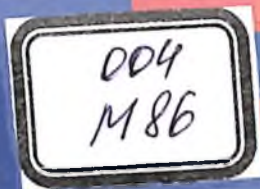


Б.Б.МУМИНОВ



**НОРАВШАН ВА СТОХАСТИК
АХБОРОТ МУҲИТЛАРИДА
МАЪЛУМОТЛАРНИ ИЗЛАШ ВА
ҚАЙТА ИШЛАШ МОДЕЛЛАРИ,
УСУЛЛАРИ**

МУҲАММАД АЛ-ХОРАЗМИЙ НОМИДАГИ
ТОШКЕНТ АХБОРОТ ТЕХНОЛОГИЯЛАРИ
УНИВЕРСИТЕТИ

Б.Б.МЎМИНОВ

НОРАВШАН ВА СТОХАСТИК
АХБОРОТ МУҲИТЛАРИДА МАЪЛУМОТЛАРНИ
ИЗЛАШ ВА ҚАЙТА ИШЛАШ МОДЕЛЛАРИ,
УСУЛЛАРИ



Тошкент
«Университет»
2018

Норавшан ва стохастик ахборот муҳитларида маълумотларни
излаш ва қайта ишлаш моделлари, усуллари. Б.Б.Мўминов.
–Т.: «Университет», 2018, 152 бет.

УЎК: 004.681.4:517(575.1)

КБК

Мазкур монографияда маълумот излашнинг ривожланиш босқичлари ва фундаментал асослари ва сўровларни қайта ишлаш усуллари, корпоратив ахборот муҳитларида МИТни яратиш муаммолари, САМда МИТнинг элементлари ва моделлаштириш омиллари, ресурс ва термин частоталарининг вазнини ҳисоблаш усуллари, МИТда мантикий семантик излаш модели, САМда МИТнинг дастурий модулини яратиш, норавшан сўровларни қайта ишлаш усули, МИИТда тегишлилик функцияларининг синфларини танлаш ва ИИда семантик ядро модели ва алгоритмларини ишлаб чиқиш, МИИТда норавшан излаш модели ва қоидаларни ишлаб чиқиш, МИТнинг дастурий модулини яратиш, корпоратив тармоқда МИҚИнинг FSV технологияси, сўровларни қайта ишлаш ҳамда маълумотларни излаш ва тақдим қилиш инструментал дастурий модулини яратиш, FSV технологиясини жорий қилишда МБсини ишлаб чиқиш, ДҚларга жорий қилиш ва самарадорлиги баён қилинган.

Техника фанлари доктори, профессор М.А.Рахматуллаев тахрири остида.

Тақризчилар:



Т.Ф.Бекмуродов – Ўзбекистон Республикаси Фанлар академияси академиги, т.ф.д., профессор, Муҳаммад ал-Хоразмий номидаги ТАТУ хузуридаги Ахборот-коммуникация технологиялари илмий-инновацион маркази бош илмий ходими

О.Ж.Бобамуродов – т.ф.д., Муҳаммад ал-Хоразмий номидаги ТАТУ “Ахборот технологияларининг дастурий таъминот” кафедраси мудири

Монография Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялар университети Илмий-техник кенгаши-нинг 2018 йил 27 ноябрдаги 6-сонли қарори билан нашрга тавсия этилди.

ISBN 978-9943-5601-0-9

© «Университет» нашриёти, Тошкент – 2018й.

ҚИСҚАЧА МАЗМУНИ

Дунёда рақамли ахборотга бўлган эҳтиёж кун сайин ортиб бориши ахборотлар ҳажми катталашшига олиб келмоқда. Бу эса катта ҳажмли рақамли ахборот тизимлари технологияларини такомиллаштириш масалаларини келтириб чиқармоқда. Бугунги кунда электрон ресурсларнинг катта ҳажмини ахборот-ресурс марказларида кўриш мумкин. Бу, ўз навбатида, ахборот-ресурс марказлари фаолиятини ахборот технологиялари орқали такомиллаштириш, сифатли хизмат кўрсатиш тизимларини автоматлаштириш масалаларининг долзарб муаммоларидан бири бўлиб ҳисобланади. Турли ахборот тизимларида, корпоратив тармоқларда маълумотларни излаш ва қайта ишлаш масалалари доимий долзарб муаммолардан ҳисобланиб келинади. Айниқса, ахборот-ресурс марказларида фақат аниқ метамаълумотлар орқали маълумотларни излаб топиш мумкин. Аммо бугунги кунда фойдаланувчилар сони ва ахборот ҳажмининг ошиб бориши, локал тилларда сўровлар яратилиши ортикча маълумотларни таҳлил қилиш учун вақт сарфлашни талаб этмоқда. Шу нуқтаи назардан қараганда, танланган мавзу долзарб бўлиб ҳисобланади.

Биринчи боб маълумот излашнинг ривожланиш босқичлари ва фундаментал моделлари, усулларининг таҳлилига, ахборот муҳитлари хусусиятлари ва классификацияларини аниқлашга, корпоратив ахборот-кутубхона тизимларида маълумотларни излаш ва қайта ишлаш тизимини яратиш муаммоларини ўрганишга бағишланган. Корпоратив ахборот муҳитларида маълумотларни излаш тенденциялари асосида маълумотларни излаш ва қайта ишлашнинг концептуал схемаси яратилган. Схема асосида маълумотларни излаш ва қайта ишлашнинг асосий модуллари ва элементлари аниқланган. Таниқли олимларнинг маълумотларни излаш бўйича олиб борган тадқиқотлари, мантиқий излаш модели, тартибли индекслаш модели, луғатларга асосланган модели, вариантли сўровларни яратиш усуллари, *k*-граммли ва масофали таҳрирлаш усуллари анъанавий излаш тизимлардаги жорий қилинганлигининг таҳлили баён қилинган. Корпоратив ахборот-кутубхона тизимларининг ривожланиш истикболлари, яратиш механизмлари, имкониятлари, архитектуралари, яратилган вариантлари бугунги кунда янги замонавий технологияларэб билан мослашиши мураккаблиги аниқланган. Корпоратив ахборот-кутубхона тизимларининг замонавий вариантларини яратиш учун 3 босқичли «Мижоз-сервер» архитектурасини танлаш таклиф қилинган. Фан ва таълимга оид ахборот

муҳитларининг элементлари тузилмаси асосида ахборот муҳитининг бошқарув схемаси ва унинг модуллари ишлаб чиқилган. Ушбу модулларнинг имконияти ва маълумотларни излаш ҳамда қайта ишлаш учун маълумот тузилмаларга асосланиб, фан ва таълимга оид ахборот муҳитлари 3 та синфга детерминанлашган, стохастик ва норавшан ахборот муҳитларига ажратилган. Ахборот муҳитларининг фаркланиши етарлича асослаб берилган, улар реал муҳитда доимий ўзаро таъсирда бирга бўлиши келтирилган. Корпоратив ахборот муҳитларида маълумотларни излаш тизимларни яратиш тенденциялари асосида мавжуд муаммолар аниқланган.

Иккинчи боб стохастик ахборот муҳитида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ва дастурий модулини яратишга бағишланган бўлиб, маълумотларни излаш тизимида метамаълумотлар асосида параметрли излашни такомиллаштириш ва унда майдонга мос индекслаш масаласи, ўқув мисолларидан фойдаланиш келтирилган. Шунингдек, электрон ресурс ва термин частоталарини ҳисоблаш, эвристик ёндашувлар оркали сўровга мос маълумотларни излашни амалга ошириш усуллари такомиллаштирилган ва алгоритмлари ишлаб чиқилган. Стохастик ахборот муҳитларида маълумотларни излашни тезлаштириш учун излаш мисолларидан фойдаланиб, мантикий семантик излаш модели ишлаб чиқилган. Ушбу модел асосида сўровларнинг ўхшашлиги ва боғланганлигини аниқлаш, уларнинг семантик маълумотларини яратиш мумкин. Стохастик ахборот муҳитларида маълумотларни излаш тизимлари учун инструментал дастурий модулнинг функционал имконияти тузилмаси, IDEF0 моделлари ва уларнинг таснифи, реляцион маълумотлар тузилмасининг IDEF1x модели ва унинг объектлари, хусусиятлари очиб берилган. Инструментал дастурий модулнинг IDEF1x моделининг моҳият-муносабат ҳолатлари ва MVC асосида яратилгани баён қилинган.

Учинчи боб норавшан ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ва дастурий модулини яратиш масалаларига бағишланган бўлиб, норавшан сўровларни қайта ишлаш модели Лутфе Заденинг лингвистик ўзгурувчиси моделидан фойдаланиб, қайта ишлаш таклиф қилинган. Параметрли тегишлилик функцияларини шакллантириш усулида норавшан термларга мос параметрларни танлаш ва шулар асосида уларни SQL стандартига шакллантириш процедуралари ишлаб чиқилган. Маълумотларни интеллектуал излашда норавшан термларга мос синфлари ва тегишлилик функцияларининг кўринишларини танлаш

бўйича тавсиялар келтирилган. Маълумотларни интеллектуал излаш ва қайта ишлаш учун семантик ядро яратишнинг 3 усули, излаш ва қайта ишлаш натижасида автоматик тарзда, электрон ресурсларга берилган ҳаволалар, цитатасини қайта ишлаш орқали, эксперт гуруҳлари ёрдамида семантик ядрони яратиш усули ва алгоритмлари ишлаб чиқилган. Бу алгоритмлар асосида яратиладиган семантик ядро модели, шунингдек, норавшан билимлар базасининг қоидалари тизими Мамдани қоидаси асосида ишлаб чиқилган. Норавшан ахборот муҳитида маълумотларни излаш тизимларининг 3 та тоифага эга IDEF0 моделлари ҳамда реляцион маълумотлар тузилмасининг IDEF1x модели, моҳият-муносабат ҳолатлари ишлаб чиқилган.

Тўртинчи боб корпоратив тармоқда маълумотларни дизлаш ва қайта ишлаш учун маълумотларни излаш тизимининг дастурий таъминотини яратиш ва жорий қилиш масалаларига бағишланган бўлиб, маълумотларни излаш тизими учун FSV технологияси, архитектураси ва F-сўровларни шакллантиришнинг дастурий модули, S-маълумотларни излашнинг дастурий модули, V-тақдим қилишнинг дастурий модулларининг функционал тузилмаси, IDEF моделлари ва функционал имкониятлари ишлаб чиқилган. Таклиф қилинган FSV технологиясини дастурий таъминотларда жорий қилиш учун маълумотлар базаси лойиҳаси, маълумотлар тузилмалари ишлаб чиқилган. FSV технологиясини тажриба синовдан ўтказиш мақсадида бир хил стандарт асосида фаолият кўрсатадиган дастурий тизимлар танлаб олинди, 3 мезон асосида таққосланган ҳамда натажалар диаграмма кўришида келтирилган. FSV технологияси жорий қилинган ARMAТ++ дастурий таъминотининг ёрдамчи ускуналари асосида ижтимоий ва иқтисодий самарадорлигини аниқлаш параметрлари ишлаб чиқилган.

Тадқиқот Ўзбекистон Республикаси Инновацион ривожланиш вазирлиги томонидан тасдиқланган А5-055-«Корпоратив тармоқларнинг норавшан ва стохастик ахборот муҳитларида сўровларни қайта ишлаш» (2015-2017); А5-066-«Битирув малакавий ва магистрлик ишлари натижалари ва ютуқлари билан алмашиш имкониятини берувчи виртуал биржа платформасини ишлаб чиқиш» (2015-2017); И2016-4-15-«Кутубхона ресурсларидан мобил алоқа асосида тармоқдан фойдаланишнинг дастурий таъминотини яратиш ва татбиқ қилиш» (2016-2017); И2017-4-4-«Ахборот тизимларида маълумотларга интеллектуал ишлов бериш, излаш моделларини яратиш ва жорий қилиш» (2017-2018) лойиҳалари доирасида бажарилган.

TASKS AND RESULTS OF THE RESEARCH

The aim of the research work. Development of models, methods for searching and processing data in fuzzy and stochastic information environments, as well as creating software for corporate information and library systems.

The objectives of the research:

the development of the classification of information environments, the scheme for searching and processing data, the stages of search technologies, as well as the method, algorithm and software module for data storage in corporate information and library systems;

the development of a method for reducing time and basic elements of data retrieval in information environments of corporate networks;

development of a method and software module for designing parametric accessory functions for fuzzy queries;

development of a method, algorithm and software module for searching and processing data in fuzzy and stochastic information environments;

development of an algorithm and a software module for creating a knowledge base for searching and processing data in fuzzy information environments;

development of modules, algorithms, databases, software architecture, IDEF models of the data retrieval system in corporate information and library systems.

The object of the research work are sets of fuzzy and stochastic information environments, electronic scientific and educational resources of corporate information-library systems.

The scientific novelty of the study is as follows:

software module has been developed and the method of storing data for the search and processing of data in corporate information and library systems has been improved;

the method of logical semantic search of data in stochastic information environments of a corporate network is developed;

the program module was developed and a method for designing parametric membership functions based on a linguistic variable for fuzzy queries was developed;

the model, an algorithm and a software module of the semantic kernel for a data retrieval system in fuzzy and stochastic information environments based on logical semantics;

the model, method, algorithm and program module for creating the knowledge base are developed in accordance with the Mamdani rule of the intelligent data retrieval system in fuzzy information environments;

developed FSV technology to create a data retrieval system in corporate networks, as well as its architect, IDEF model and algorithms, database project;

Implementation of the research results. Based on FSV-technology, created on the basis of methods of searching and processing data in fuzzy and stochastic information environments:

in the corporate networks of the main library of the Academy of Sciences, the Republican Scientific Library of Agriculture, the library of the Intellectual Property Agency, a tool program module for intellectual search and data processing, a method for determining the preliminary rating of an electronic resource on the basis of references, a semantic core, a knowledge base for ARMAT ++ software, corporate information -Library system (certificate of the Ministry for the Development of Information Technology and Communications from September 12, 2017 No. 33-8/6059). Based on the results of the research, the efficiency of work on the categorization of corporate information and library systems, the remote search by the readers of the necessary data in a single electronic general catalog, the speed of search by querying are increased. Effective use of interactive services in libraries, reduction of the time of issuance and receipt of books, efficient search and processing of data by 5% and a level of rendering services increased by 10%.

In the "ELine-press" LLC, a method of logical semantic search has been introduced, an instrumental program module for the preparation of static and dynamic reports through the intellectual search and processing, the use of the method of logical semantic search (reference of the Ministry for the Development of Information Technologies and Communications of September 12, 2017). No. 33-8/6059). Based on the research results, the conditions for the formation of a database of periodic sources, indexing, certification, search and processing of data, preparation of static and dynamic reports, and the integration of the database with the corporate information library system have been increased by 10%.

ЗАДАЧИ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Цель исследования – разработка моделей, методов поиска и обработки данных в нечетких и стохастических информационных средах, а также создание программного обеспечения для корпоративных информационно-библиотечных систем.

Задачи исследования:

разработка классификации информационных сред, схемы поиска и обработки данных, этапов поисковых технологий, а также метода, алгоритма и программного модуля хранения данных в корпоративных информационно-библиотечных системах;

разработка метода вычисления предварительного рейтинга электронного ресурса в корпоративных информационно-библиотечных системах;

разработка метода сокращения времени и основных элементов поиска данных в информационных средах корпоративных сетей;

разработка метода и программного модуля проектирования параметрических функций принадлежности для нечетких запросов;

разработка метода, алгоритма и программного модуля поиска и обработки данных в нечетких и стохастических информационных средах;

разработка алгоритма и программного модуля создания базы знаний для поиска и обработки данных в нечетких информационных средах;

разработка модулей, алгоритмов, баз данных, архитектуры программного обеспечения, IDEF моделей системы поиска данных в корпоративных информационно-библиотечных системах.

Объектом исследования являются множества нечетких и стохастических информационных сред, электронных научно-образовательных ресурсов корпоративных информационно-библиотечных систем.

Научная новизна исследования заключается в следующем:

разработан программный модуль и усовершенствован метод хранения данных осуществления поиска и обработки данных в корпоративных информационно-библиотечных системах;

разработан метод логического семантического поиска данных в стохастических информационных средах корпоративной сети;

разработан программный модуль и усовершенствован метод проектирования параметрических функций принадлежности, основанных на лингвистической переменной для нечетких запросов;

разработаны модель, алгоритм и программный модуль семантического ядра для системы поиска данных в нечетких и стохастических информационных средах, основанных на логической семантике;

разработаны модель, метод, алгоритм и программный модуль создания базы знаний в соответствии с правилом Мамдани системы интеллектуального поиска данных в нечетких информационных средах;

разработаны FSV-технологии создания системы поиска данных в корпоративных сетях, а также его архитектуры, IDEF модели и алгоритмов, проект базы данных;

Внедрение результатов исследования. На основе FSV-технологии, созданной на базе методов поиска и обработки данных в нечетких и стохастических информационных сред:

- в корпоративные сети основной библиотеки Академии наук, Республиканской научной библиотеки сельского хозяйства, библиотеки Агентства интеллектуальной собственности внедрены инструментальный программный модуль интеллектуального поиска и обработки данных, метод определения предварительного рейтинга электронного ресурса на основе ссылок, семантическое ядро, база знаний для программного обеспечения ARMAT++ корпоративной информационно-библиотечной системы (справка Министерства по развитию информационных технологий и коммуникаций от 12 сентября 2017г. №33-8/6059). На основе результатов исследования повышены эффективность работ по каталогизации корпоративных информационно-библиотечных систем, удаленного поиска читателями необходимых данных в едином электронном общем каталоге, скорость поиска путем формирования запросов. Обеспечено эффективное использование интерактивных услуг в библиотеках, сокращение времени процессов выдачи и приема книг, а также повышена эффективность поиска и обработки данных на 5% и уровень оказания услуг повышен на 10%.

- в ООО "ELine-press" внедрены метод логико-семантического поиска, инструментальный программный модуль подготовки статистических и динамических отчетов путем интеллектуального поиска и обработки, использования метода логико-семантического поиска (справка Министерства по развитию информационных технологий и коммуникаций от 12 сентября 2017 г. №33-8/6059). На основе результатов исследования созданы условия для формирования базы данных периодических источников, индексации, сертификации, поиска и обработки данных, подготовки статических и динамических отчетов, а также путем интеграции базы данных с корпоративной информационно-библиотечной системой повышен уровень оказываемых ими услуг на 10%.

КИРИШ

Тадқиқотнинг долзарблиги ва зарурати. Жаҳонда аҳолининг электрон ресурсга, ахборотга бўлган эҳтиёжларини қондириш учун корпоратив ахборот-ресурс марказларида маълумотларни интеллектуал излаш тизимларини яратишга катта эътибор қаратилмоқда. «Кун сайин катталашиб бораётган ахборот-ресурсларидан фойдаланиш технологиялари, инсоният учун бугуннинг муҳим масалаларидан бири бўлиб қолмоқда»¹. Бу борада корпоратив тармоқларда маълумотларни семантик ва норавшан қоидалар асосида излаш ҳамда қайта ишлаш технологияларини ишлаб чиқиш ва такомиллаштириш, маълумотлар ва мантиқий билимлар базасини лойиҳалаштириш муҳим аҳамият касб этади. Ушбу соҳада хорижий мамлакатларда, жумладан, АҚШ, Германия, Япония, Хитой, Австрия, Франция, Греция, Россияда ахборот-ресурслардан фойдаланиш учун маълумотларни излаш тизимининг математик ва дастурий таъминотини яратишга бағишланган илмий-амалий тадқиқотлар олиб борилмоқда.

Жаҳонда корпоратив тармоқларда маълумотларни излаш тизимларини яратиш, маълумотлар ва семантик билимлар базасини лойиҳалаш, маълумотларни интеллектуал таҳлил қилиш, излаш, сақлаш алгоритмлари ва дастурий модулларини яратиш, такомиллаштириш муҳим аҳамиятга эга. Бу борада корпоратив тармоқларнинг норавшан ва стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усулларини ишлаб чиқиш, лойиҳалаштириш, маълумотларни излаш тизимини яратувчи дастурий таъминотни яратиш ва амалиётга тадбиқ этиш муҳим вазифалардан ҳисобланади.

