

## ҒЫЛЫМИ-БІЛІМ БЕРУ ҚЫЗМЕТІН ҚОЛДАЙТЫН ТАРАТЫЛҒАН АҚПАРАТТЫҚ ЖҮЙЕЛЕРГЕ ҚОЙЫЛАТЫН ТАЛАПТАРДЫ ҚАЛЫПТАСТЫРУ

*Серикбаева С.К.<sup>1</sup>, PhD докторант*

*Тусупов Д.А.<sup>1</sup>, ф-м.ғ.д., профессор*

*Самбетбаева М.А.<sup>1,2</sup>, PhD доктор*

<sup>1</sup>*Л.Н.Гумилев атындағы Еуразия ұлттық университеті,*

*Мунайтпасов К., 5 Нұр-Сұлтан қ., 010011, Қазақстан, [inf\\_8585@mail.ru](mailto:inf_8585@mail.ru)*

<sup>2</sup>*БҒМ ҒК Ақпараттық және есептеуіш технологиялары институты, Пушкин 125, Нұр-Сұлтан қ., 010011, Қазақстан*

### Түйін

Ғылыми зерттеулердің тиімділігін арттыру барысында ғылыми жұмыстардың жүйеленген ақпараттық ресурстарға қол жеткізу қажет. Сондықтан кез келген ғылыми салада, зерттеу жұмыстарынан, ғылыми ақпараттарды іздеуден басталады, бірақ ғылыми мақалалар, кітаптар, монографиялар, патенттердің санының өсуіне қарай, ақпаратты іздеу жұмыстары қиынға түсуде. Ғалымдар өзге ғылыми зерттеулердің нәтижелерімен тез танысуға және олардың қайталануын болдырмауға мүмкіндік беретін бірыңғай ақпараттық жүйені құру. Мақалада ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің технологиялық әдістері қарастырылған. Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің моделін құрудың негізі тапсырмалары, модельдің функционалдық қызметін қамтамасыз етілуі, осы жүйеге арналған метадеректер түсінігі және метадеректер бейініне қойылатын талаптар сипатталған. Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің он екі талаптарын жүйелендіре отырып, оларға сипаттама берілді. Ұсынылған талаптар негізінде ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйесінің архитектурасы әзірленіп, құрылымы сипатталды.

**Кілт сөздер:** таратылған ақпараттық жүйе, метадеректер, Z39.50, LDAP, PostgreSQL, HDFS.

### Кіріспе

Соңғы жылдары ақпаратты беру және өңдеу салалары бойынша жоғары технологиялары қарқынды дамып келе жатыр, атап айтқанда қазіргі заманауи телекоммуникалық жүйелердің құрылуы, ғылыми-білім

беру қызметінің барлық деңгейлерін ұйымдастырудың жаңа мүмкіндіктері пайда болуына алып келді, бұл дегеніміз ғылыми-білім беру қызметін қолдайтын ақпараттық жүйенің

қажеттіліктерінің сапалы өсуіне әкелді [1, 3б.].

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 үйлесімдігін 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, 90б.]

Жұмыста ғылыми-білім беру қызметін қолдауға арналған ақпараттық жүйенің модельдерін құрудың технологиялық әдістері қарастырылады. Қарастырылатын модель, ғылыми материалдармен жұмыс жасауға арналған ақпараттық жүйенің дамыған моделі ақпараттарды ұзақ уақыт сақтауға, атрибуттар бойынша деректерді іздеуді ұйымдастыру, метадеректерді жинақтау және алмастырудың мәселелерін шешуі керек болып табылады.

Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің моделі

Ғаламдық ақпараттық және есептеу жүйелерінің қарқынды дамуында таратылған ақпараттық ресурстарды қолдануға және дамытуға көшу ретінде сипатталатын деректерді өңдеудің іргелі парадигмаларының өзгеруіне алып келеді [3, 32б.]. Сондықтан ақпаратпен жұмыс жасау технологиясымен байланысты маңызды мақсаты таратылған деректер көздерінің үйлесімділік тәсілдерін зерттеу болып табылады.

Ақпараттық ресурстардың үйлесімділігі деп олардың қолдану мақсатына қарай ұсыну, әртүрлі ақпараттың құрамында сақталуы, қолданушыға ыңғайлы интерфейстерді ұсынуды білдіреді. Бұл ретте ресурстарды біріктіру физикалық түрде жүргізілмеуі керек. Ең бастысы, ол қолданушыға қол жетімді ақпаратты бірыңғай қамтамасыз етуі керек. Атап айтатын болсақ, электронды кітапханалардағы ақпараттық жүйелерге қолжетімділік ерекшелеріктеріне қарамастан, қолданушының ақпаратты іздеудегі тиімділігін қамтамасыз ете отырып, гетерогенді деректер қорымен немесе деректермен жұмыс жасауын қамтамасыз етуі қажет [4, 12б.].

*Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің моделін құру өзекті мәселе бөліп табылады:*

- Ғылыми жұмыстар нәтижелерімен алмасу әрекетін біріздендіру;

- Ашық семантикалық кеңістікке біріктірілген деректермен және құжаттармен жұмыс жасау;

- Қолданушының қажеттіліктеріне сәйкес ресурстарды сипаттау, ұсыну, түрлендіру.

*Модель келесі функцияларды қамтамасыз етуі қажет:*

- Мақала ресурстарын тіркеу, аңдатпа рәсімдерін және форматты анықтау қадамдарын қоса сипаттау қажет:

- Ресурстарды аналитикалық талдау;

- Жарияланған ресурстарға қол жеткізу;

- Автоматтандырылған ресурстар жүйесіне бақылау және оларды метасипаттамаларын жаңарту;

- Қолданушыға жаңа ресурстардың пайда болуы және жаңарту туралы хабарлау.

