [tussupov@mail.ru](mailto:tussupov@mail.ru)

Самбетбаева Мадина Аралбаевна – PhD, ассоциированный профессор, Евразийский национальный университет им. Л.Н. Гумилева, Институт информационных и вычислительных технологий КН МОН РК, Нур-Султан, Казахстан

<https://orcid.org/0000-0001-9358-1614>, e-mail: [madina\\_jgtu@mail.ru](mailto:madina_jgtu@mail.ru)

Боранкулова Гаухар Сарсембаевна – кандидат технических наук, Таразский региональный университет им. М. Х. Дулати, Тараз, Казахстан

<https://orcid.org/0000-0001-5701-8074>, e-mail: [b.gau@mail.ru](mailto:b.gau@mail.ru)

**Аннотация.** *Статья посвящена разработке фундаментальных основ технологии создания моделей, методов для построения интегрированных в открытое семантическое пространство информационных систем поддержки научной деятельности, согласованных с мировыми стандартами и тенденциями развития как национальной, так и международной информационной инфраструктуры. Информационные системы поддержки научных исследований оперируют с различного рода информацией.*

**Ключевые слова:** *распределенные информационные системы, концептуальная модель, интеграция, онтология, извлечение метаданных, хранилище данных, Z39.50, LDAP, OAI-PMH.*

**S.K. Serikbayeva<sup>1\*</sup>, J.A. Tussupov<sup>1</sup>, M.A. Sambetbayeva<sup>1,2</sup>, G.S. Borankulova<sup>3</sup>**

<sup>1</sup>L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan

<sup>2</sup>Institute of Institute of Information and Computational Technologies CS MES RK, Nur-Sultan, Kazakhstan

<sup>3</sup>Taraz regional university named after M.KH.Dulaty, Taraz, Kazakhstan

**Information about authors:**

Serikbaeva Sandugash Kurmanbekovna – Doctoral student, L.N.Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan

<http://orcid.org/0000-0003-1729-6875>, e-mail: [inf\\_8585@mail.ru](mailto:inf_8585@mail.ru)

Tusupov Jamalbek Aliaskarovich – Doctor, Professor of physical and Mathematical Sciences, L.N.Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan

<https://orcid.org/0000-0002-9179-0428>, e-mail: [tussupov@mail.ru](mailto:tussupov@mail.ru)

Sambetbaeva Madina Aralbaevna – PhD, Associate Professor, L.N Gumilyov Eurasian National University, Institute of Institute of Information and Computational Technologies CS MES RK, Nur-Sultan, Kazakhstan

<https://orcid.org/0000-0001-9358-1614>, e-mail: [madina\\_jgtu@mail.ru](mailto:madina_jgtu@mail.ru)

Borankulova Gaukhar Sarsembaevna – Candidate of technical sciences, Taraz regional university named after M.KH.Dulaty, Taraz, Kazakhstan

<https://orcid.org/0000-0001-5701-8074>, e-mail: [b.gau@mail.ru](mailto:b.gau@mail.ru)

## **BUILDING OF THE INTEGRATED DISTRIBUTED INFORMATION SYSTEMS SUPPORTING SCIENTIFICALLY-EDUCATIONAL ACTIVITY**

**Annotation.** *The article is devoted to the development of the fundamental foundations of the technology of creating models, methods for building information systems integrated into the open semantic space to support scientific activities, consistent with international standards and trends in the development of both national and international information infrastructure. Information systems that support scientific research operate with various types of information.*

**Keywords:** *distributed information systems, conceptual model, integration, ontology, metadata extraction, data warehouse, Z39. 50, LDAP, OAI-PMH.*