Республикамиз мустақилликка эришгандан буён жамиятни ахборотлаштириш даражасини оширишда ахборот технологиялари, аппарат-дастурий воситаларни интеллектуал бошқариш тизимлари асосида ахборот-ресурс марказларининг маълумотлар базасини шакллантириш, корпоратив ахборот-кутубхона тизимлари, миллий контент яратишга алоҳида эътибор қаратилмоқда. Бу борада маълумотларни излаш тизимларида мантиқий излаш дастурий комплекслари ва маълумотлар базаларини ишлаб чиқиш йўлга қўйилди. Шулар билан бир қаторда, маълумотларни интеллектуал излаш ва тақдим қилиш жараёнларини такомиллаштириш, миллий тилда тузилган сўровларни қайта ишлаш технологияларини ишлаб

¹ <http://users.ics.forth.gr/~tzitzik/sarg/index.html>

чиқиш талаб этилмоқда. 2017-2021 йилларда Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегиясида «Миллий контентни ривожлантириш, давлат тилидаги таълим, илмий-маърифий, ёшлар эhtiёжларига мос замонавий ахборот-ресурсларини, мультимедиа маҳсулотларини яратиш ва тарғиб қилиш механизмларини такомиллаштириш, ... ахборот-коммуникация технологияларини жорий этиш»² вазифалари белгиланган. Бу борада корпоратив ахборот-кутубхона тизимларида электрон ресурсларни шакллантириш, интеллектуал излаш тизимларини ишлаб чиқиш муҳим масалалардан бири ҳисобланади.

Ўзбекистон Республикасининг «Ахборот эркинлиги принциплари ва кафолатлари тўғрисида»ги (2002 йил), «Ахборот-кутубхона фаолияти тўғрисида»ги (2011 йил) қонунлари, Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги ПФ-4947-сонли, 2017 йил 30 июндаги «Республикада ахборот технологиялари соҳасини ривожлантириш учун шарт-шароитларни тубдан яхшилаш чора-тадбирлари тўғрисида»ги ПФ-5099-сонли фармонлари, Вазирлар Маҳкамасининг 2011 йил 5 июлдаги «Ахборот-кутубхона ва ахборот-ресурс марказларида ва кутубхоналарда тўлиқ матнли электрон ахборот-кутубхона ресурслари фондини яратиш тўғрисида»ги 198-сонли, 2017 йил 14 августдаги «Интернет жаҳон ахборот тармоғида миллий контентни янада ривожлантириш чора-тадбирлари тўғрисида»ги 625-сонли қарорлари ҳамда мазкур фаолиятга тегишли бошқа меъёрий-ҳуқуқий ҳужжатларда белгиланган вазифаларни амалга оширишга муайян даражада хизмат қилади.

Тадқиқотнинг республика фан ва технологиялари ривожланишининг устувор йўналишларига мослиги. Мазкур тадқиқот республика фан ва технологиялар ривожланишининг IV. «Ахборотлаштириш ва ахборот-коммуникация технологияларини ривожлантириш» устувор йўналишлари доирасида бажарилган.

Тадқиқот бўйича хорижий илмий тадқиқотлар шарҳи³. Норавадан ва стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлашнинг математик асослари, моделлари, усуллари

² Ўзбекистон Республикаси Президентининг 2017 йил 7 февралдаги «Ўзбекистон Республикасини янада ривожлантириш бўйича Ҳаракатлар стратегияси тўғрисида»ги ПФ-4947-сонли Фармони.

³ Тадқиқот мавзуси бўйича хорижий илмий тадқиқотлар шарҳи <http://search.ebscohost.com>, <http://link.springer.com>, <https://databases.library.jhu.edu>, Thomson Reuters, EBSCO information services, ProQuest, Nature, Oxford University Press, Cambridge University Press, eIFL ва бошқа манбалар асосида ишлаб чиқилган.

асосида дастурий таъминотлар яратиш, маълумотлар ва билимлар базаларини лойиҳалаштириш бўйича жаҳоннинг етакчи илмий марказлари, жумладан, Center of Excellence in Space Data and Information Sciences (АҚШ), European Research Consortium of Informatics and Mathematics (Испания), Fern Universities Hagen (Германия), Frederick University (Кипр), Graz University of Technology (Австрия), Gulf University for Science & Technology (Кувайт), Institute for Computer Science and Control (Венгрия), National Institute of Informatics (Япония), National Technical University of Athens (Греция), University of Novi Sad (Сербия), University of Science and Technology of China (Хитой), Ҳисоблаш технологиялари институти (Россия), Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технологиялари университетига (Ўзбекистон) кенг қамровли илмий тадқиқот ишлари олиб борилмоқда.

Маълумотларни излаш ва қайта ишлаш моделлари, усулларини яратиш, маълумот излаш тизимларини такомиллаштиришга оид жаҳонда олиб борилган тадқиқотлар натижасида қатор, жумладан, куйидаги илмий натижалар олинган: терминларга асосланган веб иловаларда матнларни излашда тилга оид тўсиқларни бартараф этиш алгоритмлари яратилган (National University of Distance Education, Испания); XML синфлаш ва асосий элементларни таҳлиллаш ёрдамида излаш, замонавий кутубхоналарда матнли излаш, автоматик таҳлил қилиш усуллари ишлаб чиқилган (New Jersey Institute of Technology University Heights, North Carolina State University Library, АҚШ); маълумотларни излашда агрегатив операторлардан фойдаланиб бирлаштириш моделлари яратилган (University of Pierre and Marie Curie, Франция); веб контентларда маълумотларни излаш учун Яндекс, Google маълумотларни излаш тизимлари ва SEO тамойиллари ишлаб чиқилган (АҚШ, Хитой, Россия, Ҳиндистон, Куба, Малайзия).

Дунёда маълумотларни интеллектуал излаш ва қайта ишлаш масалаларини тадқиқ қилиш учун электрон ресурслар яратиш, жараёнларни моделлаштириш ва юқори самарали бошқариш тизимларини яратиш бўйича қатор, жумладан, куйидаги йўналишларда тадқиқотлар олиб борилмоқда: информатика ва кутубхона соҳаларидаги маълумотларни излаш ва қайта ишлашда норавшан тўпламлар назарияси асосида интеллектуал дастурий модуллар яратиш; Data mining, Text mining, Big Data модель ва усулларига асосланган маълумотларни мантикий излаш, сақлаш усулларини яратиш ва маълумотларни излаш тизимлари билан интеграция

килиш усул ҳамда алгоритмларини ишлаб чиқиш; SMART ва SEO тамойиллари асосида семантик ядро, билимлар базасини яратиш ҳамда миллий тилларга асосланган маълумотларни излаш моделлари ва дастурий таъминотини ишлаб чиқиш.

Муаммонинг ўрганилганлик даражаси. Компьютер тармоқларида маълумотларни интеллектуал излаш ва қайта ишлаш муаммоларига математик статистика, норавшан тўпламлар назарияси, Data mining, Text mining ва Big Data асосида моделлар қуриш ва усулларини яратиш, SMART ва SEO тамойилларини ривожлантириш ва ахборот-ресурсларини жорий қилиш муаммолари бўйича тижорат компаниялари Bing, Google, Rambler, Search.Mail.ru, Yandex ва дунёнинг машҳур олимлари раҳбарлигида тадқиқот ишлари олиб борилмоқда, жумладан, Y.Baeza, M.Boughanem, O.Bouidghaghen, C.Carpineto, T.J.Dickey, J.liu, A.Z.Lotfi, J.Wang, Н.Ш.Виктор, маълумотларни излаш тизимларининг таҳлили, муаммо ва ечимларига бағишланган тадқиқот ишларини Sulton, Harman, Krovetz, Hull, Шарқ мамлакатларининг тиллари асосида маълумотларни излаш бўйича Lunde, орфографик хатоларни тузатишнинг эҳтимолли моделларини яратиш бўйича Kernigan, Bill ва Моогелар, сўровларни қайта ишлаш модель ва алгоритмлари асосида сўровларни қайта тузиш ва татбиғи бўйича Cucerzan ва Brillар, республикамызда луғатда матнларни қайта ишлаш муаммолари бўйича М.И.Бадалов ва А.М.Мирзамовлар тадқиқот ишлари олиб боришган.

Маълумотларни интеллектуал таҳлиллаш ва қайта ишлаш бўйича Л.Заде, А.Холмблад, Б.Коскон, Д.Дюбуа, А.Парада, Е.Мамдани, В.В.Рыбин, Б.Лю, М.Джамшиди, Б.Фазлоллахин, Э.Мендельсон, А.Леоненков, республикамызда компьютер тармоқларида интеллектуал таҳлил асосида қарор қабул қилиш бўйича академиклар Т.Ф.Бекмурадов, М.М.Камилов, Ф.Б.Абуталиев ва Н.Р.Юсупбеков, Д.Мухаммадиева, М.А.Рахматуллаев, М.И.Бадалов, Р.А.Алиев, Н.А.Игнатов каби олимларимизнинг тадқиқот ишларида ўрганилган. Электрон ресурслар яратиш ва унда излаш масалалари бўйича назарий ва амалий тадқиқот ишлари олиб борган олимларга Henriette D.Avram, Hugh C.Atkinson, Donald S.Culbertson, корпоратив ахборот-кутубхона тизимлари учун модель ва алгоритмлар А.С.Крауш, М.А.Рахматуллаев ва Д.Ю.Копытков томонидан бажарилган илмий-амалий тадқиқот ишларида, библиографик ахборотларга ишлов бериш моделлари ва алгоритмлари Я.Л.Шрайберг, Ф.С.Воройский, М.А.Рахматуллаев, У.Ф.Каримов, R.P.Rodgers ва

А.А.Леонтьевларнинг ишларида, республикамизда корпоратив ахборот-кутубхона тизимларини яратиш ҳамда марказлашган каталоглаштириш тизимларининг модель ва алгоритмларини яратиш, жорий қилиш, маълумотларни излаш, хавфсизлигини таъминлаш, асосан, М.А.Раҳматуллаев раҳбарлигида ўрганилган.

Шунингдек, маълумотларни излаш масалалари доирасида фан ва таълимга оид корпоратив ахборот-кутубхона тизимларининг ахборот муҳитларини тадқиқ қилиш, маълумотларнинг тузилиши, хусусиятлари, сўровларни шакллантириш ва қайта ишлаш, маълумотлар йиғиш, мантиқий семантик ядро яратиш, электрон ресурслар рейтингини ҳисоблаш, маълумотларни интеллектуал излаш ва қайта ишлаш, тақдим қилиш моделлари ва усулларини яратиш, маълумотларни излаш тизимининг дастурий таъминотини ишлаб чиқишга бағишланган илмий изланишлар ҳозирги кунда етарли даражада ўрганилмаган.

Тадқиқотнинг мақсади корпоратив ахборот-кутубхона тизимининг норавшан ва стохастик ахборот муҳитларида маълумотлар излаш ва қайта ишлаш моделлари, усулларини ишлаб чиқиш ва дастурий таъминотини яратиш.

Тадқиқотнинг вазифалари:

корпоратив ахборот-кутубхона тизимларининг ахборот муҳитларининг классификациясини, маълумотларни излаш ва қайта ишлашнинг схемасини, излаш технологиялари босқичлари ҳамда маълумотларни сақлаш усул, алгоритми ва дастурий модулини ишлаб чиқиш;

корпоратив тармоқнинг ахборот муҳитларида маълумотлар излашнинг асосий элементлари асосида излаш вақтини камайтириш моделини ишлаб чиқиш;

норавшан сўровлар учун параметрик тегишлилик функцияларини лойиҳалаштириш усули ва дастурий модулини ишлаб чиқиш;

норавшан ва стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш модели, усули, алгоритми ва дастурий модулини ишлаб чиқиш;

норавшан ахборот муҳитларида маълумотларни излаш ва қайта ишлаш учун билимлар базасини яратиш усули, алгоритми ҳамда дастурий модулини яратиш;

корпоратив ахборот-кутубхона тизимларида маълумотларни излаш тизимининг IDEF моделлари, дастурий таъминот архитектур-

тураси, модуллари, алгоритмлари ва маълумотлар базасини ишлаб чиқиш.

Тадқиқотнинг объекти сифатида фан ва таълимга оид корпоратив ахборот-кутубхона тизимларининг норавшан ва стохастик ахборот муҳитларида маълумотлари излаш ва қайта ишлаш жараёнлари қаралган.

Тадқиқотнинг предмети корпоратив ахборот-кутубхона тизимларининг норавшан ва стохастик ахборот муҳитларида маълумотларни интеллектуал излаш ва қайта ишлашда, норавшан тўпламлар назарияси, математик статистика, Data mining, Text mining, SMART, MVC технологиялари, мантиқий излаш моделлари, усуллари ва дастурий воситаларидан иборат.

Тадқиқотнинг усуллари. Тадқиқот жараёнида математик моделлаштириш, норавшан тўпламлар назарияси, математик статистика, алгоритмлаштириш, функционал моделлаштириш, объектга йўналтирилган дастурлаш технологиялари ҳамда MVC, DOM, ORM технологиялари ва ҳисоблаш экспериментларини ўтказиш усуллари, дастурий модулларни синовдан ўтказиш усуллари қўлланилган.

Тадқиқотнинг илмий янгилigi қуйидагилардан иборат:

корпоратив ахборот-кутубхона тизимларида маълумотларни самарали излаш ва қайта ишлашнинг рекуррент муносабатли сақлаш алгоритми такомиллаштирилган ва дастурий модули ишлаб чиқилган;

корпоратив тармоқнинг стохастик ахборот муҳитларида маълумотларни мантиқий семантик излаш модели яратилган;

норавшан сўровлар учун лингвистик ўзгарувчига асосланган параметрик тегишлилик функцияларини лойиҳалаштириш усули такомиллаштирилган ва дастурий модули ишлаб чиқилган;

норавшан ва стохастик ахборот муҳитларида мантиқий муносабатга асосланган маълумотларни излаш тизимининг семантик ядро модели, алгоритми ва дастурий модули ишлаб чиқилган;

норавшан ахборот муҳитларида маълумотларни интеллектуал излаш тизимининг Мамдани қондасига мувофиқ билимлар базасини яратиш модели, алгоритми ва дастурий модули ишлаб чиқилган;

корпоратив тармоқда маълумотларни излаш тизимини яратувчи FSV технологияси, унинг архитектураси, IDEF моделлари ва алгоритмлари, маълумотлар базасининг лойиҳаси ишлаб чиқилган.

Тадқиқотнинг амалий натижалари қуйидагилардан иборат:

-компьютер тармоқларининг корпоратив ахборот-кутубхона тизимларида маълумотларни излаш ва қайта ишлашнинг бошқарув схемаси, ахборот муҳитларининг классификацияси ишлаб чиқилган;

-фойдаланувчиларнинг сўровларини шакллантириш алгоритми, параметрли тегишлилик функциясини лойиҳалаш алгоритми, дастурий модуллари ишлаб чиқилган;

-маълумотларни излаш тизимларини лойиҳалаштиришнинг IDEF моделлари, маълумотлар базасининг лойиҳаси ҳамда корпоратив тармоқларнинг норавшан ва стохастик ахборот муҳитлари учун FSV технологияси асосида «ARMAT++» дастурий таъминоти ишлаб чиқилган.

Тадқиқот натижаларининг ишончлилиги. Тадқиқот иши бўйича олинган солиштирма статистик маълумотлар, таклиф қилинган математик ва IDEF моделлар, усул ва алгоритмлар негизида яратилган FSV технологияси асосида «ARMAT++» дастурий таъминоти маълумотларни интеллектуал излаш ва қайта ишлаш масалаларини ечишда реал ва тажриба синовидан ўтказилиши билан таъминланган. Корпоратив ахборот-кутубхона тизимининг норавшан ахборот муҳитларида маълумотларни интеллектуал излаш тизимларининг мантиқий семантик ва норавшан билимлар базасининг аҳамияти, мослиги ва моҳияти, мезонлар асосида тажриба натижалари таққослаш усулидан фойдаланилган. Ахборот тизимларида маълумотларни интеллектуал излаш ва қайта ишлаш соҳадаги малакали олимлар ҳамда мутахассис экспертларнинг яқуний натижалар бўйича хулосалари билан изоҳланади.

Тадқиқот натижаларининг илмий ва амалий аҳамияти. Тадқиқотда олинган натижаларининг илмий аҳамияти маълумотлар тузилмаси турлича бўлган ёки тизимлаштирилмаган ахборот муҳитлари учун янги модификациялашган маълумотларни излаш ва қайта ишлаш усулларини, алгоритмларини яратиш, маълумотларни параллел излаш ва қайта ишлаш, турли хусусиятли электрон ресурсларни яратиш, илмий электрон ресурсларнинг импакт факторини ҳисоблаш, фан ва таълим, кутубхона, архив, музей ва ихтисослаштирилган корпоратив тармоқларда маълумотларни излаш тизимининг модуллари яратиш ҳамда электрон ҳукумат доирасида маълумотларни излаш масалаларини илмий ечимларини олиш билан изоҳланади.

Олинган натижаларнинг амалий аҳамияти электрон ресурсларда фойдаланувчиларнинг табиий тил асосида ёзган сўровларни қайта ишлаш, архивлар, музейлар, маълумотномалар тизими, хужжат алмашиш тизимлари, фан ва таълимга оид манбаларни сакловчи маълумотлар базасида, Ziyonet каби дастурий таъминотлар учун ички дастурий модул сифатида фойдаланиш катта аҳамиятга эга. Шунингдек, мактаб, академик лицей ва касб-хунар билим юртлири, махсус электрон ресурсга эга муассасалар, давлат ахборот-ресурсларини шакллантириш, улардан фойдаланиш ижросини таъминлаш, жаҳон ахборот-ресурсларини тўплаш, қайта ишлаш ва излаш билан изоҳланади.

Тадқиқот натижаларининг жорий қилиниши. Норавадан ва стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари, FSV технологияси асосида корпоратив ахборот-кутубхона тизимининг «ARMAT++» дастурий таъминоти негизида:

электрон ресурсларнинг дастлабки рейтингини ҳаволалар асосида ҳисоблаш усули, китобхонларнинг норавадан сўровлари учун лингвистик ўзгарувчига асосланган параметрик тегишлилик функцияларини лойиҳалаштириш алгоритми ва дастурий модули, билимлар базасини яратиш усули, алгоритми ва дастурий модули. Фанлар академияси асосий кутубхонасининг корпоратив тармоқларига жорий қилинган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2017 йил 30 октябрдаги 33-8/7293-сонли маълумотномаси). Илмий тадқиқот натижасида электрон ресурслари рейтингини ҳисоблаш самарадорлигини 30%га ошириш ресурсларни такдим қилишда китобхонларга рейтинг юқори ресурслардан фойдаланиш имконини берган. Корпоратив тармоғида сўровларни қайта ишлаш самарадорлигини 15%га ошириш, маълумотларни йиғиш, излаш ва қайта ишлаш, такдим қилиш самарадорлигини 12%га ошириш китобхонлар ва ходимларни тезкор маълумотлар билан таъминлаш сифатини яхшилаш имконини берган;

мантикий семантик излашнинг математик модели, сўровларни қайта ишлаш учун лингвистик ўзгарувчига асосланган параметрик тегишлилик функцияларини лойиҳалаштириш алгоритми ва дастурий модули, семантик ядро яратиш модели, усуллари, алгоритмлари ва дастурий модули республика илмий қишлоқ хўжалиги кутубхонаси ва унинг филиаллариаро корпоратив тармоғига жорий қилинган (Ахборот технологиялари ва коммуникацияларини

ривожлантириш вазирлигининг 2017 йил 30 октябрдаги 33-8/7293-сонли маълумотномаси). Илмий тадқиқот натижасида табиий тилда баён қилинган сўровларга мос ресурсларни танлаш имкониятини 30%га ошириш, маълумотларни излаш самарадорлигини 8%га ошириш, топилган мос электрон ресурслар сонини 2 баравар камайтириш ҳамда аниқлигини 30%га ошириш кўп сонли электрон ресурслардан зарур маълумотларни олиш ва таҳлил қилиш вақтини тежаш имконини берган;

маълумотларни рекуррент муносабатли саклаш алгоритми, семантик ядро модели, алгоритми ва дастурий модули, Мамдани қондасига мувофиқ билимлар базасини яратиш модели, алгоритми ва дастурий модули Интеллектуал мулк агентлиги ахборот-ресурс марказига жорий қилинган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2017 йил 30 октябрдаги 33-8/7293-сонли маълумотномаси). Илмий тадқиқот натижасида бошланғич маълумотларни шакллантириш ва қайта ишлаш самарадорлигини 15%га ошириш, маълумотларни излаш самарадорлигини 12%га ошириш, норавшан термли сўровларни қайта ишлаш орқали сўровга мос маълумотлар сонини 30%га ошириш ҳамда ортиқча маълумотлар сонини 50%га камайтириш имконини берган;

электрон ресурсларнинг дастлабки рейтингини ҳаволалар асосида ҳисоблаш усули ва дастурий модули, мантиқий семантик излашнинг математик усули ва дастурий модули, маълумотларни излаш тизимини яратиш учун ишлаб чиқилган FSV технологиясининг IDEF моделлари, архитектураси ва маълумотлар базасининг лойиҳаси «E-Line Press» МЧЖга жорий қилинган (Ахборот технологиялари ва коммуникацияларини ривожлантириш вазирлигининг 2017 йил 30 октябрдаги 33-8/7293-сонли маълумотномаси, «E-Line Press» МЧЖнинг 2017 йил 28 сентябрдаги Исх.№ 437/2017 маълумотномаси). Илмий тадқиқот натижасида илмий ресурс ва даврий нашрларнинг дастлабки рейтингини аниқлаш самарадорлигини 1,5 мартага ошириш, маълумотлар базасини яратиш самарадорлигини 20%га ошириш, маълумотларни излаш самарадорлигини 5%га ошириш, топилган мос ресурслар сонини 50%га камайтириш ва аниқлигини 1,3 мартага ошириш имконини берган.