Метадеректер бейінін таңдау

Интернет желісінде көптеген ғылыми деректердің көп болуына байланысты, құжатты іздеу мәселесі бірінші орында тұр, оларды жүйелендіріп беретін метадеректер қолдануға болады.

Метадеректер дегеніміз – ақпараттық объектілерді

сипаттаудың тиімді құралы болып табылады.

Метадеректер ресурстарды қолдану саласына қатысты ерекше сипатқа ие. Метадеректер ақпараттық кеңістікте және нақты әлемге қатысты нысандарды сипаттай алады. Ақпараттық ресурстардың метадеректері бөлек сақталуы мүмкін [5. 5б.].

Метадеректер – құжаттың туралы қосымша ақпараттардың сипаттамасын бере алады. Мысалға алатын болсақ, автор, атауы, қысқаша аңдатпасын және т.б. көруге болады.

Ақпараттық жүйелердің ішкі жүйелерін іске асыру халықаралық стандарттармен байланысты ашық спецификацияларға негізделуі тиіс. Таратылған ақпарат жүйелер ортасында деректерді синхрондау болуы керек, мысалы репликация түрінде таратылуы керек (1-сурет). Сонымен қатар, OAI-PMH, OAI-ORE, SRW/SRU, Z39.50, LDAP сияқты стандартты хаттамалар желілік аралық өзара әрекеттесуді қамтамасыз етуі керек.



1-сурет. Таратылған ақпараттық жүйелердің ішкі жүйелерінің өзара желілік байланысу хаттамасы

Метадеректер мынадай мәселелерді шешу үшін қажет:

- Құжаттың мазмұны, құрылымы, қолдану тәсілдері және т.б. туралы мәліметтерді ұсыну;
- Қажаттарды жүйелендіру және классификациялау;
- Ішкі жүйелерді ұйымдастыру кезеңі;
- Сыртқы ақпараттық жүйелермен алмасуды қолдау.

Метадеректер сипаттамалық, құрылымдық және әкімшілік бірнеше кластарға бөлінеді.

*Сипаттамалық метадеректер* – ақпараттық ресурстардың мазмұнын сипаттайды, мысалы Дублин ядросының метадеректер элементтерінің мәндерінің жиынтығы Dublin Core, егер ол мақалалар болса библиографиялық

деректер, андатпа, ресурстар идентификаторлары және т.б.

*Құрылымды метадеректер* – ресурстардың жалпы құрылымын және оның құрылымын, ресурстардың көлемін және басқа ұқсас қасиеттерін сипаттайды.

*Әкімшілік метадеректер* – ресурстың құру күнін, оның кім құрғанын немесе өзгерткенін, ресусқа құқық иесінің кім екенін, қолданушылардың кіру өкілеттілігін, ресурстардың сақталғанын және көшірмелері туралы деректерді, ресурсты басқару деректерді сипаттайды.

Атрибуттар бойынша ақпаратты күрделі іздеу және классификациялық функциялары үшін, толық мәтіндік іздеу, санаттар мен сөздік-классификаторлар бойынша ресурстарды қарау мүмкіндігін қамтамасыз ету қажет.

Олардың сипаттамаларын жасау кезеңінде өзара әрекеттесу мәселелеріне жеткіліксіз болуына байланысты: ұсынылған құжаттарды стандарттау жөнінде келісімдер және ұсынымдар, әртүрлі ақпараттық ресурстарды үйлесімділік құралдары әлсіз қолданылады. Ақпараттық жүйелердің өзара әрекеттестігі деп оның басқа ақпараттық жүйелермен, оның ішінде адамдармен өзара әрекеттесу қабілеттілік дәрежесі түсініледі. Бірақ егер адаммен қарым-қатынас кезінде (ақпараттық жүйелер сияқты) өзара түсіністікке жетудің негізгі ауырлығы нашар ұйымдастырылған ақпараттарды өңдей алатын соңғысына жүктелсе, онда ақпараттық жүйелердің дұрыс, арнайы технологиялық әдістері арасындағы тиімді өзара әрекеттесуді қамтамасыз ету керек және жалпы келісімдер қажет. Бұл барлық деректер сызба интерфейстердің және хаттамалардың халықаралық стандарттар мен ұсыныстарға сәйкестігін сақтау қажеттілігін тудырады.

Стандартты қолданбалы бейін білгілі бір топтар немесе қолданушылардың функционалды тапсырмалар негізінде әзірленеді. Бұл дегеніміз метадеректермен жұмыс жасау жүйелерін құруды оңтайландырады. Метадеректер бейіні белгі бір функцияны орындауға қажетті таңдалған стандарттар кластарын, ішкі жиындарын, қосымша мүмкіндіктері мен параметрлерін анықтауға болады [5, 96].

Мақалаларға арналған метадеректер саласында арнайы

ресурстардың бейінді элементтер тізімі болуы қажет, стандартқа сәйкес рұқсат етілген мәндер жиынтығын толықтыратын немесе кеңейтілген элементтердің мәндерін сипаттау үшін сөздіктер орнатылуы қажет. Сонымен қатар, элементтердің қосымша сипаттамалары ұсынылуы мүмкін.

Ғылыми жүйені әзірлеудің негізін ғылыми-білім беру жүйесінің бейінін қалыптастыратын стандарттар және халықаралық ұсыныстар құрайды, олар белгілі бір талаптарды шешуге бағытталған, нақты функцияларды орындауға қажетті стандарттар кластары таңдалған, ішкі жиындарын және опцияларын көрсете отырып, бір немесе бірнеше негізгі нормативтік-техникалық құжаттар жиынтығын білдіреді. Ақпараттық жүйе айналымында метадеректер бейіні ең маңызды болып табылады [6, 156].

Метадеректер бейіні келесі талаптарға сәйкес келуі қажет:

- Ғылыми-білім беру қызметін қолдайтын қажетті ақпараттың негізгі түрлерінің сипатамасын қамтамасыз ету;

- Метадеректер сипаттамасына сәйкес қолжетімділік ашық болуын қамтамасыз ету;

- Кеңейтілген сипаттамалардың нақтылау мүмкіндігін қамтамасыз ету;

- Ақпаратты біріктіру және ақпараттық үйлесімділігін қамтамасыз ету;

- Ақпаратты сұрыптау, жүйелендіру және классификациялауды қамтамасыз ету;

- Таратылған ортада ақпараттарды орналастыру және іздеу, басқа жүйелермен өзара үйлесу мүмкіндіктерін қамтамасыз ету;

- Ақпаратты сипаттау және заманауи технологияларды қолдануға мән беру [6, 156б.].

Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйелердің талаптары

Көптеген ақпараттық жүйелерге жоғарғы талаптар қойылады: қолданушыға қолдануға ыңғайлылығы және үйрену қолайлы; техникалық жағынан алғанда – әртүрлі ақпараттық жүйелермен өзара үйлесімділігі және стандартты хаттамалармен өзара әрекеттесуін қамтамасыз ету.

Ғылыми құжаттармен жұмыс жасау барысында бірнеше талаптар тұжырымдалған. Ғылыми-білім беру қызметін қолдайтын бірнеше функционалдық талаптар жиынтығын анықтауға болады [7, 286.]:

1. *Ақпараттық ресурстарды жинақтау.* Ақпараттарды жинақтау әртүрлі деректерді енгізуді қолдану қажет:

- қолданушы туралы деректі енгізу:

- арнайы интернет бағдарламалары арқылы деректерді жинақтау;

- басқа ақпараттық ресурстар деректерімен алмасу.

2. *Құжаттардың релеванты.* Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйесі үшін интернетте ақпаратты автоматты түрде жинақтау маңызды және маңызды емес болуы мүмкін.

Сондықтан туындаған мәселені мына жолдармен шешуге болады:

- ресурстар туралы метадеректерді және тақырыптарды классификациялауға арналған құрылымдалған анықтамаларды ұсынудың негізгі жүйеленген форматтарын құру. Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйелерінің метадеректерін сипаттауға және ресурстарды орналастыру кезінде белгілі бір форматтағы метадеректерді құру үшін қолданушыларға интерактивті веб-парақшаларын ұсыну тиіс;

- қолданушылар мен сарапшылар веб-парақшалар жүйесіне кіру нұсқасына қарай ақпараттық ресурстарды бөлу, сондай-ақ ақпараттың сенімділік дәрежесін арттырып көрсету;

- ақпаратты іздеу құралдарына іздеу және классификациялау кеңістігін көрсету, сондай-ақ енгізілген ақпараттың сапасын бағалау критерийлері;

- қолданушылардың қажеттілігіне қарай ресурстарды классификациялау сызбасын қолдану және сызбаға сәйкес ақпараттық ресурстарды классификациялау [8, 14-196б.].

3. *Құжаттардың өзектілігі, толықтылығы, түпнұсқалығы.* Ақпараттық құжаттардың өзектілігі мен толықтық мәселелері ресурстарды қамтамасыз ету мәселелерін шешу тәсілдеріне ұқсас болып келеді. Ал ақпараттың дұрыс шығу тегін анықтау тәсілдері:

-интерактивті ақпаратты енгізу үшін тек расталған (аутентификацияланған) қолданушылар енгізе алады;

-ақпаратты автоматтандырылған жинау жүйелерін орындайтын агенттің қолдану аясына шектеулер қою;

-басқа импортталған ақпарат жүйелер ресурстарымен ақпарат алмасу барысында сұрыптау жүргізу;

-барлық енгізілген ақпараттарды тексеру және классификациялау әдістерін жүргізу керек.

4. *Қолданушының сұраныстарын өңдеуде зияткерлік қызметтерді қолдану.* Қолданушының сұраныстарын өңдеу қызметтері бойынша – атрибуттарды іздеуді, толық мәтінді іздеу, санаттары бойынша ресурстарды қарауды және семантикалық іздеуді қамтамасыз етуі керек (міндетті емес).

5. *Білімді шығару.* Білімді шығарудың ішінара автоматтандыруды қолдануға болады. Ұсынылатын мәтіннің негізгі тәсілі семантикалық желі түріне негізделген, ал оның құрылу принципі мәтіндегі ұғымдардың бірлесу жиілігін қолдануға негізделген. Желіні қолданушығы тақырыптық терминдер ағаш түрінде ұсынылады, бұл дегеніміз қолданушыға навигация жасауға мүмкіндік береді және мәтінді іздеу, зерттеу процестерін айтарлықтай жеңілдетеді. Сонымен қатар, бұл тәсілмен рефераттарды, мәтіндік тақырыптарды классификациялау және кластерлеу, семантикалық іздеу тағы басқа сияқты мәселелерді шешу үшін қолдануға болады. Әртүрлі ақпараттармен жұмыс жасайтын ғылыми-білім беру қызметін қолдайтын таратылған

ақпараттық жүйесіне келесі талаптарды қарастыру керек болып табылады.