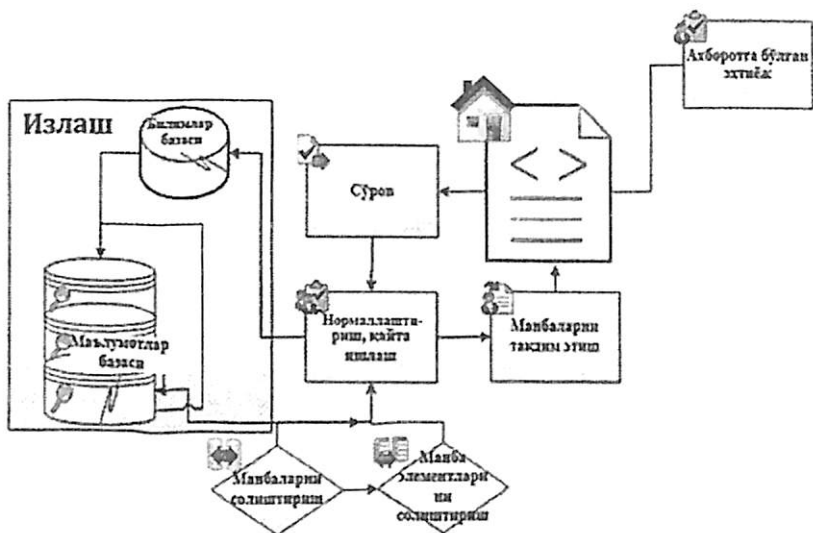
Тадқиқот натижаларининг апробацияси. Мазкур тадқиқотнинг назарий ва амалий натижалари 13 та халқаро ва 15 та республика илмий-амалий анжуманларида маъруза қилинган ва муҳокамадан ўтказилган.

I-боб. КОРПОРАТИВ АХБОРОТ МУҲИТЛАРИДА МАЪЛУМОТЛАРНИ ИЗЛАШ ВА ҚАЙТА ИШЛАШ УСУЛЛАРИ, ВОСИТАЛАРИНИНГ ҲОЗИРГИ ҲОЛАТИ

1.1. Маълумот излашнинг ривожланиш босқичлари ва фундаментал моделлари, усуллари

XX асрнинг ўрталарига келиб, информатика соҳаси ва кутубхона ишининг туташган қирраларида пайдо бўлган янги илмий ва амалий йўналиш бу – «Маълумотларни излаш» йўналиши [123; 6-б.]. «Маълумотларни излаш (инглизча – Information retrieval, русча – Информационный поиск)» сўзи Келвин Муэрс (Calvin Northrup Mooers) томонидан 1948 йилда ёзилган докторлик диссертацияси орқали фанга кириб келган ва 1960 йиллардан бошлаб мустақил илмий-техникавий йўналишга айланган. Маълумотларни излаш ва қайта ишлаш (МИҚИ) муаммосининг биринчи қатъий тенденцияси XX асрнинг 50-60-йилларига тўғри келса, глобал компьютер тармоғи Интернет тизимининг пайдо бўлиши ва жаҳон глобал ахборот тизимига айланиши унинг иккинчи қатъий тенденция билан ривожланишига сабаб бўлган. «Маълумотларни излаш» сўзининг маъноси кенг камровлидир. Чунки ахборотлашган жамиятнинг ҳар бир соҳасида МИҚИ масаласига дуч келамиз ва ёндашувлар ҳам турлича [110]. «Маълумотларни излаш» сўзининг маъноси илмий тадқиқот ва амалий йўналишига нисбатан турлича таърифланган [91; 12-б.]. Фикримизча, «*Маълумотларни излаш*» – ахборотга бўлган эҳтиёжни қондириш мақсадида тузилган сўров ва сўровни шакллантириш усули, маълумотлар базасидан (МБ) сўровга мос маълумотларни танлаш, қайта ишлаш ва тақдим қилиш жараёнидир.

Корпоратив ахборот муҳитларида (КАМ) МИҚИ асосий модуллардан бўлиб ҳисобланади [83; 42-43-б.]. Чунки, ахборотга бўлган эҳтиёжни қондириш учун унга мос электрон ресурсни (ЭР), маълумотни излаш ва тақдим қилиш лозим. Ахборот муҳитларининг имкониятлари, қўлланиш соҳалари, архитектураси, ишлаш тамойилларига тизимли ёндашувлар асосида олиб борилган таҳлил ва тадқиқот натижалари асосида «Маълумотларни излаш» сўзининг қиёсий таърифларга таяниб МИҚИнинг умумий концептуал схемасини куйидагича тасвирлаймиз (1.1-расм) [47, 48, 49, 96, 98, 103]:



1.1-расм. Маълумотларни излаш ва қайта ишлашнинг концептуал схемаси.

Корпоратив ахборот муҳитларида МИҚИнинг концептуал схемаси тўртта асосий дастурий модулдан иборат бўлиши лозим [115]:

- ахборотга эҳтиёжнинг мавжудлиги ва шу асосида сўровни шакллантириш дастурий модули;
- сўровни нормаллаштириш ва ББ асосида МБдан танлаш усулларининг дастурий модули [119];
- ЭРлар тўплами, яъни сўровга мос танланган маълумотлар тўпламини солиштириш дастурий модули;
- ЭРлар тўпламини қайта ишлаш (баҳолаш, индекслаш [104], тайёрлаш) ва тақдим этиш дастурий модули.

Корпоратив ахборот муҳитларида МИҚИда асосий элементлар «сўров» ва «сўров объекти» бўлиб ҳисобланади.

Сўров – ахборотга бўлган эҳтиёжни сунъий ёки табиий тилда формал кўриниши. Сўровларни шакллантириш бўйича турли ёндашувлар мавжуд бўлиб, ҳар бир ахборот муҳитининг ўзига хос талаблари мавжуд [53]. Махсус сўровларни шакллантириш тиллари SQL, Language Integrated Query, XQuery, XPath, шунингдек, сўровларни табиий тилларда шакллантириш имкониятлари жуда кўп фойдаланиб келинмоқда [63, 78].

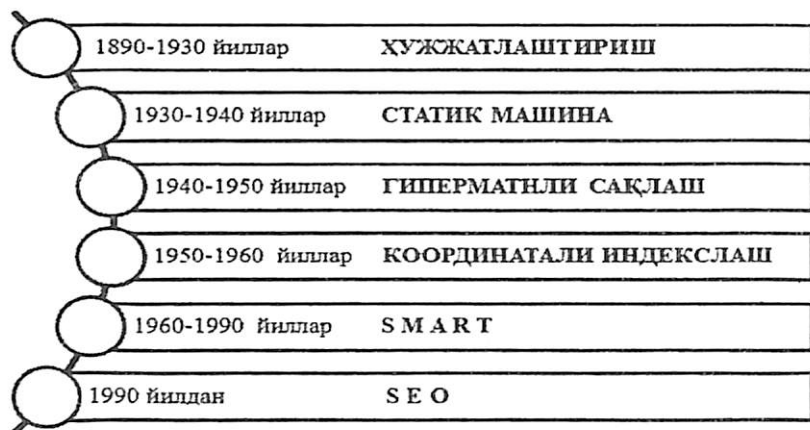
Сўров объекти – сўровга мос такдим учун тайёрланган ЭРлар, маълумотлар тўплами. Сўров объектини чегаралаб бўлмайди ва унинг турли кўриниши бўлиши мумкин.

1960 йилларнинг ўрталари ва 1970 йилларнинг бошларида ЭҲМлар асосида ахборотни йиғиш, таҳлиллаш, синфлаштириш, сақлаш, масофага узатиш, излаш ва узатишнинг автоматлашган тизимлари бўлган. Хусусан, Гарвард университети профессори Дж.Солтон бошчилигидаги тадқиқотчилар гуруҳи томонидан замонавий излаш тизимлари базавий тамойиллари учун жуда кўп қўлланиб келинаётган SMART усулига асосланган матн излаш ва таҳлил қилиш тизими яратилган ҳамда кенг фойдаланилган [91; 14-15-б.]. Унинг хусусий жиҳати ЭРлар ва сўровларни табақалаш масалаларида янгича ёндашувлар асосида ахборот мазмунини таҳлиллаш, интерактив излаш ва узатиш тамойилини илгари сурганида.

Аммо, 1970 йилларда ЭҲМларнинг қуввати ва хотирасининг кичиклиги, юқори сифатли алоқа каналларининг етишмаслиги туфайли МИҚИ соҳаси яхши ривожланмаган. Бундан ташқари, муаммоли масалалар ҳам бўлган: биринчидан, ахборот тизимларида узоқ масофадан туриб фойдаланиш учун умумий тармок протоколи йўқ эди. Иккинчидан, маълумотни узоқ юклаш масалаларини реал вақтда бу тизимлар амалга ошириб беролмас эди. Булар эса фақат локал маълумотларни излаш тизимларини (МИТ) ривожлантиришга сабаб бўлган.

1980 йилларга келиб, марказий процессор билан боғланмасдан туриб, маълумотларни қайта ишлаш ва юқори сифатли визуал тасвирлаш имкониятини берувчи шахсий компьютерлар кўпая бошлаган. Бу эса марказлаштирилган МИТларни тадқиқ қилиш ва яратиш соҳасини сезиларли даражада пасайишга олиб келган. Натижада, МИҚИ соҳасида фундаментал илмий тадқиқот ишлари тўхтаб қолган. Фақатгина, 1990 йилларга келиб, Интернет тизимининг катта ҳажмдаги ахборот сақлаш имконияти пайдо бўлиши билан глобал тармокларда 3W технологиясига асосланган контентли ахборот муҳитларида МИТларга яна эҳтиёж сезиш бошланган [91; 16-б.].

Тадқиқот доирасида олиб борилган таҳлилларга асосланиб, ахборот муҳитларида МИҚИнинг ривожланиш босқичлари ва технологиялари қуйидаги кетма-кетлик асосида ишлаб чиқилган (1.2-расм) [26, 92].

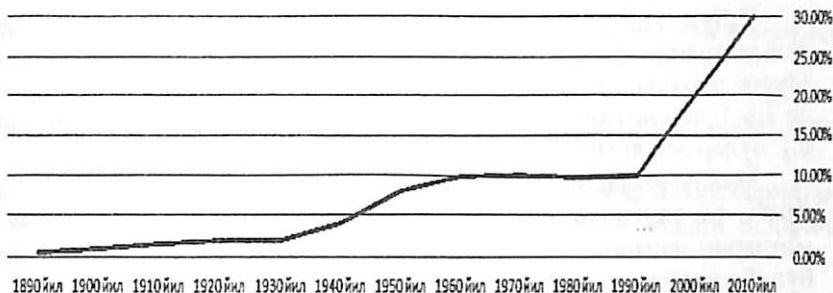


1.2-расм. Маълумот излашнинг ривожланиш босқичлари ва технологиялари

Интернет инфраструктурасининг ривожланиши МИҚИ усулларига ҳам ўз таъсирини кўрсата бошлаган. Чунки, Интернетда очик, ёпик ЭР тизимлари кенг тарқала бошлаган ва маълумот излаш, қайта ишлаш учун махсус МИТлар пайдо бўлган (Alta Vista, Excite, Lycos, Yahoo, Yandex, Google) ва тармоқларини ҳам кенгайтириб бормокда [34; 291-299-б., 26]. Бу тизимлар бугунги кунда SEO тамойилига асосланган ҳолда веб контентларда МИҚИ, тақдим этишни нормаллаштириш ва оптималлаштириш усулларидан фойдаланилади [16, 31].

Интернет тизими DOM технологияси асосида гиперҳаволали матнлар ва 3W технологиясига асосланади [23;10-б.]. МИТларнинг фойдаланувчилари, веб контентлари, интерфейс тиллари, сервер компьютерлари сонининг ортиб бориши тизимлар орасида рақобат ва уларга эҳтиёж борлигини билдиради. Бу эса маълумотларни излаш тизимларида МИҚИ моделларини, усуллари ва дастурий модулини доимий такомиллаштириш, ривожлантирилишини тақозо этади.

Жамиятда инфраструктура ва ахборот тизимларининг ривожланиши асосида МИҚИ технологияларининг татбиқ қилинганлигининг микдорий графигини келтирамиз (1.3-расм).



1.3-расм. Маълумот излашнинг технологияларини тадбиқ қилинганлигининг миқдорий графиги

Корпоратив тармоқларда МИҚИ технологияларининг ривожланиши фуқароларнинг турли ахборотга эҳтиёжлари ортиб боришини билдиради. Телекоммуникация ва ахборот тизимларининг мунтазам ривожланиши МИТ ва электрон ресурсларга янгидан-янги талаб ҳамда муаммоларни келтириб чиқаради.

Айни пайтда энг ривожланган ва тараққий этган давлатларда жуда катта соҳали ва очик ахборот муҳитлари мавжуд: Европада CORDIS нинг [3] бир қисми бўлган ERGO [62] интеграциялашган тизим, Америкада кенг ҳажмли ахборот билан Конгресс кутубхонаси [33], Россияда «Информатика» тизими, Ягона илмий ахборот муҳити (ЕНИП РАН) [56], eLibrary [64], EBSCOPost DataBase [9], EBSCO Information services [8] ва Socionet [52] тизимлари. Бундай тизимларнинг аксарияти фойдаланувчиларни ахборотга бўлган эҳтиёжини қондириб келиш билан бирга уларнинг ҳар бирини маълум йўналишдаги муаммолари мавжуд. Бу муаммоларга:

- замонавий ахборот тизимлари учун дастурий таъминотлар мослашиш имкониятининг йўқлиги (интеграция масаласи);

- энг сўнги долзарб маълумотлар, ЭРлар билан бойитиб боришнинг ташкилий, иқтисодий, технологик ва инфраструктуравий босқичидаги муаммолар;

- МИТнинг мураккаб тузилганлиги. Баъзиларида аънавий излаш усулларидан фойдаланилса ҳам (кутубхоналар учун муаллиф, номи ва бошқалар) муаммолар борлиги;

- ЭРларнинг тўлиқ матнларидан излаш муаммолари;

- каталоглаштириш ва тўлиқ матнларни электрон кўринишга ўтказишнинг ташкилий, иқтисодий, технологик ва инфраструктуравий муаммолари.

Х.Лун (H.P.Luhn) моддий эхтиёжини қондириш учун мантикий излаш тизимларига ўзи билмаган ҳолда асос солган [60; 802-б.]. Муерс фикрича, маълумотларни излаш тизимларини яратиш асосида Дж.Бул яратган алгебранинг амалларига таяниш лозим [60; 804-б.]. Булар эса янги муаммоларга сабаб бўлган.

Маълумотларни излашда сўровлар учун иккита мантикий излаш ва тартибли индекслаш моделлари мавжуд бўлиб, маълумотларни излашнинг ҳозиргача энг кўп фойдаланилган моделлари бўлиб ҳисобланади [19, 21].

Маълумотларни излаш ва қайта ишлаш соҳасида энг кўп фойдаланиладиган ва таянч излаш моделлари мантикий сўровга асосланган мантикий излаш моделлари бўлиб ҳисобланади [91; 22-б.].

Маълумотларни амалий излаш соҳасида «ВА», «ЁКИ» амалларининг ўзаро мослашувчанлигини аниқлаш ва аргенал маълумотни тақдим этиш усулларини Ли ва Фоксаларнинг (Lee and Fox) тадқиқот ишларида [25], маълумотларни излаш тезлиги ва мантикий излаш моделлари ҳамда тартибли излаш моделларини таққослаш бўйича Уитттен бошчилигидаги (Witten and all) жамоанинг тадқиқот ишларида [68], Регуляр муносабатларнинг амалий асосларини яратиш ва амалиётга қўллаш муаммоларини Фридл (Friedl), назарий асосларини Хопкрофтнинг (Hopcroft) илмий-амалий тадқиқот ишларида ёритилган [18].

Агар мантикий излаш сўровида сўров терминлар кетма-кетлиги кўринишида мантикий амалларсиз берилган бўлса, бу терминларнинг ҳар бири учун тегишли ЭРлар, маълумотлар рўйхатини тўғридан-тўғри маълумотлар тўпламидан излаш орқали тақдим этади. Яъни, терминлар мантикий «ЁКИ» амал мазмуни билан берилган деб ҳисобланади. Аммо сўровда мантикий «ВА», «ИНКОР» мазмунидаги амаллар ҳам берилган бўлса, сўровга мос маълумотларни излаб топиш учун «Термин-маълумот» матрицасидан фойдаланиш тавсия этилади. Бу матрица мантикий излаш моделининг асоси бўлиб ҳисобланади [91; 21-б.].

Мантикий излаш модели – мантикий амаллар билан ифодалаш мумкин бўлган барча сўровлар учун маълумотларни излаш модели. Мантикий излаш моделининг ютуқларини санаб ўтамиз:

- кўп сонли маълумотлар тўпамини қайта ишлаш. Маълумотлар тўпамининг сони компьютернинг тезлигига нисбатан паст даражада ўсиб боришга нисбатан;

- такқослашда мослашувчанлик. Терминлар орасида мантикий амалларни қўллаш орқали амалга оширилишига нисбатан;

- тартибли излаш. Берилган сўровга МБдан энг мос маълумотларни, яъни айнан сўровдаги терминларга эга бўлган маълумотларни тақдим этиши;

- сўров тилларига мослиги. Мантикий амаллар билан берилган сўровни SQL тилига шакллантириш тез ва осон амалга ошириш мумкинлиги.

Мантикий излаш модели камчиликларини санаб ўтаемиз:

- мураккаб сўровларни мантикий сўровга акслантириш;

- «Термин-маълумот» матричасининг ҳажми, яъни маълумотлар турли терминларга эга бўлганда «Термин-маълумот» матричаси ҳажми катталашшига олиб келади. Бу эса оддий сўров учун кўп амални бажаришни талаб этади;

- фақатгина мантикий сўровлар билан ишлаши;

- компьютерда бажариладиган мантикий амалларнинг такрорий бажарилиши.

Ўз вақтида мантикий излаш моделининг асосий камчилиги «Термин-маълумот» матричасининг ҳажми билан боғлиқ бўлган [65].

Бу муаммонинг олдини олиш учун ҳар бир термин учун махсус тўплам ташкил қилиш ғояси илгари сурилган. Бу ғоя маълумотларни излаш учун асосий ва биринчи концепция вазифасини ўтаган «Тартибли индекс» моделининг пайдо бўлишига олиб келган.

IBM ходимлари (Tauble and Wooster) томонидан автоматик тартибли индекслаш учун биринчи дастурий таъминот яратилган [54].

«Тартибли индекс» моделида маълумотлар тўпламига нисбатан луғат (терминлар рўйхати) тузилади ва ҳар бир термин учун маълумотлар рўйхати тузилади. Рўйхатнинг ҳар бир элементи терминнинг координатаси ҳақидаги маълумотни сақлайди. «Тартибли индекс» модели босқичлари куйидагича:

1. Маълумотлар тўпламини йиғиш ва шакллантириш.

2. Терминлар рўйхатини тузиш (луғат тузиш).

3. Маълумотлар координатасини аниқлаш.

4. Терминга мос координаталар тўпламини тузиш.

Маълумотларни шажарали излаш, луғатларда В-шажарали излаш каби назарий маълумотларни Д.Кнутнинг (D.Knuth) моногра-

фиясида [80], тартибли индекслаш бўйича Гарфилднинг (Garfield) ишларида ўз аксини топган [21,104].

«Тартибли индекс» модели (ингл. inverted index) маълумотларнинг махсус тузилмаси бўлиб, унда маълумотлар тўплами ҳар бир термини учун маълумотларнинг мос рўйхатлари сакланади.