6. *Орталықтырылмаған ақпараттық жүйелердің сәулетін қолдау.* Бұл талап ақпараттың толықтылығы, түпнұсқалығы және өзектілігі үшін қажетті шарт болып табылады. Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйесін қолдану тәжірибесі ғылымның бір саласында немесе қандай да бір ғылыми ақпаратты қамтитын орталықтандырылған ғылыми жүйелерді құрудың күрделілігін көрсетті.

7. *Ақпараттық кеңістің құрылымдылығы.* Ақпаратты іздеу және классификациялаудың күрделі функцияларын қолдану үшін толық мәтінді сақтаудан басқа, ақпаратты атрибуттар бойынша іздеу, толық мәтінді іздеу, санаттар мен классификалық сөздіктер бойынша ресурстарды қарау қажет болып табылады. Бұл жағдай классификаторды таңдауда мамандандырылған жүйенің дәрежесімен сипатталады.

8. *Ақпаратты бейімдеп ұсыну.* Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің қолданушыларының сұраныстары бойынша ақпаратты іздеу сапасын жоғалтпай, ақпаратты іздеу жылдамдығы мен таңдау дәлділігін арттыру, жүйеде жұмыс жасау кезінде құзыреттілігін, уақыт шектеулерін ескеру қажет. Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйені қолданушыларға ақпаратты жылдам іздеу үшін қысқаша сипаттамадан бастап, ақпараттың объектілерін

нақты сипаттауға дейін ұсыну кезінде абстаркцияның әртүрлі деңгейлерін алуға мүмкіндік беруі керек болып табылады.

9. *Ақпараттың тарихилылығы.* Ғылыми ақпараттардың ерекшелігі – оның тез ескіруі және өзектілілін жоғалту болып табылады. Ақпараттық ресурстардың көптеген түрлері үшін деректер қорында ақпараттың барлық өзгерістерін сақтау маңызды және қалпына келтіру мүмкіндігі қажет. Мысалға алатын болсақ, адамның тегі, жұмыс орны өзгерген кезде авторлар туралы ақпарат уақыт өте келе өзгеруі мүмкін. Сол себепті уақыт факторын ескере отырып, уақыт аралықтарымен байланысты субъектілер үшін заманауи ақпаратты қолдану қажет.

10. *Мұрағат.* Жоғарыда атап айтылғандай, ғылыми ақпараттық көп бөлігінің өзектілігі ескіреді. Бірақ ұзақ уақыт ақпараттық ресурстарға қолжетімділікті қамтамасыз ету қажет. Мысалға, олар ұзақ мерзімді күші бар заңды құжаттар, патенттер немесе мультимедиалық ақпаратты кез келген уақытта талап етілуі мүмкін. Сонымен қатар, ғылым саласындағы ғалымдардың ғылыми баяндамалары, ғалымдардың сөйлеген сөздері уақыт өте келе маңызын ала отырып, үлкен тарихи құндылыққа ие болуы мүмкін. Сондықтан ақпараттық ресурстарды ұзақ мерзімді қалпына келтіру мүмкіндігін қолдауы керек.

11. *Ақпаратты іздеуде ақпараттың жіктелуін қолдану.* Ақпаратты іздеу мен классификациялауға арналған күрделі функцияларды пайдалану

үшін толық мәтіндік сипаттамаларды сақтаудан басқа, атрибуттар бойынша іздеуді, толық мәтінді іздеуді, санаттар және сөздік-жіктегіштер бойынша қарауды жүзеге асырылуы қажет. Жіктегіштерді мамандандырылған жүйе арқылы таңдау дәрежесі анықталады. Ол функцияларды жүзеге асыру үшін ғылыми-білім беру қызметін қолдайтын ресурстардың сәйкестендірілуі мен жіктелеуін қамтамасыз ететін сөздік-жіктегіштер болуы керек.

12. *Таратуды қолдау.* Әлемде ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйелерді дамыту қарқынды өсуіне қарай, келесі талаптар қойылады:

- қабылданған деректердің экспорты мен импорты үшін метадеректер стандарттарын қолдау;
- басқа ақпараттық жүйелер арасында ақпарат алмасу хаттамаларын қолдау;
- қолданушы ішкі ресурстармен байланысу интерфейстерін де, жүйелік деңгейде мүмкіндіктерді қолдау [9, 28-33бб.].

Жоғарыда атап айтылған талаптарға сәйкес, ғылыми-беру қызметін қолдайтын таратылған ақпараттық жүйенің прототипін құру технологиясын қарастырамыз. Ақпараттық жүйелердің негізгі міндеттері ақпараттық ресурстарды жинау, сақтау, өңдеу, тіркеу, жаңарту және қолданушылардың сұраныстарын өңдеуді қамтамасыз ету болып табылады.

1. *Ақпараттық ресурстарды жинақтау және тіркеу.* Ақпараттық ресурстарды жинақтау және тіркеу қызметтері бір мезгілде жүзеге асырылуы мүмкін. Ақпараттық



ресурстарды жинақтау мен тіркеу функцияларын жүзеге асыруда, келесі міндеттерді шешу қажет:

- Деректерді сұрыптау міндеті. Ақпараттық жүйеге кез келген деректер көздерінен ақпарат келеді, сондықтан олардың көбісі артық болуы мүмкін. Ақпаратты сұрыптау нәтижесінде жүйеге кіретін ақпараттық жиынтықтан қажетті деректер таңдалады;

- Деректерді тексеру міндеті. Деректердің шынайлығы мен логикалық тұтастығын тексеруді қамтамасыз етеді. Деректер қорына енгізілген ақпараттардың сәйкестігіне сараптамалық талдау және тексеру негізінде жүзеге асырылады. Деректердің логикалық тұтастығын бақылау құралдарының көмегімен алдын-ала ақпаратты өңдеу кезеңінде жүзеге асырылады;

- Деректерді сығу (сжать) міндеті. Жадтағы ақпараттық ресурстарды сақтауды барынша азайту, сонымен қатар байланыс арналары арқылы деректерді жіберу құнын төмендету мақсатында жүзеге асырылады;

- Деректерді бір форматтан екінші форматқа түрлендіру міндеті. Деректерді бір ақпараттық жүйеден екінші ақпараттық жүйеге ауыстыру қажет болғанда және жүйедегі әртүрлі ақпараттар арасында деректерді жіберу кезінде жүзеге асырылады.

Аталған функцияларды жүзеге асыру барысында деректерді тазалау, тексеру, сығу, деректерді бір форматтан екінші форматқа түрлендіру мәселелерін шешу кезінде Z39.50 және LDAP хаттамаларын қолдануға негізделген. Әртүрлі жүйелердегі

кең таралған деректер 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 қолданушыға 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 жеткізуге жол ашады. Пайдаланушыларды әртүрлі есептеу платформаларын ұсынатын ресурстармен тікелей байланыстыру қабілеті Z39.50 хаттаманың институционалды жүйелерді байланыстыратын кітапханалар үшін тартымдылығын арттырды. Осы хаттаманы пайдалану нәтижесінде әртүрлі ұйымдардың деректер қорын қамтитын таратылған ақпараттық жүйелерді құруға болады.

2. *Ақпараттық ресурстарды сақтау.* Жүйе таратылған ақпараттық жүйелердің басқару және деректерді сақтауды, әртүрлі деректердің сақталатын құрылымын, олардың тұтастығын және оларға қолжетімділікті қамтамасыз ету.

3. *Ақпараттық ресурстарды өңдеу.* Деректер қорында сақталған ақпараттық ресурстарды өңдеп қолданушыға өңделген ақпаратты ұсынуға мүмкіндік береді. Жүйеде ақпаратты өңдеу функцияларын орындау кезінде – деректер тұтастырығын, сөздіктер мен индекстерді іздеу жұмыстарын жүзеге асырады.

4. *Ақпараттық ресурстардың өзектілігі.* Ақпараттық жүйелерде ақпараттық ресурстардың пәндік облысын қолдау және модельдерін сипаттауға мүмкіндік береді. Ақпараттық жүйелердің динамикалық модельдерін пәндік облысы бойынша өзектілігін арттыру қажет. Ақпараттық ресурстардың өзектілігіне құжаттарды енігізу, жою, құжаттар арасындағы байланыстарға сілтемелерді өзгерту арқылы жүзеге асырылады. Ақпараттық жүйелердің пәндік облысының құрылымы өзгерген кезде ақпараттарды жаңарту кезінде деректер қорының сызбасы өзгертіледі.

5. *Қолданушыларға ақпараттық ресурстарды ұсыну.* Ақпараттық жүйелерді құрудың негізгі мақсаты – қолданушының сұранысы бойынша қажетті ақпараттық ресурстарды ұсыну және ақпараттық қажеттіліктерін қанағаттандыру. Ресурстарды қамтамасыз ету үшін pull және push-технологиялар қолдану арқылы жүзеге асырылуы мүмкін:

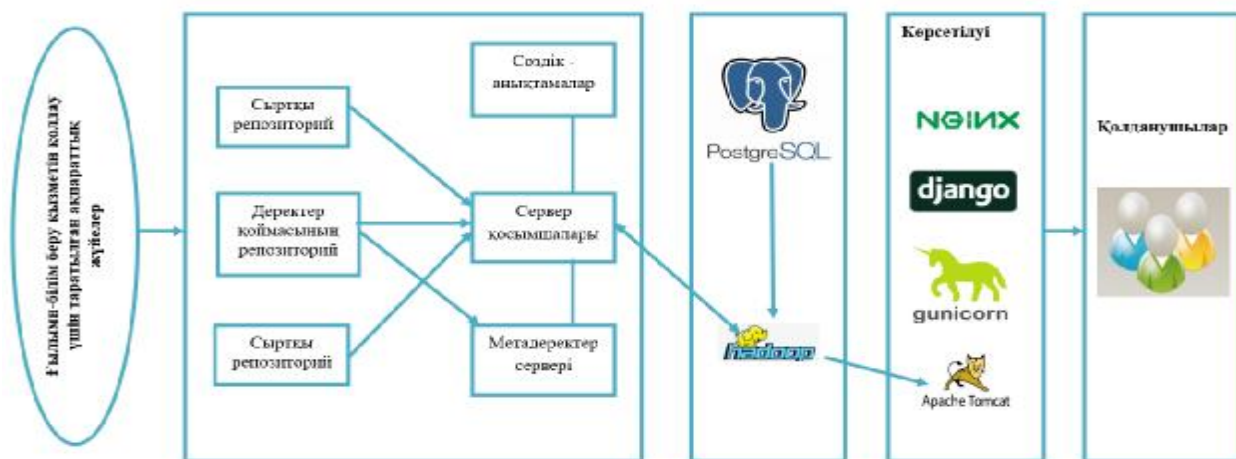
- Pull-технологиясы арқылы қолданушының бастамасы негізінде қолданушы интерфейтері арқылы ақпараттық ресурстарды іздеу және навигациясы механизмдерін қамтамасыз ете отырып жүзеге асырылады.

- Push-технологиясы – белгілі бір ережелерге сәйкес және белгілі бір қолданушы шеңбері негізінде қолданушылар арасында әртүрлі

ақпараттарды тарату мақсатында қолданылады. Жүйеде тіркелген қолданушыларға ғылыми-білім беру қызметінде жаңа құжаттардың түскендігін туралы хабарлайды [10, 15б.].