«Тартибли индекс» моделининг ютуқлари қуйидагилардан иборат:

- аниклик;
- натижанинг тиниклиги;
- муайян тартиб асосида эффектив саралаш имконияти;
- юкори аниклик – «ВА» учун пастдан ва «ЁКИ» учун юкоридан аникланган;

«Тартибли индекс» моделининг мантикий сўровларни қайта ишлашда инобатга олинмаган жиҳатлари қуйидагилардан иборат хисоблаймиз[97]:

- фақатгина луғатдаги сўзлардан фойдаланилади, сўровдаги имловий хатолар ва норавшан терминлар учун шакллантирилмаган;
- сўровдаги бир-бирига ёзилиши яқин бўлган терминлар;
- сўровдаги синоним терминлар;
- сўровдаги аноним терминлар;
- маълумотда терминнинг частотаси;
- сўровга боғланган ҳолда тартиблаш имконияти;

Тартибли индекс модели асосида тартибли излаш модели яратилган ва унинг босқичларига асосланган ҳамда фарқли равишда қуйидаги босқичлардан иборат:

- маълумотлар тўпламини йиғиш ва мавзули индекслаш [20];
- манти термин (сўз)ларга ажратиш;
- терминларни бошланғич лингвистик таҳлил қилиш;
- маълумотни ҳар бир термин бўйича индекслаш;

Маълумотларни индекслаш бўйича Уиттен (Witten et all) раҳбарлигида олимлар томонидан индекслар тузиш, хотира ва дискда индекслаш ҳамда вақтинчалик тартибланган рўйхатлар учун алгоритмлари бўйича жуда чуқур илмий тадқиқот ишлари олиб борилган [12]. Бундай алгоритмларга BSBi каби саралашга асосланган параллел жамлаш (sort-based multi-way merge) алгоритмини олишимиз мумкин. Бу алгоритмнинг фарқли томони луғатлар тузилиши ва уларни сиқиш тамойилларидадир. Хайнц (Heinz et al.) раҳбарлигида олимлар хотирада луғатларни тўплаш учун маълумотлар тузилишини ўрганиб туриб, таккослашган [66]. Бютчер ва

Кларкелар (Buttcher and Clarke) индекслаштиришда кўп фойдаланувчи тизимда ахборот хавфсизлигини ўрганишган [61].

Луғат асосида излаш моделлари маълумотларни излаш ва қайта ишлаш усулларининг асосини тартибли индекслаш таъкил қилиб, мантикий ва сўзли сўровлар учун мўлжалланган. Джонсон ва унинг ҳамкасблари томонидан илгари сурилган ғоя асосида сўровларни ёзишдаги имловий хатолар ёки сўровдаги терминларга альтернатив вариантдаги терминларни аниқлаш масалаларига эътибор кучайтирилган. Бундай масалаларни ҳал қилиш «Луғатлар ва толерант⁴ излаш» сарлавҳаси орқали маълумотларни излаш ва қайта ишлаш жамоатчилигига эълон қилинган [11]. Луғатлар ва толерант излаш ўз ичида бир нечта модель ва усулларни камраб олади.

Вариантли сўровлардан куйидаги ҳолатларда фойдаланилади [102]:

- фойдаланувчи терминнинг тўғри ёзилишини билмас (масалан, олма ёки алма учун *лма);

- фойдаланувчи терминни бир нечта вариантлар билан ёзилиши ва аниқ биттасини билмаслиги (масалан, район, туман, ноҳия);

- фойдаланувчи терминни тўлиқ билмас (масалан, автомат*);

- фойдаланувчининг терминларни аниқ билмаслиги (масалан, ахборот * тузилмаси).

Умумий кўринишдаги вариантли сўровларда берилган q_v вариантли сўров мантикий q_i сўровига махсус алгоритм орқали ўтказилади. Терминлар тўпламидан q_i сўровга мос терминлар тўплами ҳосил қилинади ва q_v ҳақиқий сўров билан солиштирилади. Агар мос келса, қайта ишланади ва тартибли индекслаш амалга оширилади.

Умумий кўринишдаги вариантли сўровларда индекслашда ўрин алмаштириш биринчи махсус амал бўлиб ҳисобланади. Чунки у ўзида тартибли индекслашнинг турли кўринишларини сақлайди. Одатда, терминлар охирида *Null* ёзилади ва бу кўринмайди.

Индекслашда ўрин алмаштириш бўйича биринчи марта Гарфилднинг ишларида тўлиқ маълумотлар келтирилган [27].

⁴ Инглиз тилида «Tolerant» ёки «Fuzzy», рус тилида «Сходство», «Толерантный» кейинчалик «Нечеткий» деб таржималар келтирилган. Ўзбек тилида ҳам «Норавшан» термини билан таржима қилинган. Аммо биз «Толерант» деб таржима қилдик. Чунки, Норавшан терм асосида норавшан тўпламлар назарияси назарда тутилган. Бу терм ўхшашлик, вариантлик терминларига мос келади.

Индекслашда ўрин алмаштириш жараёнида хотиранинг кўп қисмини банд қилмаслик масалалари Феррагин ва Вентуринларнинг тадқиқот ишларида ўз аксини топган [13].

Индекслашда ўрин алмаштириш усули ёрдамида яратилган терминлар ўрин алмаштирилган терминлар луғати деб юритилди.

Сўровлар учун k -граммли индекслашнинг асосий ғояси берилган сўровни k тадан белгили терминларга ажратиб чиқишдир.

Масофали таҳрирлашда k -граммли индексдан фойдаланиш бўйича тўлиқ маълумот ва усулларнинг батафсил изоҳлари ҳамда тажриба натижалари Цобел ва Дартнинг ишларида келтирилган [65]. Уларнинг фикрича, масофали таҳрирлаш учун энг самарали усул бу k -граммли индекс ҳисобланади [6].

Агар бу индекслашдан фойдаланилса, k -граммли индекслаш учун универсал луғатни ҳосил қилиш лозим. Луғатдаги ҳар бир термин ўзининг k -граммли терминиға мос тартибланган рўйхати тузилади.

Маълумотларни излашда сўров хатоларини тузатиш муҳим аҳамиятға эға. Жуда кўп ҳолатларда фойдаланувчилар томонидан сўровларни авто хатолик (билиб-билмай) билан тузадилар. Бундай сўровларни носоз, хатоли сўровлар деб аташ мумкин. Шундай хатоли сўровларни қайта ишлаш ва хатоларини тузатиш муаммоси учун 2 та усул мавжуд: масофали таҳрирлаш (edit distance), k -граммли бирлашма (k -gram overlap). Бу усулларни қўллашдан олдин сўров хатоларини тузатишни ташкиллаштириш лозим.

Орфографик хатоларни тузатиш усуллари бўйича илк тадқиқотлар Дамеро ишларида татқиқ қилинган [38]. «Масофали таҳрирлаш» тушунчасини Левенштейн таклиф қилган [30] бўлсада, унинг алгоритмини Вагнер ва Фишер яратганлар [22]. Масофали таҳрирлаш алгоритмларига асосланган усуллар Перетсон ва Кукичлар ишларида келтирилган [17, 22].

Кўплаб сўровдаги хатоларни тузатиш алгоритмлари иккита фундаментал тамойилға асосланади.

1. Хатоли сўровларни тузатиш учун кўп ҳолларда сўровға яқин (яқин ўхшашликка эға бўлишлик) тизимда фойдаланилган сўровлар олинади. Бунинг учун икки сўровнинг ўзаро ўхшашлик тушунчасини тушуниш керак.

2. Агар иккита бир-бириға яқин боғлиқ сўровлар бўлса, уларнинг умумийға яқин бўлган шакли танланади. Масалан, [бедо] сўровига мос яқин сўровлар [бедод] ва [бедор] бўлиши мумкин.

Хатоларни тузатиш усулининг иккита модели мавжуд:

- хавфсиз ҳолда муддатли тузатиш (isolated-term correction);

- контекст нуктаи назар асосида тузатиш (context-sensitive correction);

Хавфсиз ҳолда муддатли тузатиш моделининг усуллари масофали таҳрирлаш ва k -грамм бўлиб ҳисобланади. Масофали таҳрирлашда иккита S_1 ва S_2 терминларининг энг кам таҳрирлаш амаллари (edit operations) тушунилади. Агар таҳрирлашда амалларнинг вазнлари бирлаштирилса, масофали таҳрирлашни умумлаштириш мумкин. Хатоларни тузатиш учун k -граммли индекслашда сўровлар ва луғатда мавжуд терминлар орасидаги таҳрирлаш масофасини топиш учун k -граммли индексдан ҳам фойдаланиш мумкин. Бу усул сўровга мос луғатдаги терминлар орасидаги таҳрирлаш масофасининг энг яқинини топишга имкон беради.

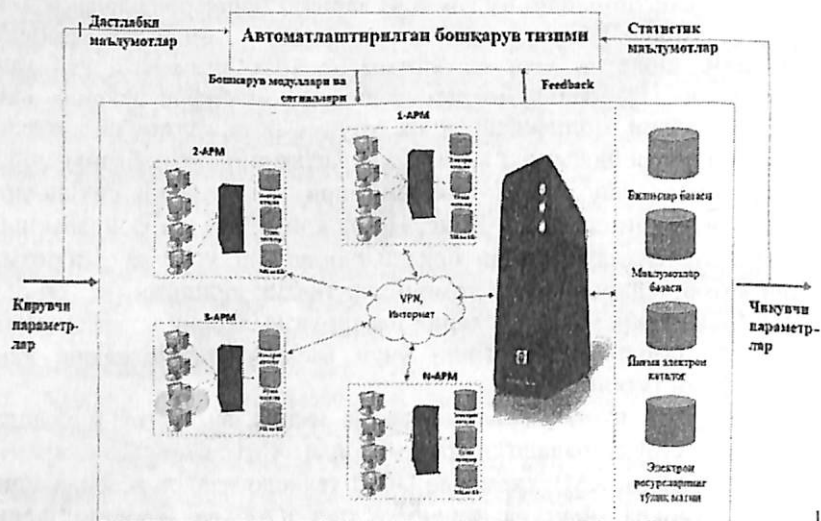
Вариантли излаш усулларида фонетика асосида хатоларни тузатиш (phonetic correction) ҳам муҳим усуллардан биридир. Бунинг учун ҳар бир термин фонетик хеш ёрдамида генерацияланади. Яъни, бир хил эшитиладиган терминлар бир хил сонлар билан кодлаштирилади. Бу ғоя ХХI асрнинг бошларида хавфсизлик хизмати билан шуғулланувчи ходимлар томонидан таклиф қилинган. Фонетик хеш алгоритми Soundex алгоритми деб ҳам юритилади [1]. Маълумотлар излашни оптималлаштириш ёки бирор моделни модификациялаш учун кўп тарқалган индекслаш усулларида фойдаланиш керак. Индекслашнинг энг кўп тарқалган ва ўрганиладиган муаммоларидан бири индексларни сиқишдир. Маълумотларни сиқишда Хипс, Ципф конунларидан фойдаланиш, луғатларни сиқиш, кодлаш орқали сиқишнинг усул ва алгоритмлари қатор тадқиқотчилар томонидан тадқиқ қилинган [91; 60-65-б.]. КАКТларида маълумотларни излаш тизимларини профессионал даражада фаолият кўрсатиши учун маълумотларни сиқиш ҳам муҳим ўрин тутди.

Бугунги кунда маълумотларни излаш ва қайта ишлашда уларнинг структуралашган тизимларида XML схемалари муҳим ўрин эгаллайди. XML схемалар DOM технологиялари асосида кенг ривожланмоқда. Фан ва таълимга оид КАКТда маълумотларни излаш ва қайта ишлаш моделлари, усул ва алгоритмларини тадқиқ қилиш ҳамда ишлаб чиқиш учун корпоратив тармокнинг ахбороткутубхона тизимлари ва унда сўровларни қайта ишлаш усулларини тадқиқ қилиш лозим.

1.2. Корпоратив ахборот-кутубхона тизимлари ва сўровларни қайта ишлаш усуллари

Корпоратив ахборот-кутубхона тизимларининг (КАКТ) ахборот муҳитларида мазкур тадқиқот иши натижасида МИҚИ учун олинган математик, ахборот моделлари ва усуллар, алгоритмлар, дастурий таъминотни тажриба синовидан ўтказиш учун энг мақбул муҳитдир. Бунинг учун КАКТлари учун компьютер тармоқларининг архитектурасини аниқлаш, яратилган дастурий таъминотларнинг имконияти, фарқлари, мослашувчанлиги, айникса, МИҚИ модулларининг тамойилларини тадқиқ қилиш лозим. КАКТни қуйидагича тасвирланади (1.4-расм).

КАКТ учун дастурий маҳсулотлар бозори 1990 йиллардан бошлаб фаол ривожлана бошлаган. Республикамизда 1990 йилларнинг ўрталарига қадар Россия Федерациясида ишлаб чиқарилган КАКТлар фақат локал тармоқларда ишлашга мўлжалланган, бу тизимлар файл-серверли архитектураси асосида фаолият олиб борган [113; 59-б.].



1.4-расм. Корпоратив ахборот-кутубхона тизими

Дунёдаги катта кутубхоналар, чунончи Оксфорд, Россия Давлат кутубхонаси, ГПИБ, ВГБИЛ, Томск университети кутубхонаси ва қатор хорижий мамлакатлар кутубхоналари VTLS, LIBER, TINLIB, ALEPH дастурий таъминотлардан фойдаланишади. Аммо кутубхоналарга мўлжалланган бу дастурий таъминотлар

Мустакил Давлатлар Ҳамдўстлиги мамлакатларида кенг тарқалмади, бунинг сабаблари, биринчидан, уларнинг нархи анча қиммат бўлса, иккинчидан, уларнинг Россия кутубхона технологияларига (стандартлари ва форматларига) мосланмаганлигида эди. Бир жиҳатдан олиб қараганда, Россияда ишлаб чиқилган КАКТларнинг қатор афзалликлари мавжуд. Уларда мамлакатимиз учун хос бўлган кутубхона технологияси асос қилиб олинган, хорижий мамлакатлар КАКТларига қараганда нархи ҳам анча арзон эди. Аммо Россияда яратилган КАКТлар бир қатор параметрлар бўйича (ишончлилиқ, технологик, Интернет билан боғлана олишлиқ, фойдаланувчи интерфейсидаги қулайликлар) хорижий мамлакатлар КАКТларидан ортада қолмоқда [113;70-б.].

1990 йилларнинг ўрталарида Россияда клиент-сервер архитектураси ва МББТнинг юқори синфига тегишли SQL2 стандарти асосида ORACLE, SQLServerларга асосланган замонавий тизимларнинг янги авлодлари яратилган. Улар MARC форматлар ва истикболли XML, SGML стандартлар билан ишлай олади. Бу соҳада эришилган муваффақиятлардан бири сифатида Россиянинг «ДИТ-М» компанияси томонидан ишлаб чиқилган OPAC-Global дастурий таъминотидир. У Россиянинг LibNet дастури талаблари асосида Россия Миллий кутубхонаси мутахассислари билан яқин ҳамкорликда ишлаб чиқарилган [114].

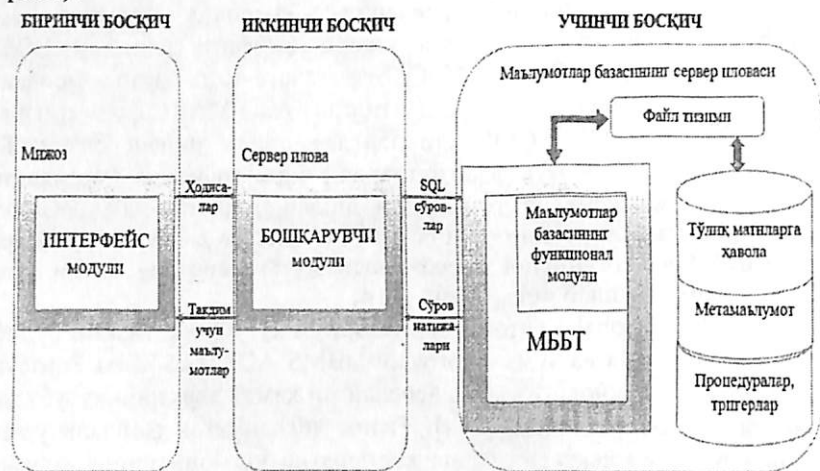
OPAC-Global – автоматлаштирилган кутубхона тизими бўлиб, Web-технология ва system software DBMS ADABAS, Text Retrieval System каби технологияларга асосланган ҳамда электрон кутубхона яратишга мўлжалланган [114]. Тизим чекланмаган фойдаланувчилар учун реал вақт режимида корпоратив иш юритишни таъминлайди. OPAC-Global ўз ичига қуйидаги тизим остиларга эга: каталоглаштириш, излаш, буюртма ва бошқалар. Россиянинг қатор кутубхоналари OPAC-Globalдан фойдаланган. РНБ, ГЦНМБ, Саха-Якутия, Свердловск ва Челябинск илмий кутубхоналарида бу маҳсулот татбиқ қилинган. Россия Миллий кутубхонасининг электрон каталоги OPAC-Global муҳитида шаклланган.

Дастурий маҳсулотлар бозорида кутубхона технологияларини автоматлаштиришга мўлжалланган кўплаб КАКТлар таклиф қилинмоқда, улар орасида Ўзбекистоннинг ARMAT++⁵ дастурий таъминоти ҳам бор [28, 29, 82, 85].

⁵ тадқиқот муаллифи томонидан MVC ва ORM технологиялари асосида яратилган.

Ахборот-ресурс марказларида (АРМ) кенг татбиқ қилинаётган, кутубхоналарни автоматлаштиришга мўлжалланган тизимлардан бир нечасининг техник ва дастурий характеристикаларидан кўринадики, уларнинг деярли барчаси кутубхоналарни комплекс автоматлаштиришга мўлжалланган бўлиб, функционал имкониятлари бир хил. Дастурий таъминотларнинг архитектураси ҳам деярли бир хил: 2 босқичли «Клиент-сервер» тамойили [113].

Тадқиқот натижасида фан ва таълимга оид норавшан ва стохастик ахборот муҳитларида (НСАМ) МИТни яратиш учун компьютер тармоқларининг 3 босқичли «Клиент-сервер» тамойилига асосланган архитектурасини куйидагича тасвирлаймиз (1.5-расм):



1.5-расм. КАКТлари учун компьютер тармоғининг 3 босқичли «Клиент-сервер» архитектураси.

Таҳлил қилинган КАКТларнинг барчаси библиографик ёзувларни алмаштиришнинг халқаро талабларига жавоб беради, штрих кодлардан фойдаланиш имконияти мавжуд. Тизим томонидан тайёрланадиган чиқиш формаларини ҳам ихтиёрий шаклда қоғозда босиб чиқиш имконияти мавжуд. Шунча ўхшаш имкониятлари бири-бирига яқин бўлса ҳам уларда сўровларни нормаллаштириш, маълумотларни излаш усуллари, моделлари ва излашни ташкиллаштириб берувчи мулоқот ойналарининг турли хилдалиги уларни рақобатбардош дастурий маҳсулотларни яратишга олиб келмоқда. Аммо ҳозирги кунгача универсал КАКТнинг дастурий таъминоти яратилмаган. Бунинг турли сабаблари бор, ўйлаймизки, сабаб-

ларидан бири фойдаланувчиларнинг ахборот эҳтиёжини қондира олмаслик, МИҚИ натижасининг ҳаддан кўп ортикча турли-туман маълумотлар билан тақдим қилиш имкониятларининг жуда юкори эканлигида.

Республикамизда ИБС, АРМАТ (КАРМАТ), АРМАТ++ каби КАКТлар яратганки, бу дастурий таъминотлар 60га яқин олий таълим муассасалари ва 200дан ортиқ касб-хунар ва академик лицейларда, қатор ташкилотларда ишлатилиб келинмоқда. Бу тизимлар фойдаланувчилар томонидан тузиладиган сўровларни қайта ишламайди ёки шакллантирмайди, кўпинча бир хил жавоб қайтаради. Аммо, амалиётда жуда кўп сўровлар ноаниқ ёки норавшан (сонли характери бўлмаган) кўринишда тузилади. Бунини амалга ошириш учун АРМАТ++ тизимида норавшан ва стохастик ахборот муҳитларида маълумотларни интеллектуал излаш ва сўровларни эркин тузиш ҳамда уни нормаллаштириш имкониятлари бўлиши лозим, деб ҳисоблаймиз.

КАКТларда сўровларни қайта ишлаш – МБсидан маълумотларни олиш учун фойдаланиладиган энг асосий вазифалар сирасига кирадиган амаллардан бўлиб ҳисобланади [4, 7]. Бу жараён юкори босқичли дастурлаш муҳитларида шакллантириладиган сўровларни МБ учун лойиҳалаштириш [115], мослаштириш усуллари, акслантириш, оптималлаштириш ва сўровларни баҳолаш тизимларини қамраб олади.

Сўровларни қайта ишлашнинг асосий босқичлари қуйидагилардан иборат [60; 537-б.]:

1. Таҳлил ва таржима.
2. Нормаллаштириш.
3. Баҳолаш.