Аталған талаптарды қанағаттандыру үшін метадеректерді – ақпараттық ресурстар және оларға қол жеткізу қағидалары туралы құрылымдалған ақпаратты ұсыну және алмасу үшін ақпараттық қызмет немесе ортаны құру қажет. Қазіргі уақытта метадеректерді жинаумен және таратумен айналысатын көптеген ақпараттық орталықтар олардағы қорлармен алмасу мақсатында өзара іс-қимылды ұйымдастыруға белсенді қызығушылық танытуда. Әдетте, қорлардың мұндай үйлесімділігінің негізі метадеректерді ұсыну үшін стандартты әзірлеу болып табылады, сонымен бірге нормативтік-анықтамалық ақпарат массивтерін біріктіру болып табылады [11, 20б.].

Қойылған міндеттер шеңберінде ақпараттық жүйенің архитектурасы әзірленді (2-сурет), электрондық кітапхананың ресурстарын жүйелеу үшін деректер қоймасынан, репозиторийден, метадеректер серверінен, қосымшалар серверінен, анықтамалықтар сөздігінен тұратын көп деңгейлі ЭБ архитектурасы, сондай-ақ әзірленген архитектураның бағдарламалық іске асырылуы пайдаланылады.



2-сурет. Ғылыми-білім беру қызметін қолдауға арналған таратылған ақпараттық жүйенің архитектурасы

Үлкен көлемді деректерді сақтау және өңдеу, сондай-ақ машиналық (оның ішінде терең) оқытудың шығындық алгоритмдерінің ресурстарын қолдану қажеттілігіне байланысты архитектураны масштабталу және таратылу қажеттілігін ескере отырып әзірленді. 2-суретте көрсетілген осы жүйенің негізгі компоненттерін қарастырамыз:

Деректер қоймасы коллекциялар мен олардың құрылымы туралы қосымша метадеректерді сақтауға арналған. Олар мақалалардың, кітаптардың және т.б. электрондық нұсқаларының сақталуын қамтамасыз етеді және оларға сыртқы жүйелер мен пайдаланушыларға рұқсат береді [12, 178-8266.].

Электрондық кітапхана материалдарын пайдалану мазмұнды тиімді және дәл қарауды қамтамасыз ету үшін метадеректердің болуына байланысты. Сандық кітапханаға мазмұн қосылған кезде метадеректер жасалуы керек. Метадеректер мен деректер бір-бірімен қисынды түрде

байланыстырылуы керек және уақыт өте келе, платформалар мен шамадан тыс географиялық бөліну арасында логикалық байланысты басқарудың сенімді негізгі технологиясы болуы керек, барлығы желілік таратылған жүйеде жүзеге асырылады.

Қосымшалар серверінің тиімді жұмыс істеуі үшін жіктеу белгілері де, материалды жүйелеу және жіктеу жүргізілетін негізгі терминдер жиынтығы (тәртіп қатынастарымен) бар классификатор сөздіктер жиынтығын пайдалану қажет.

Сөздік анықтамалық-бұл құжат мазмұнын сипаттау сөздігін құрайтын терминдер жиынтығы. Оны мұрағатқа кіретін материалдарды санаттаудың стандартты әдісін жасау үшін стандарттау органдары қолдайды.

Анықтамалық сөздіктің көмегімен материалды нақты санаттау іздеу нәтижесінде бір немесе бірнеше электрондық кітапханаларда іздеуді ұйымдастырған кезде іздеу өрнегіне қатысты құжаттардың табылу ықтималдығын арттырады.

PostgreSQL - құрылымдық деректерді тұрақты сақтау орны ретінде қызмет етеді. Осы мәліметтер базасында сақталатын мәліметтердің негізгі түрлері: а) жаңалықтар және метадеректер; б) векторизацияны, лемматизация, тазалау нәтижелерін және т.б қоса алғанда, талдаудың әртүрлі негізгі бірліктері деңгейіндегі (лексема/сөз/фраза/сөйлем/мәтін) өңделген мәліметтер; в) Тақырыптық модельдеу нәтижелері; Жаңалықтарды әр түрлі критерийлер бойынша жіктеу нәтижелері (тоналдылық, саясаттану, әлеуметтік маңыздылық және т.б.)

HDFS (Hadoop Distributed File System) - есептеуіш кластердің түйіндері бойынша блоктық бөлінген ақпаратқа ағындық қолжетімділік мүмкіндігі бар үлкен көлемдегі файлдарды сақтауға арналған Hadoop үлестірілген файлдық жүйе, ол еркін аппараттық қамтамасыз етуден тұруы мүмкін. Hadoop Distributed File System, кез келген файлдық жүйе сияқты-бұл ішкі каталогтар мен файлдар салынған каталогтардың иерархиясы [12, 286].

Hadoop MapReduce негізгі тұжырымдамаларын келесідей тұжырымдауға болады:

- деректердің үлкен көлемін өңдеу/есептеу;
- ауқымдылығы;
- тапсырмаларды автоматты түрде параллельдеу;
- сенімсіз жабдықта жұмыс істеу;
- тапсырмаларды орындаудан бас тартуды автоматты түрде өңдеу.

Деректерді өңдеу:

Деректерді өңдеу үшін архитектураны әзірлеу кезінде келесі негізгі қажеттіліктер анықталды:

1) есептеулерді, оның ішінде бірнеше машиналарда параллельдеу мүмкіндігі;

2) деректерді өңдеу бойынша әртүрлі міндеттерді орындауды икемді жоспарлау мүмкіндігі;

3) міндеттерді нақты уақытта орындау мониторингінің мүмкіндігі, оның ішінде ерекшеліктер туралы жедел ақпарат беру;

4) қолданылатын құралдар мен технологиялардағы икемділік.

Суретте келесі ішкі жүйелер болады:

- Ішкі жүйе - цифрлық объектілер мен коллекцияларға қол жеткізу үшін пайдаланушылық және әкімшілік WEB-интерфейстерді, сондай-ақ ашық халықаралық стандарттар негізінде басқа кіші жүйелермен үйлесімділік интерфейстерін ұсынатын цифрлық объектілердің репозиторийі.

-Қызметкерлердің мақалаларын, олардың конференцияларға қатысуы және зерттеу жобаларын орындау туралы ақпаратты қамтитын ағымдағы ғылыми-зерттеу ақпаратын басқарудың ішкі жүйесі.

- Ішкі жүйе қолданушылар және әкімшілік интерфейстерді, сондай-ақ ашық халықаралық стандарттар негізінде басқа ішкі жүйелермен үйлесімділік интерфейстерін қамтитын болады.

- Apache Hadoop технологиясы негізінде таратылған ақпараттық ресурстарды біріктірудің ішкі жүйесі.

- Таратылған ақпараттық ресурстарға қол жеткізудің ішкі

жүйесі технологияның негізі-Nginx,

Djang, Apache Tomcat.

### Қорытынды

Ғылыми-білім беру қызметін қолдайтын таратылған ақпараттық жүйенің құжаттармен жұмыс жасаудың негізгі функционалдық талаптары тұжырымдалды. Бүгінгі таңда сапалы жаңа ғылыми ақпараттық жүйені құруға қажетті барлық компоненттер бар және нақты жүйелініп пысықталынды.

Ғылыми орталықтанған таратылған жүйелерді көп бөлігі ғылыми ақпаратты алмасудың бірыңғай ортасын құруға мүмкіндік береді.

«Бұл зерттеуді Қазақстан Республикасы Білім және ғылым министрлігінің Ғылым комитеті қаржыландырды (Грант № АР09057872)».

### Әдебиеттер тізімі

1. Федотов А. М. Методологии построения распределенных систем // Вычислительные технологии. 2006, Т. 11, Избранные доклады X Российской конференции. Распределенные информационно-вычислительные ресурсы. (DICR-2005), Новосибирск 6-8 октября 2005 г. С. 3–16.

2. Федотов А. М., Жижимов О. Л., Федотова О. А., Барахнин В. Б. Модель информационной системы для поддержки научно-педагогической деятельности // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2014. Т. 12, № 1. С. 89–101.

3. Федотов А. М., Барахнин В. Б., Жижимов О. Л., Федотова О. А. Технология создания корпоративных информационных систем учета трудов научных работников // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2011. Т. 9, вып. 2. С. 31–41.

4. Жижимов О. Л., Федотов А. М., Федотова О. А. Построение типовой модели информационной системы для работы с документами по научному наследию // Вестн. Новосиб. гос. ун-та. Сер. Информ. технологии. 2012. Т. 10, вып. 3. С. 5–14.

5. Когаловский М. Р. Метаданные, их свойства, функции, классификация и средства представления // Труды 14-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» — RCDL2012, Переславль-Залесский, Россия, 15-18 октября 2012 г.

6. S.K.Serikbayeva, D.A.Tussupov, M.A.Sambetbayeva, A.S. Yerimbetova, Tazhurekova ZH.K., Borankulova G.S. EduDIS construction technology based on Z39.50 protocol: Journal of Theoretical and Applied Information Technology, 31st May 2021 г.. - Т. - Vol. 99. No. 10, P.2224-2255

7. Serikbayeva S.K, Batyrhanov A.G., Sambetbayeva M.A., Sadirmekova Zh.B., Yerimbetova A.S. Development of technology to support large information storage and organization of reduced user access to this information: (IJACSA) International Journal of Advanced Computer Science and Applications, 2021 г.. - 7 : Т. 12. - стр. 493-503

8. 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. - Vol.86. - Issue 1. - pp.44-61.

9. Tusupov J.A., Fedotov A.M., Sambetbayeva M.A., Fedotova O.A., Sagnayeva S.K., Bapanov A.A., Tazhibayeva S.Z. Classification model and morphological analysis in multilingual scientific and educational information systems // Journal of Theoretical and Applied Information Technology. - 2016. - Vol.86, issue 1, - P.96-111.

10. Allan M Zarembski Some Examples of Big Data in Railroad Engineering//IEEE International Conference on Big Data, 2014. - 15 p.

11. Егоров А.А., Чернышова А.В. Исследование инструментов распределенной системы Hadoop//Конференция Современные информационные технологии в образовании и научных исследованиях (СИТОНИ-2017). Донецк, 2017 г. – Сборник научных трудов. – ДонНТУ.

## REFERENCES

1. Fedotov A. M. Methodology of distributed systems // Computational technologies. 2006, vol. 11, Selected reports of the X Russian Conference. Distributed information and computing resources. (DICR-2005), Novosibirsk, October 6-8, 2005, pp. 3-16.

2. Fedotov A.M., Zhizhimov O. L., Fedotova O. A., Barakhnin V. B. Model of information system for support of scientific and pedagogical activity // Vestn. Novosibirsk State University. Ser. Inform. technologies. 2014. Vol. 12, no. 1. pp. 89-101.

3. Fedotov A.M., Barakhnin V. B., Zhizhimov O. L., Fedotova O. A. Technology of creating corporate information systems for accounting of scientific workers ' works. Novosibirsk State University. Ser. Inform. technologies. 2011. Vol. 9, issue. 2. pp. 31-41.