Сўровларни қайта ишлашда сўровни ўзидан фойдаланиш мумкин бўлган кўринишга келтириш керак, яъни SQL стандартидан унумли фойдаланиш лозим. Аммо SQL стандарти баъзи муаммоларни келтириб чиқаради. Тизимнинг ички сўров тузилишига келтириш асосий муаммо ҳисобланади.

Шундай қилиб, биринчи вазифа тизим сўровни қайта ишлаб, уни тизимнинг ички формал кўринишига келтириш керак. Бу жараён бошқа тизимлардаги парсер компилятор каби амалга оширилади. Сўровнинг ички формал кўринишини шакллантириш учун фойдаланувчи сўровини таҳлилловчи модул ёрдамида синтаксис таҳлил қилиш керак. Сўровларнинг кўпчилиги МБларнинг муносабатлари билан боғлиқ. Модул сўрови таҳлилидан сўнг уни

шажара кўринишда тасвирлайди. Сўровнинг тузилиши тўлиқ муносабат амаллари билан бир хил маънони саклаб қолиши зарур. Кўп компляторлар таҳлиллашнинг турли усулларида фойдаланади.

Жавобларни олиш учун сўровга боғлиқ ҳолда кўп қоидалар, турли усуллар мавжуд [7, 90]. Сўровларнинг ҳар бирида SQL сўрови ўзга усул ёрдамида муносабат кўринишида ўтказишнинг бир нечта усуллари мавжуд. Бундан ташқари, муносабат сўровларнинг амалий баҳолаши имконияти билан ҳам боғланади.

КАКТда сўровларни қайта ишлаш учун ёзилган қадамларнинг ўзаро боғлиқлиги асосий омил бўлиб ҳисобланади [107]. Ҳамма МБ ҳам бундай аниқликка асосланган қадамларга таяна олмайди. Реляцион алгебранинг қоида ва ифодаларидан фойдаланиш учун баъзи МБларида SQL сўровнинг тузилишида статик шажаралар мажмуасидан фойдаланилади. Аммо бу қоида МБларида сўровларни қайта ишлашнинг асосини ташкил қилади.

1.1-жадвал

Танлаш амали асосида маълумотларни излаш усуллари.

	Усул номи	Баҳоси	Изоҳ
A1	Чизиқли излаш	$T_s + b_r \cdot T_r$	b_r файлдаги блоклар сони.
A2	Чизиқли излаш, калитгга тенглаштириш	Ўртача қиймат $T_s + \left(\frac{b_r}{2}\right) \cdot T_r$	Берилган шарт каноатлангирганича изланади, калитгга мос топилгандан кейин тўхтатилади.
A3	Бирламчи В+-шажарали индекс, калитгга тенглаштириш	$(h_i + 1) \cdot (T_r + T_s)$	h_i -индекснинг юқори қиймати. Излаш индексини киритиш чиқаришни танлаш учун шажаранинг қийматини бирга оширади.
A4	Бирламчи В+-шажарали индекс, калитгга тенглаштирамаслик	$h_i \cdot (T_r + T_s) + b \cdot T_i$	Ҳар бир шажара босқичи учун излаш, биринчи блок учун излаш.
A5	Иккиламчи В+-шажарали индекс, калитгга тенглаштириш	$(h_i + 1) \cdot (T_r + T_s)$	Бирламчи калитгга мос иккиламчи калитгни тенглаштириш.

1.1-жадвал давоми

A6	Бирламчи В+- шажарали индекс, калитга тенглаштирмаслик	$(h_i + n) \cdot (T_r + T_s)$	Бу ерда n мос келмаганлар сони. Бунда А3 ўхшаш индекслардан фойдаланилади. Ўзувларнинг бошқа блокка жойлашганлиги инобатга олинади.
A7	Бирламчи В+- шажарали индекс, такқослаш	$h_i \cdot (T_r + T_s) + b \cdot T_i$	А3 алгоритмининг калитга тенг бўлмаган ҳоли.
A8	Иккиламчи В+- шажарали индекс, такқослаш	$(h_i + n) \cdot (T_r + T_s)$	А4 алгоритмининг калитга тенг бўлмаган ҳоли.

Сўровларни нормаллаштириш учун уларни оптималлаштирувчи ҳар бир амал учун ажратилган вақтни олдиндан билиш шарт. Аммо бу вақтни баҳолаш муҳим, машаққатли вазифа ҳисобланади. Чунки, бу иш турли параметрлар, маълумотларни қайта ишлашга мўлжалланган хотира воситалари билан боғлиқ ҳамда ҳар бир амал бажарилишини баҳолаш имкониятини бериш зарур.

Сўровни бажаришни баҳолаш дискдан маълумотларни узатишда блоклар сони ва дисклар сони билан ўзаро боғлиқ ҳисобланади. Агар дискдаги блокка маълумот узатиш тизим остисининг ўртача вақти T_i ва блокка рухсат вақти T_c бўлса, унда b блокка маълумот узатиш ва s блокка рухсат вақтининг умумий кўриниши $b \cdot T_i + s \cdot T_c$ ҳисобланади. Техник воситалар учун T_c ва T_i ларни бугунги энг сўнгги технологиялар асосида $T_c = 4ms$ ва $T_i = 0.1$ деб олиб, $4KB$ ли блок ва маълумот узатишнинг тезлиги секундига $10MB$ га тенг бўлади.

Танлаш амали сўровларни қайта ишлашда маълумотларга муурожаатнинг муҳим босқичи бўлиб ҳисобланади. Танлаш орқали излаш усуллари реляцион муносабатли МБсида жойлашган барча муносабатларни кўздан кечириб чиқиши лозим. Бунда муносабатли кортежларни сақлаган ҳолда танлаш амалидан фойдаланиш лозим.

Юқорида келтирилган КАКТлар таҳлили ва уларда сўровларни қайта ишлаш усуллари маълумотларни излашда мантикий излаш

моделларига асосланган. Шунингдек, турли ахборот муҳитларида МИҚИда синфлаш масаласидан кенг фойдаланилади. Фан ва таълимга оид корпоратив ахборот-кутубхона тизимларида МИҚИда синфлаш масаласининг, ўзига хос хусусият ва усулларини келтирамиз.

1.3. Фан ва таълимга оид ахборот муҳитларини классификациялаш

Фан ва таълимга оид ахборот муҳитларида МИҚИ учун ахборот муҳитининг ўзига хос хусусиятини, маълумотлар тузилмасини тадқиқ қилиш лозим. Бунинг учун «Ахборот муҳити» деганда, нимани тушунамиз?», деган саволга жавоб излаш керак.

Ахборот муҳити – маълум мезонга асосланган ахборот билан таъминлаш учун етарли шароитлар, воситалар (техник, иктисодий, ташкилий ва бошқалар)ни қамраб оладиган муҳит. Бошқача айтганда, ахборотни сақлаш, шакллантириш, излаш ва узатишни амалга ошириш мумкин бўлган яхлит тизим [87].

Фан ва таълимга оид ахборот муҳитининг элементлари сифатида қуйидагиларни танлаб оламиз:

- классификациялаш тизими ва стандартлар (УДК, ББК, классификаторлар, MARC формати);
- ахборот-ресурсларини шакллантириш тизими;
- маълумотларни излаш тизими;
- маълумот ва билимлар базаси (электрон каталог, тўлиқ матнларнинг базаси, йиғма электрон каталог, статистик маълумотлар базаси);
- маълумотларни тақдим қилиш тизими (интерфейс, web-саҳифа);
- маълумотлар базасининг администратор тизими (АИЎ «Администратор»);
- ахборот-ресурслари яратувчилари (муаллифлар, кутубхоначилар, операторлар, администратор ва бошқалар);
- ахборот-ресурс фойдаланувчилари (китобхонлар);
- телекоммуникацион инфраструктура (локал тармоқ, Интернет).

Фан ва таълимга оид ахборот муҳитининг объект ҳамда субъекти сифатида мутахассис ва фойдаланувчилар, назорат қилиш ва бошқариш усуллари, ахборот оқими, тартиб, регламент, ҳуқуқий нормалар, аппарат ва дастурий таъминот, телекоммуникация воситалари, билиш босқичлари ва назоратнинг қуйи тизимини ўз ичига оладиган инфраструктураларни танлаш мумкин. Ахборот

муҳити яхлит тизим сифатида бутун ҳаёт алмашиш давомида ахборот алмашиш ва ахборот тизимларига кўрсатадиган барча омилларнинг элементлар йиғиндиси сифатида қараш ҳам мумкин.

Корпоратив тармоқнинг фан ва таълимга оид ахборот муҳитларида шарт ва зарур бўлган имкониятларни санаб ўтамиз:

1. Маълумотни излаш қондаси, танлаш ва изоҳлаш.

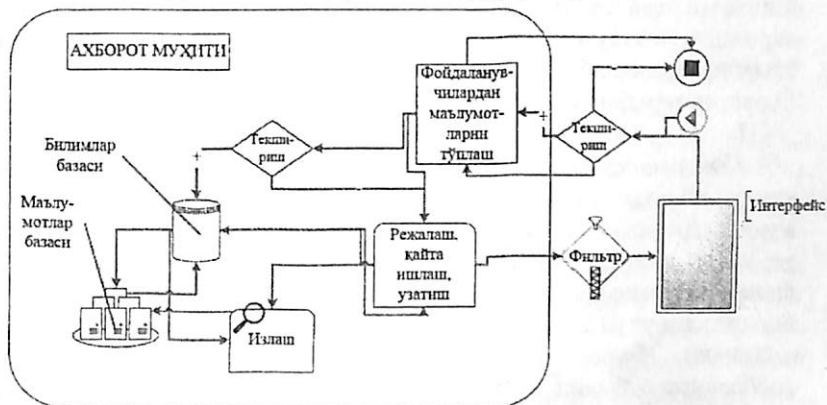
2. Маълумотни яратиш, кўриш ва аслик, олинган ахборотни тарқатиш.

3. АРларни шакллантириш, таълимий ахборотлар ва ахборотли интерактив хизматларнинг таъминланганлиги.

4. Ахборот тизимларини (ААТ, МБ, ББ) яратиш ва бошқа инфраструктураларда қўллаш, интеграция қилиш.

Инфраструктура хавфсизлиги учун ахборот хавфсизлиги, шунингдек, компьютер ва компьютер тармоқларида ахборот-дастурий ва аппарат таъминотни ҳимоя қилиш учун усул, восита ва механизмларни қўллаш ва ривожлантириш учун хизмат қилади.

Бу имкониятлар корпоратив тармоқ ахборот муҳитларининг яхлит тизимли узлуксиз ривожланишини таъминлайди. Шундай экан, бу вазифани бажариш учун ахборот муҳити аниқ бошқариш схемаси асосида ишлаб чиқарилиши зарур (1.6-расм).



1.6-расм. Фан ва таълимга оид ахборот муҳитининг бошқарув схемаси

Ушбу схемада фойдаланувчилар томонидан фан ва таълимга оид ахборот муҳитида сўровлар ёрдамида керакли маълумотни излаш, танлаш ва тақдим қилиш ўз аксини топган.

Схеманинг модулли босқичлари куйидагилардан иборат:

- мухитга кириш модули. Бу модул аутентификацияни амалга оширади;

- фойдаланувчилардан керакли маълумотларни тўплаш модули. Бу модулда маълумотларга нисбатан фойдаланувчининг ахборот мухитидаги ўрни аникланади;

- режалаш, қайта ишлаш, узатиш модули. Бу модул маълумотлар йўналишини бошқариш, назорат қилиш ва статистикани олиб боришдан иборат;

- излаш модули. Сўровга мос маълумотларни излаш жараёнини назорат қилади;

- маълумотлар базаси модули. Бу барча маълумотларни сақлаш ва сўровларни қайта ишлашни таъминловчи модулдир [15];

- билимлар базаси модули. Бунда маълумотларни сўровларга мос равишда чиқариб берувчи қоидалар мажмуасини ташкил қилиш ва амалга ошириш имкониятини таъминловчи модул;

- фильтр модули. Маълумотлар тўпламини фойдаланувчи талабига асосан саралаб, тақдим қилишни амалга ошириш модули.

Ушбу модуллар корпоратив тармокнинг фан ва таълимга оид ахборот мухитларида маълумотларни сақлаш, излаш ва қайта ишлашни амалга ошириш, маълумотларнинг қайта ишланганлик даражаси, маълумот хусусиятлари асосида ахборот мухитларини классификациялаб, 3 турга ажратиш мумкинлигини кўрсатади. Булар, детерминан, стохастик ва норавшан ахборот мухитларидир [111].

Детерминанли ахборот мухити (ДАМ) – МИҚИ учун аниқ қоида асосида катъий тартибланган маълумотларга эга бўлган мухит. ДАМда маълумотларни сақлаш, излаш ва қайта ишлаш усуллари ҳам олдиндан аникланиб, зарур ресурс учун мутахассис ёрдамида тизимлаштирилади. ДАМдан бирор маълумотни олиш ёки сақлаш учун мавжуд тизимлаштирилган комбинациялар таклиф қилинади. Бирор маълумотни излаш учун ҳам олдиндан тайёрланган сўровлар тўпланидан танлаб олиш мумкин. Танланган t – сўров учун L_i –мос келадиган ресурслар тўплами мавжуд бўлади ($i = 1 \dots n, n \rightarrow \infty$).

Стохастик ахборот мухити (САМ) – МИҚИ учун тасодифий характерли ва унинг қоидалари асосида тақсимланган маълумотларга бўлган мухит. САМдан бирор маълумот олиш учун олдин фойдаланилган тизимлаштирилмаган комбинациялар ҳамда янги комбинацияларни тузиш таклиф қилинади. Мухит таклиф қилаётган ЭРлар эҳтимолли, тасодифий қоидалар хусусиятига эга. Бирор

маълумотни излаш учун ҳам олдин фойдаланилган сўровлар тўплами ва янги сўров тузиш имконияти таклиф этилади. Танланган, олдин фойдаланилган t – сўров учун L_i – мос келадиган ресурслар тўплами мавжуд бўлади ($i = 1 \dots n, n \rightarrow \infty$). Агар L_i тўпландан фойдаланилса, унинг эҳтимоллик даражаси ошади. Агар, t сўров N та фойдаланувчи томонидан M марта танланган ресурс бўлса, t сўровга L_i тўпландан ресурсларининг эҳтимоллиги $\frac{M}{N}$ га тенг. Бундан N нинг ошиб бориши [инкремент (+1)]га M нинг ошиши [инкремент (+1)]/ошмаслиги [(+0)] натижаси учун ўринли бўлади.

$$\lim_{\substack{N \rightarrow N++ \\ M \rightarrow M++}} \frac{M}{N} \rightarrow 1 \quad \text{ва} \quad \lim_{\substack{N \rightarrow N++ \\ M \rightarrow M}} \frac{M}{N} \rightarrow 0 \quad (1.1)$$

Агар йиллар ва фойдаланувчиларнинг фаоллиги ҳисобга олинса, САМнинг маълум ресурслар тўпламига ДАМ сифатида қараш мумкин. Яъни, агар Y САМнинг фаолият йиллари ва N фойдаланувчилар сони бўлса, Y нинг ошиши N нинг ошишига олиб келади ва (1.1)дан (1.2) ифода ўринли бўлади.

$$ДАМ \subseteq САМ \text{ ёки } САМ \rightarrow ДАМ \neq 100\% \quad (1.2)$$

Норавшан ахборот муҳити (НАМ) – МИҚИ учун норавшан характерли қатъий тартибланмаган маълумотларга эга бўлган муҳит [114, 115, 116]. НАМдан бирор маълумот олиш учун сизга олдин фойдаланилган тизимлаштирилмаган мутахассисларнинг комбинациялар ва мутахассис сифатида янги комбинациялар тузиш таклиф қилинади. Агар $\{ \{ \text{“йўналиш”} \Rightarrow \text{фан} \} \rightarrow \{ \text{“блок”} \Rightarrow \text{математика} \} \}$ бўлса, $\{ \text{“бўлим”} \Rightarrow [\text{олий математика, математик анализ, элементар математика, ..., планиметрия}] \}$ каби олдин фойдаланилган ресурслар таклиф қилиниши мумкин. Аммо мутахассислар томонидан математика фанининг бўлимида “бўлим” учун келтирилганларнинг μ тегишлилик функцияси ёрдамида аниқланган қиймати бўлади. A билан $\{ \text{“блок”} \Rightarrow \text{математика} \}$ ни, B билан $\{ \text{“бўлим”} \Rightarrow [\text{олий математика, математик анализ, элементар математика}] \}$ белгиласак, унда норавшан тўпландан C мавжуд бўлади, қуйидагича аниқланади [112]:

$$C = \{ (x, \mu_C(x)); x \in B \} \text{ бу ерда } \mu_C : B \rightarrow [0,1]$$

Муҳит таклиф қилаётган ЭРлар норавшан муносабатлилиқ хусусиятига эга [113]. Худди шундай бирор маълумотни излаш учун ҳам олдин фойдаланилган сўровлар тўплами ва янги сўров тузиш имконияти таклиф этилади. Танланган, олдин фойдала-

нилган t сўров учун мутахассислар томонидан таклиф қилинган L_i мос келадиган ресурсларнинг норавшан тўплами мавжуд бўлади ($i=1\dots n, n \rightarrow \infty$). Агар шу L_i норавшан тўпландан фойдаланиш ва реал муносабат қолдирилса, L_i норавшан тўпланинг t сўров учун тегишлилик даражаси ошади. Агар t сўров учун N та мутахассис томонидан L_i тўпланинг ҳар бир элементи учун M марта таъкидланган ресурс бўлса, t сўровга L_i тўплам K_i норавшан тўплами қуйидагича:

$$K_i = \left\{ (M_{j,i}, \mu_{K_i}(l_{j,i})); l_{j,i} \in L_i \right\} \quad (1.3)$$

Бунда: $\mu_{K_i}(l_{j,i}) = \frac{M_{j,i}}{N}$, $M_{j,i}$ - N та мутахассис томонидан таъкидланган $l_{j,i}$ дир.

(1.3) ифодада N нинг ошиб бориши $[(+1)]$ га $M_{j,i}$ нинг ошиши $[\frac{1}{N}]$ ва $\mu_{K_i}(l_{j,i}) \rightarrow [0,1]$ натижаси учун ўринли бўлади:

$$\sum_j \lim_{\substack{N \rightarrow N++ \\ M_{j,i} \rightarrow \frac{1}{N}}} (M_{j,i}, \mu_{K_i}(l_{j,i})) \rightarrow [0,1] \quad (1.4)$$

Бунда: $\sum_j K_i$ ларнинг норавшан тўплами.

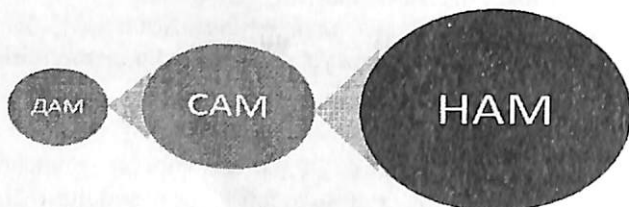
Йиллар ва мутахассисларнинг фаоллигини ҳисобга олиб, НАМнинг маълум ресурслар тўпламига САМ сифатида қараш мумкин. Яъни, Y агар НАМнинг фаолият йиллари ва N мутахассислар сони бўлса, Y нинг ошиши N нинг ошишига олиб келади, (1.4) дан (1.5) ифода ўринли бўлади.

$$САМ \subseteq НАМ \text{ ёки } НАМ \rightarrow САМ \neq 100\% \quad (1.5)$$

Фан ва таълимга оид ахборот муҳитлари синфларини биргалликда тўплалар шаклида изоҳласак, қуйидагича бўлади:

$$ДАМ \subseteq САМ \subseteq НАМ \text{ ёки } НАМ \rightarrow САМ \rightarrow ДАМ \neq 100\% \quad (1.6)$$

(1.6) ифодани визуал тасвирласак, қуйидагича бўлади (1.7-расм):



1.7-расм. Ахборот муҳитларининг ўзаро таъсири

(1.6) ифода ва 1.7-расмдан фан ва таълимга оид ахборот муҳитлари доимо ўзаро уйғунлашган ҳолда учрайди. Чунки, мавжуд ахборотлар ДАМга ўтса, янги ахборотлар кўпайиб, САМ ва НАМни ташкил қилаверади.