4. Zhizhimov O. L., Fedotov A.M., Fedotova O. A. Building a typical model of an information system for working with documents on scientific heritage. Novosibirsk State University. Ser. Inform. technologies. 2012. Vol. 10, no. 3. S. 5-14.

5. Kogalovsky M. R. Metadata, their properties, functions, classification and means of representation // Proceedings of the 14th All-Russian Scientific Conference " Electronic Libraries: Promising Methods and Technologies, Electronic Collections — - RCDL2012, Pereslavl-Zalessky, Russia, October 15-18, 2012

6. S.K.Serikbayeva, D.A.Tussupov, M.A.Sambetbayeva, A.S. Yerimbetova, Tashurekova Zh.K., Borankulova G.S. EduDIS construction technology based on Z39.50 protocol: Journal of Theoretical and Applied Information Technology, 31st May 2021 г.. - Т. - Vol. 99. No. 10, P.2224-2255

7. Serikbayeva S.K., Batyrhanov A.G., Sambetbayeva M.A., Sadirmekova Zh.B., Yerimbetova A.S. Development of technology to support large information storage and organization of reduced user access to this information: (IJACSA)

International Journal of Advanced Computer Science and Applications, 2021 г.. - 7 :  
T. 12. - стр. 493-503

8. 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. - Vol.86. - Issue 1. - pp.44-61.

9. Tusupov J.A., Fedotov A.M., Sambetbayeva M.A., Fedotova O.A., Sagnayeva S.K., Bapanov A.A., Tazhibayeva S.Z. Classification model and morphological analysis in multilingual scientific and educational information systems // Journal of Theoretical and Applied Information Technology. - 2016. - Vol.86, issue 1, - P.96-111.

10. Allan M. Zarembski Some Examples of Big Data in Railroad Engineering/ / IEEE International Conference on Big Data, 2014. - 15 p.

11. Egorov A. A., Chernyshova A.V. Research of tools of the distributed Hadoop system//Conference Modern Information Technologies in Education and Research (SITONI-2017). Donetsk, 2017-Collection of scientific papers. - DonNTU.

## **ФОРМИРОВАНИЕ ТРЕБОВАНИЙ К РАСПРЕДЕЛЕННЫМ ИНФОРМАЦИОННЫМ СИСТЕМАМ ПОДДЕРЖКИ НАУЧНО- ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ**

*Серикбаева С.К.<sup>1</sup>, PhD докторант  
Тусупов Д.А.<sup>1</sup>, д.ф-м.н., профессор  
Самбетбаева М.А.<sup>1,2</sup>, PhD доктор*

<sup>1</sup>*Евразийский национальный университет им. Л.Н.Гумилева,  
Мунайтпасов К., 5, г.Нур-Султан, 010011, Казахстан, [inf\\_8585@mail.ru](mailto:inf_8585@mail.ru)*  
<sup>2</sup>*Институт информационных и вычислительных технологий КН МОН РК,  
Пушкина 125, г.Нур-Султан, 010011, Казахстан*

### **Аннотация**

Для повышения эффективности исследования необходимо иметь доступ к систематическим информационным ресурсам научной работы. Поэтому в любой области науки он начинается с исследования, поиска научной информации, но с ростом количества научных статей, книг, монографий, патентов поиск информации становится все сложнее. Создание единой информационной системы, позволяющей ученым быстро знакомиться с результатами других научных исследований и предотвращать их дублирование. В статье рассматриваются технологические приемы распределенных информационных систем, обеспечивающих научную и образовательную деятельность. Описаны основные задачи создания модели распределенной информационной системы, поддерживающей научно-образовательную деятельность, функциональные возможности модели, понятие метаданных и требования к профилю метаданных. Определены задача, предметная область, субъекты, объекты, основная функциональность информационной системы,

приводится список основных видов информационных ресурсов. В работе выполнен анализ функциональных требований, предъявляемых к таким системам.

**Ключевые слова:** распределенная информационная система, метаданные, Z39.50, LDAP, PostgreSQL, HDFS.

## **FORMATION OF REQUIREMENTS FOR DISTRIBUTED INFORMATION SYSTEMS TO SUPPORT SCIENTIFIC AND EDUCATIONAL ACTIVITIES**

*Serikbayeva S. K.<sup>1</sup>, Ph.D. doctoral student*

*Tussupov J. A.<sup>1</sup>, doctor of ph-m.s., Professor*

*Sambetbayeva M. A.<sup>1,2</sup>, Ph.D. doctor*

*<sup>1</sup>L. N. Gumilyov Eurasian National University,*

*Munaytpasov K., 5, Nur-Sultan, 010011, Kazakhstan, [inf\\_8585@mail.ru](mailto:inf_8585@mail.ru)*

*<sup>2</sup>Institute of Information and Computational Technologies CS MES RK,*

*Pushkin 125, Nur-Sultan, Kazakhstan*

### **Abstract**

To increase the effectiveness of research, it is necessary to have access to systematic information resources of scientific work. Therefore, in any field of science, it begins with research, the search for scientific information, but with the growing number of scientific articles, books, monographs, patents, the search for information becomes more and more difficult. Creating a unified information system that allows scientists to quickly get acquainted with the results of other scientific research and prevent their duplication. The article discusses the technological techniques of distributed information systems that provide scientific and educational activities. The main tasks of creating a model of a distributed information system that supports scientific and educational activities, the functional capabilities of the model, the concept of metadata, and the requirements for the metadata profile are described. The task, subject area, subjects, objects, the main functionality of the information system are defined, a list of the main types of information resources is provided. The paper analyzes the functional requirements for such systems.

**Keywords:** distributed information system, metadata, Z39.50, LDAP, PostgreSQL, HDFS.