Фан ва таълимга оид ахборот муҳитлари жамиятнинг ажралмас бўлагидир. Жамият ривожланиши билан ахборот муҳитларининг тузилиш ва имкониятлари мосланувчан бўлиши лозим. Юқорида келтирилган 5 та имконият ва бошқарув схемаси асосида ишлаб чиқиладиган ахборот муҳитларида ДАМ, САМ, НАМларни шакллантиришни инобатга олиш муҳим касб этади. Фан ва таълимга оид ахборот муҳитларининг тутган ўрни фойдаланувчиларнинг сони билан ўлчанади. Шундай экан, фойдаланувчилар сонини кўпайтириш ва эҳтиёж бўлган ахборотни олиш учун маълумотларни излаш тизимларининг технологияларини ишлаб чиқиш, қолаверса, маълумотларни гурухлаштириш, фойдаланувчилардан эксперт сифатида фойдаланиш лозим.

1.4. Корпоратив ахборот муҳитларида маълумотларни излаш тизимини яратиш муаммолари

Фан ва таълимга оид КАМларида маълумотларни излаш ва қайта ишлаш соҳаси шу давргача ўз бошидан 2 тенденцияни ўтказиб, 3 тенденцияни ўтказиш арафасида турибди. Бунинг сабаблари жуда кўп, масалан, ахборотга бўлган эҳтиёж, техника ва технологияларнинг янги тенденцияларга эришишганлиги ва ҳ.к. МИҚИ учун SMART усули, синфлаш масаласи, Содда Байес усули, мультиномиал модель варианти ва баъзи синфловчилар учун хусусиятларни танлаш усуллари, Бернулли, Роккио, К энг яқин кўшнилар, kNN , чизикли ва чизиксиз синфловчилар, кластеризация гипотезаси асосида кластеризациялаш усулларининг К-ўртачалар, EM-алгоритм, Иерархик, Бир ва тўлиқ алоқа, Центроидлар, Пастга юрувчи кластеризация, Латент-семантик индекслаш усуллари техника ва ахборот тизимларида (АТ) ишлатилмоқда. Шунинг учун МИҚИ илмий-техникавий илм-фаннинг долзарб муаммоларидан. XX асрнинг охириги 10 йилликларида маълумотларни қайта ишлаш ва узатиш соҳасида яратилган ҳамда ривожланган юқори сифатли технологиялар, замонавий телекоммуникацион тизимлар, айниқса, Интернет тизими кутубхоналар учун янгидан-янги имкониятлар пайдо бўлишига олиб келди. Бу омиллар ахборотга бўлган эҳтиёжни тубдан ўзгартирди ва сифатли хизматларни талаб қила бошлади.

А.Г.Спиркиннинг ёзишича, МИҚИ тизимларининг асосий вазифаларидан бири эҳтиёж сезилаётган маълумотни билим

даражасида таъминлаб бериш ва инсон тафаккурида “ғоя, тушунча, фикр, назария” шаклида етарли акс этишидир [52].

Ҳақиқатан ҳам, ЭРлар кенг қамровли ресурсларининг пайдо бўлиши ва улардаги маълумотлар ҳажмининг кескин ошиб бориши дунё олимлари олдига янгидан-янги илмий-техникавий муаммоларни қўймоқда. Бу муаммолардан бири корпоратив ахборот-кутубхона тизимида МИҚИ моделлари, усулларини такомиллаштириш, модернизациялаш ва дастурий таъминотларини яратишдир.

«Маълумотларни излаш муаммоси – фуқаролик жамиятининг доимий муаммоларидан биридир», деб ёзади Ю.И.Шокин [123]. Чунки, янги билимларни шакллантириш, инновацион ғоялар ишлаб чиқиш, илмий-техникавий тадқиқотларни бажариш орқали жамиятни раванак топиши учун турли ахборот тизимлари ва муҳитларидан МИҚИ муҳим аҳамият касб этади. Яъни, янги даврнинг таракқий этиши ундан олдинги илмий-оммавий маълумотларни излаб топиш ва ўрганишга боғлиқ.

Юқоридагилардан корпоратив ахборот-кутубхона тизимида МИҚИ масалалари, маълумотларни излаш тизимларини яратиш ва фойдаланиш муаммоси ахборотлашган жамиятнинг аниқ муҳим муаммоларидан бўлиб ҳисобланади. Шунинг учун семантик ахборот хусусиятлари ва тузилиши ҳақидаги илмий қарашлардан иборат назарий информатика соҳасида фаол ривожланиш бошланган [109]. Бу илмий соҳанинг асосий масалаларидан бири маълумот излаш муаммосидир [32]. Чунки фойдаланувчиларнинг ахборотга бўлган эҳтиёжларини кондириш энг муҳим ва аҳамиятли масала ҳисобланади.

Ҳукуматимиз томонидан дастурий таъминотларни, ЭРлар базасини яратиш бўйича амалга оширилаётган чора-тадбирлар, тадқиқ қилинган ЭРлар ва амалий ишлар, юқорида келтирилган фикр-мулоҳазаларга асосланган ҳолда фан ва таълимга оид КАКТида маълумотларни излаш ва қайта ишлаш масалаларида куйидаги муаммолар мавжуд:

1. КАКТда маълумотларни излаш ва қайта ишлаш тамойилини классификациялашни тадқиқ қилиш.

2. Катта ҳажмли КАКТда ахборотни сақлаш, киритиш, таҳрирлаш усулларини доимий такомиллаштиришни тадқиқ қилиш.

3. КАКТларда фойдаланувчиларнинг ахборотга бўлган эҳтиёжи – сўровларни шакллантириш, гуруҳлаш, таҳлил қилиш, нормаллаштириш масалаларини тадқиқ қилиш.

4. АТларда маълумотларни излаш ва қайта ишлаш технология ҳамда боскичларини доимий таҳлил қилиш ва модификациялаш масалаларини тадқиқ қилиш.

5. Дастурий таъминотларнинг қўлланиш соҳаларига мос МИҚИ элементларини ривожлантиришни тадқиқ қилиш.

6. КАКТларда ЭРнинг сўров-тўплам жуфтлигининг долзарблигини тадқиқ қилиш.

7. КАКТларда маълумотлар излашни тезлаштириш, параллеллаштириш масалаларини тадқиқ қилиш.

8. МИҚИ масалалари билан узвий боғлиқ бўлган ахборот хавфсизлиги, маълумотларни индекслаш, сиқиш масалаларини, техник воситаларнинг имкониятларини тадқиқ қилиш.

9. Сўровларни мазмунли (динамик) қайта ишлаш, маълумотларни излаш учун норавшан тўпламлар назарияси асосида математик аппаратни ва у асосида янги усул ҳамда моделларни ишлаб чиқиш ва амалиётда жорий қилиш масаласини тадқиқ қилиш.

10. Фойдаланувчиларнинг ахборот эҳтиёжини кондириш учун сўровларни ўзгармас ёки статик эмас, балки динамик қайта ишлаш усуллари яратиш масаласини тадқиқ қилиш.

11. НСАМларида маълумотларни излаш учун статистик ва динамик маълумотларни йиғиш, қайта ишлаш масаласини тадқиқ қилиш.

12. Статистик ва динамик маълумотлар учун тегишлилик функцияларини лойиҳалаштириш, синфни танлаш масаласини тадқиқ қилиш.

13. КАКТларда маълумотларни интеллектуал излашда семантик ядрони яратиш масаласини тадқиқ қилиш.

14. КАМлар учун норавшан моделларга асосланган усул ва алгоритмлар ишлаб чиқиш масаласини тадқиқ қилиш, ускунавий дастурий таъминот яратиш.

15. Миллий тилларда баён қилинган сўровларни мазмун ва моҳиятан қайта ишлаш масаласини тадқиқ қилиш.

16. МИҚИ ускунавий воситасини интеграция қилувчи платформани яратиш масаласини тадқиқ қилиш.

МИҚИ ускунавий дастурий таъминоти алоҳида амалиётга татбиқ қилинмайди. Чунки бу UI ёки иловада ости тизим ускунавий таъминот (дастурий модул) сифатида фойдаланилади. Шунинг учун бу каби модулларни амалиётга жорий қилиш давомида ахборот муҳитининг ўзига хос хусусиятларидан келиб чиққан ҳолда ёрдамчи ускуналар ҳам яратилади.

Шунингдек, МИТнинг дастурий модулини яратиш усул ва алгоритмларини тадқиқ қилиш муаммолари мавжуд, жумладан:

биринчидан, жуда кўп усул ва алгоритмлар турли тижорат компаниялари томонида ишлатилади, яратилган ёрдамчи ускунавий таъминотлар асосида иктисодий самарадорликка эга бўлади. Дунёда 3W технологияси асосидаги веб иловаларда МИҚИ бўйича машхур фирмалар жуда кўп, шунингдек, пулли ихтисослаштирилган МИТлари (DuckDuckGo, Нигма, not Evil, YaCy, Pip!, FindSounds) ҳам мавжуд;

иккинчидан, МИҚИ тизими UI ёки иловада ости тизим ускунавий таъминот сифатида фаолият олиб боради. Шунинг учун усул ва алгоритмлари фойдаланувчининг фақат тахминига асосланиб шакллантирилади;

учинчидан, МИҚИ масалалари доимо тадқиқ қилишни талаб қилувчи долзарб масалалардан ҳисобланади. Чунки, жамиятда ахборот инфраструктурасининг ўзгариши, инсонлар ахборотга бўлган эҳтиёжларининг омиллари, техник ва телекоммуникация воситалари ривожланишини тўхтатишнинг иложи йўқ.

Умуман олганда, КАКТда маълумотларни излаш тизимларини яратиш муаммоларини тадқиқ қилиш фойдаланувчимутахассисларда янги билимларни шакллантириш, инновацион ғоялар ишлаб чиқиш, илмий-техникавий тадқиқотларни бажариш орқали жамиятни раванқ топиши, қолаверса, шахсий тафаккурини ривожлантириш учун хизмат қилади.

КАКТларда маълумотларни излаш ва қайта ишлаш масаласида учта асосий элемент мавжуд, буларга q -сўров, d -ЭР, $f(q, d)$ -ЭРнинг сўровга нисбатан долзарблигини ҳисобловчи функцияси. $f(q, d)$ функцияси ҳар доим ҳам q , d га боғлиқ бўлмаслиги мумкин.

КАКТда маълумотларни излаш ва қайта ишлаш элементларига асосланган ҳолда турлича тақикотни амалга ошириш мумкин [77, 78, 79]. Фан ва таълимга оид КАКТ изчил, хусусиятлари, имкониятлари, усул, восита ва механизмларни қўллаш ҳамда ривожлантириш, маълумотларни сақлаш, излаш ва қайта ишлашни амалга ошириш асосида қуйидаги хулосага келамиз:

1. Корпоратив тармоқда сўровларни шакллантириш ва нормаллаштириш, маълумотларни излаш элементлари ва долзарблиликни ҳисоблаш.

2. КАКТнинг стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ва дастурий модулини яратиш.

3. КАКТнинг норавшан ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ва дастурий модулини яратиш.

4. КАКТларида маълумотларни излаш ва қайта ишлашнинг семантик ядро яратиш усуллари, билимлар базасини ишлаб чиқиш.

5. Корпоратив тармоқларнинг КАКТларида маълумотларни излаш ва қайта ишлашнинг инструментал дастурий таъминотини яратиш, жорий қилиш.

Бу масалалар КАКТнинг норавшан ва стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ҳамда дастурий модули учун ягона архитектурали платформа яратишни талаб қилади. Бу платформа корпоратив тармоқларнинг ахборот тизимларида МИҚИ масалалари амалий ечимларини олишга хизмат қилиши лозим.

Биринчи боб бўйича хулосалар

Бобда МИТларида маълумотларни излаш ва қайта ишлашнинг таянч моделлари, усуллари, КАКТларининг имкониятлари, хусусиятлари, фан ва таълимга оид ахборот муҳитининг ўзига хос хусусиятлари ва мавжуд муаммолари тадқиқ қилинган бўлиб, куйидаги хулосаларга олиб келади:

1. “Маълумотларни излаш” тушунчасининг қиёсий таърифи ва умумий концептуал схема ишлаб чиқилди. Бу схема турли характерли ахборот муҳитларида маълумотларни излаш ва қайта ишлашнинг модель ҳамда дастурий модулларини яратишга хизмат қилади.

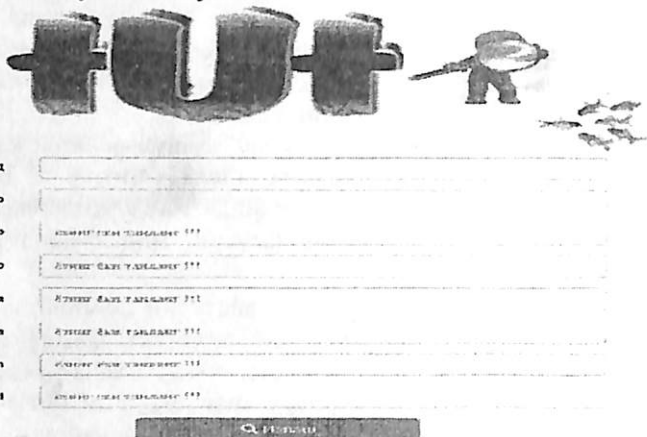
2. КАКТларида мавжуд дастурий таъминотларнинг имкониятлари, ишлаш тамойиллари, архитектуралари асосида фан ва таълимга оид корпоратив ахборот-кутубхона тизимида МИҚИ учун 3 босқичли «Мижоз-сервер» архитектураси танлаб олинди. Бу эса КАКТларида маълумотларни излаш ва қайта ишлаш самарадорлигини оширишга хизмат қилади.

3. Фан ва таълимга оид ахборот муҳитларининг шарт ва зарур бўлган имкониятлари, модулларининг вазифаси, саклаш, излаш ва қайта ишлаш тамойиллари, хусусиятларига асосланиб классификацияси, ривожланиш босқичлари ишлаб чиқилди. Ахборот муҳитининг классификацияси, ривожланиш босқичлари катта ҳажмли, турли характерли ахборот-ресурсларидан фойдаланиш моделлари, усул ҳамда технологияларини ривожлантириш истиқболларини аниқлаб бериш учун хизмат қилади.

II-боб. СТОХАСТИК АХБОРОТ МУҲИТЛАРИДА МАЪЛУМОТЛАРНИ ИЗЛАШ ВА ҚАЙТА ИШЛАШ МОДЕЛЛАРИ, УСУЛЛАРИ ВА ДАСТУРИЙ МОДУЛИНИ ЯРАТИШ

2.1. Стохастик ахборот муҳитларида маълумотларни излаш тизимининг элементлари ва моделлаштириш омиллари

КАКТда маълумотларни излаш ва қайта ишлашнинг мантикий излаш усулларида сўровлар терминлар кетма-кетлиги сифатида ўрганилди. Шунингдек, САМда берилган терминлар кетма-кетлигини ЭРлар метамаълумоти билан боғлаймиз. ЭРнинг метамаълумоти деганда, унинг аниқ характеристикалари тушунилади (муаллиф, номи, яратилган сана) [58]. Бу метамаълумотлар метамаълумот майдонларини билдиради, улар маълумот атрибутлари деб ҳам юритилади [58]. Бу атрибутларнинг типлари ва қийматлари олдиндан чегаралаб қўйилган бўлади.



2.1-расм. Параметрли излаш интерфейси

Фойдаланувчи томонидан тузиладиган сўровлар айнан шу атрибутларга мослаштириш орқали амалга оширилиши сўров ва сўров орасидаги эҳтимолликни таъминлайди. Шунингдек, сўров тартибланган рўйхат ва параметрик индекс орқали қайта ишланади. Унда ҳар бир атрибут учун битта параметрик индекс мавжуд. Параметрик индекс айнан мавжудларини аниқлаш имкониятини беради. Агар, сўровда вақт берилган бўлса, фақатгина параметрик индексда вақт атрибути мавжуд бўлган маълумотларни олади. Бундай сўровларни шакллантириш, яъни параметрик излашни амалга ошириш учун мулоқот ойнасидан фойдаланилади (2.1-расм).

Атрибут майдони ихтиёрий матнларни саклаш учун мўлжалланган бўлиб, одатда, ЭРнинг сарлавҳаси ёки аннотациясини саклайди.

Сўровни қайта ишлашни қўллаб-қувватлаш учун сарлавҳа, муаллиф, аннотация бўйича тартибланган рўйхатни куйидагича шакллантириш мумкин (2.1-жадвал).

2.1-жадвал

Атрибут майдонига мос индекслаш

Аннотация.термин → 1 → 2 → 3 → 6 → 45 → 89 → 600
Сарлавҳа.термин → 5 → 6 → 8 → 45 → 89 → 500
Муаллиф.термин → 6 → 9 → 10 → 15 → 56 → 500

Кодлаш орқали индексларни куйидагича содалаштириш мумкин:

Термин → 2. *сарлавҳа*, 2. *аннотация* → 3. *Муаллиф* → 5. *аннотация*

САМлариди маълумотларни излаш учун параметрли ва зонали майдонларнинг долзарблигини ҳисоблаш усуллари муҳим аҳамият касб этади.

Вазли майдонларнинг долзарблигини ҳисоблаш усули учун параметрли ва майдонли излашда майдон, атрибут орқали излашнинг вариантыни кўриб чиқамиз.

Бунда сўзли сўровни q билан, ЭРни эса d билан белгилаймиз. Долзарбликнинг вазли майдон усулида (q, d) жуфтлик $[0, 1]$ ораликқа акслантирилади. ЭРлар зонаси мантикий қийматларни қабул қилади.

l зонали ЭРлар тўпламини қараймиз. Унда $g_1, g_1, \dots, g_l \in [0, 1]$ бўлса, куйида келтирилган (2.1) ифодани қаноатлантиради.

$$\sum_{i=1}^l g_i = 1 \quad (2.1)$$

Бунда $1 \leq i \leq l$ да q сўров ва i -чи зонанинг муносабати s_i – эҳтимолли катталиқ бўлсин. Агар сўровнинг барча терминлари битта зонага тегишли бўлса, 1 га тенг акс ҳолда, 0 га тенг. Ҳақиқатан ҳам буни $\{0, 1\}$ тўпланининг зонасида терминларга асосланган ихтиёрий мантикий функция қаноатлантириши мумкин. Бундай вазли зонани

$$V = \sum_{i=1}^l g_i s_i$$

каби белгилаб оламиз. Долзарбликнинг вазли зоналарини долзарблик асосида эҳтимолли излаш деб айтаемиз.

Вазнли зонанинг долзарблилигини ҳисоблаш учун ҳар бир ЭРнинг долзарблигини кетма-кет ҳисоблаш орқали ҳар хил зоналарнинг йиғиндиси кўринишида амалга оширамиз. Тартибланган индекс орқали ҳам вазнли зонанинг долзарблигини ҳисоблаш мумкин. q_1 ва q_2 терминлардан иборат q сўров берилган бўлсин. Агар иккала термин ҳам бир зонага тегишли бўлса, 1 акс ҳолда 0 (2.1-алгоритм). Алгоритмни мураккаб сўровлар ва мантиқий функциялар учун кенгайтирилган вариантларини ҳам ишлаб чиқиш мумкин.

2.1-алгоритм. Вазнли зонанинг долзарблилигини ҳисоблаш.

```
ZonaVazn( $q_1, q_2$ ) {
    double vazn[N] = {0}; constant g[];
     $p_1 = \text{postings}(q_1); p_2 = \text{postings}(q_2);$ 
    while ( $p_1 \diamond \text{null} \ \&\& \ p_2 \diamond \text{null}$ ) {
        if  $\text{docID}(p_1) == \text{docID}(p_2)$  {
             $\text{score}[\text{docID}(p_1)] = \text{VaznZone}(p_1, p_2, g);$ 
             $p_1 = \text{next}(p_1); p_2 = \text{next}(p_2);$  }
        else if  $\text{docID}(p_1) < \text{docID}(p_2)$   $p_1 = \text{next}(p_1);$ 
            else if  $\text{docID}(p_1) > \text{docID}(p_2)$   $p_2 = \text{next}(p_2);$ 
    } return vazn; }
```

Вазнли зоналарнинг долзарблилигида g_i ларнинг қийматини қандай ҳисоблаш муаммоси мавжуд? Бундай ҳисоблашни экспертлар (фойдаланувчилар) орқали амалга ошириш мумкин. Лекин, жуда кўп ҳолларда қийматлар олдиндан баҳоланган, ўқув мисоллари асосида аниқланади. Бу усул маълумотларни излашда машинали ўқувга асосланган долзарблик усулларининг умумий 2 та хусусиятига тааллуқлидир. Булар:

1. Ўқув мисолларига асосланган мисоллар тўплами мавжуд бўлсин. Унда ҳар бир элемент аниқ ифодадан ташкил топади. Аниқ ифода эса q сўров, d ЭР ҳамда q ва d лар учун долзарблик қиймати бўлсин. Энг содда вариантда долзарблик бинарли деб қаралади.

2. g_i қийматлар ўқув асосида аниқланади, чунки олинаётган қиймат долзарблик қийматини аппроксимациялаган бўлиши керак.

Долзарбликнинг вазнли зоналари учун ушбу жараёни сараланган коэффицентли чизикли функция сифатида қараш мумкин. Коэффицентлар эса зонага тегишли мантиқий қийматлардир. Бу ёндашувнинг машаққатли томони маълумотлар тез-тез ўзгариб турса (вебга ўхшаш) ўқув асосида долзарбликнинг қийматини йиғишдир.

Долзарбликнинг вазли зонаси учун куйидаги масалани оламиз. Ҳар бир ЭР сарлавҳа ва аннотация зоналарига эга бўлсин. Берилган q сўров ва d ЭР учун $s_T(d, q)$ ва $s_B(d, q)$ мантикий функцияларни аниқлаш керак бўлсин. Қайсики, мантикий функциялар q сўров d ЭРнинг сарлавҳаси (аннотацияси) бўлиши ёки бўлмастлигини аниқласин. $[0,1]$ оралиғидаги долзарблик қийматини ҳар бир маълумот-сўров жуфтлиги асосида ҳисоблаш учун $s_T(d, q)$ ва $s_B(d, q)$ ларнинг қийматлари ва $g \in [0,1]$ ўзгармасдан фойдаланиб куйидагича ёзамиз:

$$r(d, q) = g \cdot s_T(d, q) + (1 - g) \cdot s_B(d, q) \quad (2.2)$$

$\Phi_j = (d_j, q_j, r(d_j, q_j))$ учлик асосида аниқланган ўқув мисоли асосида g ўзгармасни аниқлаймиз. Ҳар бир ўқув мисоли d_j маълумот, q_j сўров ва $r(d_j, q_j)$ долзарблик қиймати асосида аниқланади (2.2-жадвал).

2.2-жадвал

Ўқув мисоллари

Ўқув мисоли	DocID	сўров	s_T	s_B	Долзарблик қиймати
Φ_1	1	Тинчлик	1	1	1
Φ_2	28	Компьютер	0	1	0
Φ_3	80	Китоб	0	1	1
Φ_4	231	Маълумот	1	0	0
Φ_5	344	Ахборот	1	0	1
Φ_6	1991	Излаш	0	1	1
Φ_7	1177	Долзарб	0	1	0

Ҳар бир Φ_i ўқув мисоли учун $s_T(d, q)$ ва $s_B(d, q)$ мантикий қийматлар мавжуд. Шунинг учун долзарбликни ҳисоблаш учун (2.2) формуладан фойдаланамиз.

$$\text{Result}(d_j, q_j) = g \cdot s_T(d_j, q_j) + (1 - g) \cdot s_B(d_j, q_j) \quad (2.3)$$

Инсон фактори оркали келтирилган долзарблик қиймати ва (2.3) долзарблик қийматларини таққослаш мумкин. Қиймат хатолигининг функциясини куйидагича аниқлаймиз.

$$\varepsilon(g, \Phi_j) = (r(d_j, q_j) - \text{Result}(d_j, q_j))^2$$

Бунда инсон фактори 0 ёки 1 билан ифодаланади. Бу хатоларни йиғиб, ўқув мисоллари учун куйидагига эга бўламиз.

$$R = \sum_j \varepsilon(g, \Phi_j) \quad (2.4)$$

Ўқув мисоллари асосида g ўзгармасни аниқлаш мумкин. Бунинг учун (2.4) ифодани энг кичик бўлган қийматларидаги g ни аниқлаш керак.

Ҳақиқатан ҳам, (2.4) ифода ёрдамида g нинг энг яхши қийматларини аниқлаш мумкин.

g нинг энг яхши қийматини ҳисоблашда ихтиёрий Φ_j учун $s_T(d, q) = 0$ ва $s_B(d, q) = 1$ берилган бўлсин. Долзарблик (2.3) ифода орқали ҳисобланса $1 - g$ га тенг бўлади. (2.4) ифодани $s_T(d_j, q_j)$ ва $s_B(d_j, q_j)$ ларнинг бошқа вариантлари учун ҳам ёзиш мумкин (2.3-жадвал).

2.3-жадвал

s_T ва s_B ларнинг комбинацияси

$s_T(d_j, q_j)$	$s_B(d_j, q_j)$	Долзарблик
0	0	0
0	1	$1 - g$
1	0	g
1	1	1

$s_T(d, q) = 0$ ва $s_B(d, q) = 1$ бўлган, инсон фактори долзарб (1) (мос равишда, долзарбмас (0)) ўқув мисоллари сонини n_{01r} (мос равишда n_{01n}) билан белгилаймиз. Шунда (2.4) ифода $s_T(d, q) = 0$ ва $s_B(d, q) = 1$ учун қуйидагича:

$$R = (1 - (1 - g))^2 n_{01r} + (1 - g)^2 n_{01n}$$

Худди шундай $s_T(d_j, q_j)$ ва $s_B(d_j, q_j)$ ларнинг бошқа комбинациялар учун (2.4) ифода орқали аниқланган йиғиндини ёзиш мумкин, яъни:

$$R = (n_{01r} + n_{10n})g^2 + (n_{10r} + n_{01n})(1 - g)^2 + n_{00r} + n_{11n} \quad (2.5)$$

(2.5) ифодани g бўйича соддалаштириб, нолга тенглаштириб, g нинг энг самарали қийматини аниқлаймиз.

$$R = \frac{n_{10r} + n_{01n}}{n_{10r} + n_{10n} + n_{01r} + n_{01n}}$$

Корпоратив ахборот-кутубхона тизимларининг САМларда маълумотларни излаш ва қайта ишлаш усулларида вазнли майдонларга асосланган ЭРнинг долзарблигини ҳисоблаш ва g баҳолашни аниқладик.

ИМнинг элементлари бўлмиш терминларга асосланиб, долзарбликни ҳисоблаш учун ЭР частотаси ва векторли тўпламлар моделидан самарали фойдаланиш мумкин. Шунинг учун корпоратив ахборот-кутубхона тизимларининг САМларда маълумотларни излаш ва қайта ишлаш учун ЭР частотаси ва векторли тўплам моделига асосланган усулларни тадқиқ қилиш ҳамда модификацияланган вариантларини ишлаб чиқишга эҳтиёж мавжуд.

2.2. Маълумотларни излаш тизимларида электрон ресурс ва термин частоталарининг вазнини ҳисоблаш усуллари

КАКТларда электрон ресурснинг долзарблиги унинг зоналарига боғлиқ. Сўровларда кўп учрайдиган терминнинг зонаси ёки маълумоти учун сўровнинг долзарблигини инобатга оламиз ва унга катта қиймат берамиз. Бу фараз сўровларни эркин киритиш учун қулай, яъни МИТларда сўровларнинг эркин, ёрдамчи амалларсиз киритиш (мантикий амаллари каби) назарда тутилган. Бу ёндашув кўпроқ веб технологияларга асосланган ва унга сўров терминларнинг тўплами сифатида қаралади. Ҳақиқатан ҳам, ЭРлар кўрсаткичларини ҳисоблаш учун сўровдаги терминларнинг маълумот қийматларини йиғиш кифоя [63].

Ушбу терминнинг ҳар бир маълумот қийматини бирлаштириш, яъни ЭРдаги терминнинг учрашиш сонига боғлиқ бўлиб, унинг вазнини аниқлашга олиб келади. Бундан эса, d ЭРдаги t терминнинг вазни орқали q сўров ва d ЭРнинг ўзаро муносабати аниқланади. Бу вазни d ЭРдаги t терминнинг сони билан ҳам белгилаш мумкин. Бу алгоритм терминнинг частотаси (term frequency) деб юритилади ва $t_{t,d}^f$ каби белгиланади [90; 79-б.]. Бунда индексдаги t – термин ва d – ЭРни билдиради.

Бирор d ЭР учун D вазнлар тўпамини (ихтиёрий бирор функция ёки d ЭРдаги t терминнинг сони орқали аниқланган ва ҳар доим мусбат, бутун сонлар) d ЭРнинг ифодасини сонлар кўринишида изоҳлаш мумкин. Бундай ёндашув илмий журналларда термин копи (bag of words model) деб юритилади [90; 80-б.]. Бу модель асосида ҳар бир терминнинг ЭРда учрашиш сони ЭРдаги терминлар тартиби билан изоҳланади. Бунда фақат ҳар бир терминнинг учрашиш сони ҳақидаги маълумотдан фойдаланамиз. Маълумотдаги барча терминларнинг вазнлари бир хилми, деган саволга жавоб излайлик. Ҳақиқатан ҳам, маълумотдаги терминларнинг вазнлари бир хил эмас. Чунки унда қўшимча сўзлар (stop

words), индекслаштириш инобатга олинмайдиган терминлар, излаш ва долзарбликка таъсир этмайдиган терминлар мавжуд.

Одатда, сўровдаги терминларнинг барчаси муҳим ўрин тутати, деб қаралади. Юқорида келтирилганлар асосида бундай ёндашувда камчиликлар мавжудлигини кўриш мумкин. Аммо, аслини олганда сўровдаги терминларнинг аҳамияти турлича, яъни кам ва баъзан ноль қийматли терминлар ҳам учрайди. Фараз киламиз, ахборот технологиялари учун ЭРлар тўплами мавжуд бўлсин. Бу тўпламнинг деярли барча маълумотларида “ахборот технологиялари” терминини учратиш мумкин. Ушбу камчиликни бартараф этиш учун “кам таъсирчан термин” тушунчасини киритамиз, унда ЭРда тез-тез учраб турувчи ва ЭРнинг долзарблигига таъсир кўрсатмайдиган термин назарда тутилади. Тўпламда катта частотали терминларнинг вазнини камайтириш ғояси пайдо бўлади. Бу ғоянинг асоси, t^f термин вазнини коэффиценти оркали кичиклаштириш, яъни унинг тўпламдаги частотасининг катталашини ҳисобига.

Бу ғоянинг ўрнига кўп ҳолларда d_i^f “ЭР частотаси” (document frequency) тушунчаси ишлатилади [90; 80-б.], унда t терминдан иборат бўлган d_i^f тўпламдаги маълумотлар оркали ифодаланadi. Бу сўров долзарблигига нисбатан ЭРлар орасидаги фарқни топиш оркали изоҳланади. Аммо, маълумотларнинг ўз статик кўрсаткичлари оркали фойдаланиш самаралироқ. Агар, берилган терминга мос ЭРлар сони тўпламдаги термин частотаси ва ЭР частота 3.4-жадвалдагидек берилган бўлса, унда t^f тўпламдаги термин частотаси ва d^f ЭР частотаси турлича бўлиши мумкин. 2.4-жадвалда хусусий ҳол учун терминларнинг частотаси бир-бирига яқин, аммо ЭР частоталари анча фарқли. Бундан эса, “биринчи” терминни сакловчи маълумотлар шу терминли сўровлар муносабатига караганда “иккинчи” терминни сакловчи ЭРларнинг шу терминли сўровлар муносабати катта аҳамиятга эга.

2.4-жадвал

Тўплам ва электрон ресурсдаги частота

Термин	t^f – тўпламдаги частота	d^f – ЭРдаги частота
Биринчи	128	111980
Иккинчи	125	772005

ЭР частотаси вазнини яхшилашни қандай амалга ошириш керак? деган савол пайдо бўлади. Бунинг учун тўпламдаги ЭРлар

сонини N билан белгилаймиз ва t терминнинг тескари ЭР частотасини қуйидагича аниқлаймиз:

$$d_i^{i,f} = \log \frac{N}{d_i^f} \quad (2.6)$$

Кўп учрайдиган терминларга нисбатан кам учрайдиган терминларнинг тескари электрон ресурс частотаси катта бўлиб ҳисобланади (2.5-жадвал). Бунда тўпلامдаги электрон ресурслар сони $N = 32021177$ ва (2.6) ифода билан ҳисобланган.

2.5-жадвал

Тескари маълумот частотасига мисол

Термин	t_i^f	d_i^f
Бир	1980	4.209
Икки	1981	4.21
Уч	2004	4.20355
Тўрт	2005	4.2033

Терминнинг тескари электрон ресурс частотаси ва частоталар комбинациялари асосидаги вазни ҳисоблашда Ҳар бир d ЭРда ҳар бир t терминнинг вазнини аниқлаш учун d ЭРда t терминнинг комбинациялашган частотаси t^f ва тескари ЭР частотасини d^f ҳисоблаш лозим. Бунинг учун $f^{t,d}$ вазни ҳар бир t терминга мос d ЭР вазнини бирлаштириш асосида аниқлаймиз.

$$f^{t,d} = t_{i,d}^f \times d_{i,d}^f \quad (2.7)$$

d ЭРдаги t терминнинг $f^{t,d}$ вазни қуйидаги хусусиятларга эга:

1. Агар катта сонли бўлмаган ЭРларда t термин кўп марта учраса, у максимал қийматга эга бўлади. Бу шартнинг ўзи бошқа ЭРлардан фаркланишини билдиради.

2. Агар қайсидир ЭРда фақатгина бир неча марта учраса ёки кўп ЭРларда учраса, у кичиклашади.

3. Агар термин барча ЭРларда учраса, у минимал қийматга эга бўлади.

Ҳар қандай ЭРни вектор сифатида шакллантириш мумкин, унда ҳар бир элемент учун мос луғатдаги термин ва вазн (2.7) ифода билан ҳисобланади. ЭРда учрамайдиган луғатдаги терминларнинг вазни 0 га тенг. Бу векторли ифодалаш вазн ва долзарблик учун муҳим. d ЭРнинг аҳамиятлилиги шундаки, сўровнинг барча терминлари жорий ЭРга тегишлигидадир. Аммо, сўровдаги ҳар бир

терминнинг ЭРга тегишлилигини йиғиш эмас, балки уларнинг $f^{t,d}$ вазини йиғиш керак, яъни:

$$R(q, d) = \sum_{t \in d} f^{t,d} \quad (2.8)$$

ЭРнинг долзарблиги учун векторли тўплам моделини аниқлаймиз. Юқорида ЭРни вектор кўринишида тасвирлашни келтирдик, яъни терминларнинг ЭРга бўлган аҳамиятлилик даражаси кўринишида. Векторли тўпламда ЭРлар тўпланининг вектор кўринишида тасвирланиши векторли тўплам модели (vector space model) деб юритилади [90; 83-6.]. МИТларнинг кўп масалалари учун фундаментал бўлиб, сўровга нисбатан ЭРнинг долзарблиги, ЭРларни синфлаштириш ва кластеризациялаш масалаларида асос ҳисобланади. Аввал, векторли тўпламда ЭРнинг долзарблигини тадқиқ қиламиз. Бунда асосий масала ЭРлар тўплами учун аниқланган векторли тўпламда сўровларни векторли кўринишда шакллантиришдир.

Луғатнинг ҳар бир t термини учун алоҳида элементга эга d ЭР асосида қурилган векторни $\vec{V}(d)$ билан белгилаймиз, ҳар бир элемент $f^{t,d}$ вазн асосида ҳисобланади. Аммо бу вазн ҳеч қандай аҳамиятга эга эмас. Агар ҳар бир термин учун алоҳида ўқ мос бўлса, тўпламдаги ЭРлар тўпланини векторли тўпламда векторли кўринишда тасвирлаш мумкин. Бундай ҳолда ҳар бир ЭРдаги термин тартиби ҳақидаги маълумот йўқолиши мумкин, худди термин копидек.

Векторли тўпламда сонли кўринишда икки ЭРнинг ўхшашлигини қандай ажратиш мумкин? Агар ҳар хил ЭРлар учун векторлар фарқининг модули кўринишида тахмин қилинса, бундай ўхшашлик, агар улардан бири иккинчисидан анча узоқ бўлиши, икки ЭР учун векторли фарқ жуда яқин мазмун билан бўлиши мумкин. Бошқача айтганда, икки ЭРдаги терминлар бир хил бўлиши мумкин, аммо терминнинг частотаси уларнинг бирига нисбатан катта, бирига нисбатан кичик бўлади.

Иккита ЭР d_1 ва d_2 учун маълумот узунлиги таъсирини сусайтириш учун уларнинг $\vec{V}(d_1)$, $\vec{V}(d_2)$ векторли кўринишлари учун конусли ўхшашлик (cosine similarity) асосида ҳисобланади [90, 84].

$$\text{sim}(d_1, d_2) = \frac{(\vec{V}(d_1), \vec{V}(d_2))}{\|\vec{V}(d_1)\| \|\vec{V}(d_2)\|} \quad (2.9)$$

Бунда ифодада ($\bar{V}(d_1), \bar{V}(d_2)$) векторларнинг скаляр кўпайтма кўриниши тасвирланган ва махражда бу векторларнинг Евклид нормасининг кўпайтмаси келтирилган. (\bar{x}, \bar{y}) векторларнинг скаляр кўпайтмаси куйидагича аниқланади:

$$(\bar{x}, \bar{y}) = \sum_{i=1}^M x_i y_i$$

d ЭРнинг $\bar{V}(d)$ вектор элементларини $\bar{V}_1(d), \dots, \bar{V}_M(d)$ каби белгилаймиз ва векторнинг Евклид нормаси учун куйидаги ифода ўринли бўлади:

$$D_{norm} = \sqrt{\sum_{i=1}^M \bar{V}_i^2(d)}$$

(2.9) ифоданинг махражидagi $\bar{V}(d_1)$ ва $\bar{V}(d_2)$ векторларнинг узунлик асосида нормаллаштирсак, яъни уларнинг узунликлари 1 га тенг ва $\bar{v}(d_1) = \frac{\bar{V}(d_1)}{\|\bar{V}(d_1)\|}$, $\bar{v}(d_2) = \frac{\bar{V}(d_2)}{\|\bar{V}(d_2)\|}$ ифода ўринли бўлсин.

Бундан (2.9) ифодани куйидагича ёзиш мумкин:

$$sim(d_1, d_2) = (\bar{v}(d_1), \bar{v}(d_2))$$

Фараз қилайлик 2.6-жадвал асосида 3 та ЭР берилган бўлсин.

2.6-жадвал

Бошланғич маълумотлар

Терминлар	doc_1	doc_2	doc_3
Биринчи	28	16	25
Иккинчи	1	11	1
Учинчи	1980	2005	2004
Тўртинчи	25	12	1981

Улар термин частотасининг векторига Евклид нормасини қўллаймиз. Натижада

$$\bar{V}(doc_1) = \{28, 1, 1980, 25\} = 1980.35,$$

$$\bar{V}(doc_2) = \{16, 11, 2005, 12\} = 205.12,$$

$$\bar{V}(doc_3) = \{25, 1, 2004, 1981\} = 2817.97$$

кийматлар ўринли бўлади. Бу векторларнинг Евклид нормаси куйидагича бўлади (2.7-жадвал):

Натижалар

Терминлар	doc_1	doc_2	doc_3
Биринчи	0.014	0.0079	0.0088
Иккинчи	0.0005	0.0054	0.0003
Учинчи	0.9998	0.9999	0.711
Тўртинчи	0.01	0.005	0.702

ЭРларни вектор сифатида ифодалашнинг яна бир аҳамиятли томони бор. Бу сўровларни ҳам вектор сифатида қараш мумкин. Ҳар бир ЭРнинг долзарблик кийматига қўшиш учун скаляр кўпайтириш амалга оширилади.

$$\text{Result}(q, d) = (\bar{V}(q), \bar{V}(d))$$

Сўровга энг мос аҳамиятлилигини танлаш учун конусли ўхшашликдан фойдаланамиз, (3.8) ифодани куйидагича ёзиб оламиз:

$$\text{Result}(q, d) = \frac{(\bar{V}(q), \bar{V}(d))}{\|\bar{V}(q)\| \|\bar{V}(d)\|}$$

Одатда, кўп ҳолларда ЭРлар тўплами векторлар орқали ифодаланади, худди шундай эркин матнли сўровлар учун ҳам векторли ифодалаш амалга оширилади[99]. Бунда сўров учун долзарбликка эга бўлган K та ЭРлар тўплами бўлиб, аниқлашни кўриб чиқамиз. K та ЭРлар тўплами долзарблик бўйича камайиб бориш тартибида сараланган бўлади. Жуда кўп МИТларида $K = 10$ деб олинади, дастлабки 10 та ЭР сўров учун долзарб ҳисобланади.

Берилган ЭР учун 20 та терминнинг ЭРда 20 мартадан кўпроқ учратиш эҳтимоли кам бўлади. Бунда термин частотасини санаш орқали ҳисоблаш ўринсиз. Шунинг учун уни ҳисоблашга оид жуда кўп тажрибалар олиб борилган, тадқиқотлар натижасида термин частоталари умумий ўзгаришини логарифмик термин частоталарига асосланиб, ҳисоблаш амалга оширилган[90].

$$w^f = \begin{cases} 1 + \log(t_{i,d}^f), & \text{агар } t_{i,d}^f > 0 \\ 0, & \text{акс ҳолда} \end{cases}$$

Шунингдек, юқорида келтирилган ёки ихтиёрий t^f термин частотасини w^f логарифмик термин частотаси билан ўзгартириш мумкин:

$$f^{w,d} = w_{i,d}^f \times d_{i,d}^f$$

Бундан эса (2.11) ифодани ўзгартириш мумкин бўлади.

Вазнларни ҳисоблашда ЭРдаги терминларнинг вазнларини катта қийматли терминларга нисбатан нормаллаштириш кўп тадқиқ қилинган [90]. Тўпламдаги ҳар бир d ЭР учун қуйидаги ифода ўринли:

$$t_{\max}^f(d) = \max_{t \in d} t_{t,d}^f$$

Бунда $\tau - d$ ЭРдаги терминларнинг ўтиш индекслари. d ЭРдаги t терминнинг нормаллаштирилган частотасини қуйидагича аниқлаш мумкин:

$$t_{t,d}^{f_n} = a + (1-a) \frac{t_{t,d}^f}{t_{\max}^f(d)} \quad (2.10)$$

Бу ерда $a - 0$ ва 1 оралиғидаги катталиқ (кўп ҳолларда $a=0.4$ олиш тавсия қилинади [90; 87-б.]). (2.10) ифодада a сил-ликлаштириш коэффиценти деб юритилади. Унинг мақсади икки хаднинг қийматини камайитиришдан иборат.

Термин частоталарни нормаллаштириш катта ЭРлардаги термин частоталарини ҳисоблашдаги ноқулайликлардан қочиш учун ишлатилади. Бунинг учун d ЭРни нусхалаймиз ва d' билан белгилаймиз. Берилган сўровга d' ЭРнинг долзарблиги d ЭРниқидан катта бўлмайди.

Катта термин частотага нисбатан нормаллаштиришнинг камчиликлари мавжуд. Буларга:

- бу усул тургун эмас, доим бир хил бўлавермайди. Агар боғловчи сўзлар рўйхатини ўзгартириш, ЭРдаги терминлар частотасига таъсир қилса, долзарблигига ҳам таъсир қилади. Буни таъминлаш мураккаб жараёндир.

- ЭР мазмунига таъсир қилмайдиган нормал бўлмаган катта термин частотасининг бўлиши.

- Кам такрорланадиган терминлар кам учрайди. Аммо уларнинг термин частотаси ЭР учун муҳим аҳамиятли бўлиши мумкин.

Векторли тўпламнинг долзарблиқни аниқлаш усулларида $\bar{V}(q)$ ва $\bar{V}(d)$ векторларнинг вазнларини аниқлаш бир-биридан фарқланади (2.8-жадвал).

$f^{t,d}$ вариантлар учун SMART усуллар

Термин частотаси	Маълумот частотаси	Нормаллаштириш
$t_{t,d}^f$	I	I
$1 + \log(t_{t,d}^f)$	$\log \frac{N}{f^{t,d}}$	$\frac{1}{\sqrt{w_1^2 + w_2^2 + \dots + w_M^2}}$

2.8-жадвалда $\bar{V}(q)$ ва $\bar{V}(d)$ векторлар вазнини ҳисоблашнинг асосий схемаларига асосланган вазнларни ҳисоблаш ва ҳар хилли танлашлар SMART деб юритилади ва *ddd.qqq* кўринишга эга. Унда биринчи учлик ЭР вектори учун терминлар характерлари, иккинчи учлик сўров учун терминлар характерларини билдиради. Ҳар бир учликдаги биринчи белги термин частотаси, иккинчиси ЭР частотасини, учинчиси эса нормаллаштириш усулини билдиради. Кўп ҳолларда $\bar{V}(q)$ ва $\bar{V}(d)$ векторлар турлича нормаллаштирилади. Агар илғор схемалардан *ddd.qqq* ҳисобланса, унда маълумотларнинг векторини яратиш учун логарифмик вазнли частотадан фойдаланилади, лекин тескари маълумотли частотадан фойдаланилмай, конусли нормаллаштиришдан фойдаланилади. Аммо сўровларда логарифмик вазнли частотадан, тескари ЭР частота ва конусли нормаллаштириш ишлатилади.

ЭРларнинг векторлари бир хил ягона узунликка эга бўлиши учун ҳар бир вектор Евклид нормаси билан нормаллаштирилди. Шунингдек, жорий ЭРнинг барча маълумотларини истесно тариқасида бирлаштирилади. Бу маълумотлар узун маълумотларни қайта ишлашда керак бўлиши мумкин.

Биринчидан, узун маълумотларда кўп терминлар бўлади, шунга етарли даражада уларнинг t^f термин частоталари ҳам катта бўлади.

Иккинчидан, узун маълумотлар ҳар хил терминларга эга бўлади. Бу фактлар узун бўлган маълумотларга бўлган эҳтиёжни кондириши мумкин. Узун маълумотларни икки тоифага бўлиш мумкин: 1) кўп сўзли маълумотлар мазмунан баъзан қайтариладиган, яъни бу ЭРларда узунлик терминларнинг вазнига таъсир қилмайди; 2) маълумотлар кўп мавзули бўлса, яъни ЭРдаги терминлар вазни турли-туман.

Булар учун самарали бўлган узун маълумотни нормаллаштириш усулидан фойдаланилади, бу термин частотаси ва ЭР

частотасига ҳам боғлиқ эмас. Бунинг учун тўпламдаги ЭРларнинг векторлари учун нормаллаштириш киритилади, яъни векторлар бирга тенг бўлмаган “нормаллаштирилган” тушунчаси киритилади. Шунда сўров вектори ва нормаллаштирилган ЭР векторининг скаляр кўпайтмасини ҳисоблаётганда узун маълумотлигини инobatга олиш учун аниқлаштирилади. Бу усул узун маълумотларни таянч нормаллаштириш дейилади.

2.3. Тўлиқ функционалли маълумотларни излаш тизимларида долзарбликни ҳисоблаш усулларига эвристик ёндашув

Олдинги параграфларда КАКТнинг ЭРдаги терминларнинг вазнлари орқали уларнинг долзарблигини аниқлаш усуллари келтирилди. Унда конусли ўхшашликни ҳисоблашнинг асосий алгоритмлари оддий векторли тўплам модели орқали ифодаланишга олиб келинди. Ушбу ҳисоблашни эвристик ёндашув асосида тезлаштириш усуллари ўрганамиз. Бундай усулларнинг кўпчилиги излаш тезлигини K та энг яхши ЭРлар тўпланими аниқлашдан эмас, балки сўровга нисбатан тасодифий танлаш ҳисобига ошади. Баъзи эвристик усуллар конусли ўхшашликка асосланган долзарблик усуллари умуллаштиради. Биринчи бўлиб, фан ва таълимга оид тўлиқ функционалли МИТларнинг барча элементларини аниқлаймиз. Бунинг учун МИТларда долзарблик масаласини умумий ечимлари учун конусли ўхшашликни ҳисоблашга қайтамиз. Унда долзарбликни қўллаб-қувватлаш учун фақатгина конусли ўхшашлик асосидаги индекс ва тузилмадан эмас, балки долзарбликнинг умумий омиллари, терминларнинг яқинлигидан фойдаланилади. Векторли тўплам модели ва оммабop сўровлар тилининг операторлари асосида эркин матнли сўровларнинг ўзарo боғлиқлигини келтирамиз.

Долзарбликни ҳисоблашнинг тезкор усулида $q = \{t_1, t_2\}$ сўров берилган бўлсин. Бунда қуйидагиларни аниқлаш мумкин:

1. Ягона $\bar{v}(q)$ вектор фақатгина 2 та нолдан фарқли элементни сақлайди.

2. Агар сўровдаги терминларнинг вазни қўлланилмаса, уларнинг қийматлари бир хил қийматга тенг бўлади.

Бу сўровга мос ЭРларнинг долзарблиги учун тўпламдаги ЭРнинг вазни керак. Бунинг учун $\bar{v}(d_1)$ вектор ва ягона $\bar{v}(q)$ векторлар учун конусли ўхшашликни ҳисоблаш етарли. Бунда $\bar{v}(q)$ векторнинг нолдан фарқли барча элементлари 1 га тенглаштирилади. Ихтиёрий 2 та d_1 ва d_2 ЭР учун қуйидаги ифода ўринли:

$$(\bar{V}(q), \bar{V}(d_1)) > (\bar{V}(q), \bar{V}(d_2)) \Leftrightarrow (\bar{v}(q), \bar{v}(d_1)) > (\bar{v}(q), \bar{v}(d_2))$$

Ихтиёрий d ЭР учун $(\bar{v}(q), \bar{v}(d))$ конусли ўхшашлик маълумотда мавжуд сўровдаги терминларнинг вазни йиғиндисини тасвирлайди. Буни эса сўз жойларининг бирлашмасини аниқлаш орқали ҳисоблаш мумкин. Бу алгоритмнинг модификациялангани 3.2-алгоритмда келтирилган. Бу алгоритмнинг фарқи шундаки, 7 кадамдаги $w_{i,d}$ ни бирга тенглаштириш, кўпайтириш ва бўлиш амаллари фақатгина бўлиш амалига айлантирилган. Алгоритмда тартибланган индексдан сўровдаги терминларнинг жойларига ўтиш зарур, ҳар бир ЭР учун аҳамиятликнинг жами қийматини аниқлаш керак. Сўнг мантикий сўровларни қайта ишлагандек [97], ЭРларнинг долзарблик қийматлари аниқланади. Ҳар бир ЭР тўпламининг лугати учун тескари d^f ЭР частотаси ва ҳар бир термин жойи учун эса t^f термин частотаси аниқлан. Бу имконият ихтиёрий сўров термини учун термин жойлари мавжуд ихтиёрий маълумотнинг долзарблигини аниқлаш имкониятини беради.

Долзарбликни ҳисоблаш ва фойдаланувчига тақдим этиладиган ЭРларнинг рўйхатини чиқаришни ҳам инобатга олиш керак. Бунинг учун K та ЭРлар тўпламини аниқлаш лозим. ЭРларни долзарблик бўйича саралаш мумкин, аммо уни пирамидали саралаш орқали энг яхши K та ЭРни танлаши лозим. Конусли ўхшашлик ноль бўлмаган J сондаги ЭР берилган бўлсин. Бунда пирамидали саралаш учун $2J$ муносабат амали, ундан K та ЭРни ажратиб олиш учун $\log J$ муносабат амалини бажариш керак.

2.2-алгоритм. Векторли тўпланда тезлаштирилган долзарбликни ҳисоблаш алгоритми.

```
Vazn(q){
double vazn[N]=0;
foreach (q as items=>d)
  lenD = length(d);
while t ∈ q do
  while (d,tft,d) ∈ InvList do vazn[d] += wft,d;
Len = length(d);
While d do vazn[d] = vazn[d]/Len;
Return sort(k, vazn[d]); }
```

Шу пайтгача сўровни қаноатлантирувчи K та энг яхши ЭРларни излашга эътибор қаратган эдик. Энди K та энг яхши ЭРлар рўйхатини шакллантириш усуллари кўриб чиқамиз. Фойдаланувчилар учун ушбу имконият сўров натижаларини шакллантириш

ва ҳеч қандай таъсирсиз излаш тезлигини пасайтирмаслиги керак. Ҳақиқатан ҳам кўп дастурларда етарлича энг яхши K та ЭРларни жуда яқинликлари, ўхшашлари мавжудлигини кўриш мумкин.

Фойдаланувчи нуктаи назаридан оладиган бўлсак, K та ЭРни тахминий излаш ҳар доим ҳам ёмон эмас. Бундан эса конусли ўхшашлик асосида K та ЭРлар рўйхатига ишонч қолмайди. Конусли ўхшашлик фақатгина ўзича ЭРнинг сўровга аҳамиятлилигини изоҳлайди, холос.

Эвристик усулларда конусли ўхшашлик асосида аҳамият-лилик билан энг яхши K та ЭРнинг ўхшашлик, яқинлиги асосида K та ЭРларни аниқлаш муҳим. Асосий ҳисоблашлар сўров ва ЭРлар тўпламининг ўзаро аҳамиятлигига боғлиқ. Аслида катта ЭРлар тўпланимига ўхшаш, яқин ЭРлар учун пирамидали саралаш қўлланилса, алгоритмнинг амаллар сони ошиб кетади.

Катта сонли ЭРлар тўпламида конусли ўхшашликни ҳисобла-масдан туриб, K та ЭРни аниқлаш мумкин. Бундан эвристик усуллар икки босқичда амалга оширилади:

1. Шундай A ЭРлар тўплами мавжуд бўлса, энг яхши K та ЭР учун талабгор бўлади. $K < |A| < N$. Бу тўплам q сўровга нисбатан K та ЭР аҳамиятлигининг таъсири йўқ. Аммо унда етарлича шартни каноатлантирувчи ЭРлар бўлиши керак.

2. A тўпламда K та ЭРларга аҳамиятлиси мавжуд бўлсин.

Тадқиқотлар натижасида КАКТларининг тўлиқ функционалли МИТларида долзарбликни эвристик ҳисоблаш усуллари учун мезонлар:

индексни қисқартириш усули учун кўп терминдан иборат q сўров учун шу терминлардан бири учраган ЭРларни танлаш етарли. Бу жиҳатни эвристик коидалар билан бойитган ҳолда фойдаланиш мумкин.

Энг яхши рўйхат усули энг яхши рўйхат ёки топ ЭР деб юритилади. Бунинг асосий ғояси d ЭРлар тўпламидан луғатдаги t терминга муносабатли бўлган энг юқори вазнли ЭРларнинг N (олдиндан аниқланган ва $N < d$) таси. q сўров берилган бўлса, энг яхши рўйхатни A билан белгилаймиз. A рўйхатни q сўровдаги ҳар бир термин учун энг яхши рўйхатини бирлаштириш оркали аниқлаймиз.

Статик ва тартибланган индекс усули энг яхши рўйхатнинг кенгайтирилгани бўлиб, сўровга боғлиқ бўлмаган ҳолда танланади.

Аҳамияти бўйича тартибланган усул ЭРнинг термин частота орқали параллел саралаш йўли билан амалга оширилади. Ҳар бир термин учун аҳамиятли ЭРлар тўплами ҳосил қилинади ва бирлаштирилади.

Кластерли танлаш усулида ЭРлар векторларини кластерлаш орқали амалга оширилади. Сўровни қайта ишлаш вақтида фақатгина кластерланган тўплам қаралади [15].

Парсер сўров ва долзарблик усулида кўп ҳолларда сўровлар эркин киритилади. Чунки МИТ кўпроқ корпоратив тармоқда фойдаланилади, уларнинг фойдаланувчилари учун яхлит тизимли ёндашувни амалга ошириш мушкул. Эркин киритилган сўровларни индекслаш орқали излаш тезлигини ошириш мумкин. Бунинг учун қайси ҳолларда сўров индекслаштирилади, деган савол пайдо бўлади. Бунинг учун K та ЭР етарли бўлган сўровлар олинади. Агар берилган сўров учун (3 та терминли) ЭРлар тўплами олинса, унда K та ЭР бўлса, индексга ёзилади. Агар мавжуд бўлмаса, кейинги қадамга ўтилади.

Берилган сўровни 1-қадам қаноатлантирмаса, уни (3 терминлини) 2 та терминлига ($a, b, c \rightarrow a, b$ ва b, c ва a, c) кўринишдаги сўров шаклида олиб, аввал парсер сўровлардан, сўнгра ЭРлар тўпамидан изланади. Агар жами K та ЭР бўлса тугатилади. Агар мавжуд бўлмаса, кейинги қадамга ўтилади.

Берилган сўровни 2-қадам қаноатлантирмаса, уни (3 терминлини) 1 та терминлига ($a, b, c \rightarrow a$ ва b ва c) кўринишдаги сўров шаклида олиб, аввал парсер сўровлардан, сўнгра ЭРлар тўпамидан изланади.

Векторли тўпламда долзарбликка сўровлар тили операторларини таъсири ҳам мавжуд. Сўров тили операторлари МИТдаги сўровларни гуруҳлашга олиб келади. Буларга қуйидагилар киради:

1. Буль (мантиқий) сўров.
2. Вариантли сўров (* орқали).
3. Эркин сўров (ихтиёрий, табиий тил ёрдамида).

ЭР частотаси, термин частотаси, ЭР долзарбликлари кўп ҳолларда эвристик усуллар билан аниқланади. Шунинг учун МИТда излаш тезлигини ошириш, тўлиқлик ва аниқлик мезонларини сақлаб қолган ҳолда статистик маълумотлар асосида фан ва таълимга оид маълумотларни излаб, тадқиқ қилиш лозим.

2.4. Маълумотларни излаш тизимларида мантикий семантик излаш модели

Инсониятни фан ва таълимга оид энг ишончли, қимматли маълумотлар билан таъминлаш мақсадида фан ва таълимга оид КАКТнинг ЭРларини бошқариш ва улардан фойдаланиш жараёнларини автоматлаштириш масалалари билан бирга, бу тизимларда маълумотларни излаш, тақдим этиш каби масалалар ўз ечимини кутмоқда [70, 81, 125].

МИТни қуйидаги жуфтлик билан белгилаймиз:

$$\langle q, r \rangle$$

Бунда: q – сўров, яъни фойдаланувчининг ахборотга бўлган эҳтиёжи, r – сўровнинг объектлари, яъни сўровга мос аниқланган ЭРлар тўплами.

Тизим фойдаланувчилари томонидан табиий ва сунъий тилда шакллантириладиган сўровларни қайта ишлаш, керакли ЭРларни излаш муҳим ва мураккаб амалий ишдир. Чунки фойдаланувчилар кам сонли, аммо керакли, долзарб ЭРларни талаб қилишади.

МИТлари турли платформали ва технологияли ахборот инфраструктурадаги тизимлардан фойдаланиб яратилмоқда. Тадқиқотлар асосида кўпгина фойдаланувчилар МИТни баҳолашда қуйидаги омилларга эътибор берадилар:

- сўровга нисбатан топилган маълумотлар сони (n);
- сўровни қайта ишлаш вақти (v);
- сўровни шакллантириш ёки тизимнинг интерфейси каби параметрлар.

Фикримизча, ушбу параметрлар МИТни баҳолашда етарли омил эмас ёки иккинчи даражали баҳолаш параметрлари бўлиши мумкин ҳолос [96, 99]. Бизнинг тадқиқотларимиз асосида АРМАТ++ КАКТда маълумотлар излашни баҳолаш учун қуйидаги 3 та элементдан фойдаланиш лозим деб ҳисоблаймиз:

1. Тизимдаги ЭРлар сони (N).
2. Сўровлар (q).
3. Сўровларга ЭРнинг мослиги (M).

Бу элементлар асосида 2 та аниқлик ва тўлалик мезони асосида МИТни баҳолаш мумкин.

Аниқлик (P) деганда, топилган ЭРлар тўпламида айнан мос ЭРларнинг улуши тушунилади [45], яъни:

$$P = \frac{M}{N} \quad (2.11)$$

Тўлалик (K) эса, барча мос ЭРлар тўпламидан сўровга мос ЭРлар сони [45], яъни:

$$K = \frac{R}{M} \quad (2.12)$$

Бунда: R – сўровга мос электрон ресурслар сони. Мезонларни янада аниқроқ ифодалаш учун жадвал шаклида келтирамиз (2.9-жадвал).

2.9-жадвал

Аниқлик ва тўлалик мезони

	Мос ЭР сони	Мос келмаган ЭР сони
Топилган ЭРлар сони	T_p	F_p
Топилмаган ЭРлар сони	F_n	T_n

Ушбу жадвал асосида (2.11) ифода

$$P = \frac{T_p}{T_p + F_p}$$

ва (3.12) ифода эса

$$K = \frac{T_p}{T_p + F_n}$$

каби ёзиб оламиз. Жами тизимдаги ЭРлар сони эса

$$N = (T_p + F_p + T_n + F_n)$$

МИТнинг чинлиги куйидагича аниқланади:

$$T = (T_p + T_n) / N$$

Ушбу келтирилган мезонлар МИТни баҳолаш учун биринчи даражали параметрлар бўлиб ҳисобланади. Агар МИТ автостатик, интеллектуал излаш тизимига ўтказилса, ушбу мезонлар ўзини оқлайди.

Реал амалиётда оладиган бўлсак, бундай излаш фойдаланувчиларни қаноатлантирмайди. Чунки аниқ, аҳамиятли ЭРни излаш учун вақт керак, бу эса, ўз навбатида, қаралаётган ЭРлар тўпламининг куйидаги элементларига боғлиқ:

1. ЭРлар сонига.
2. Техник воситаларига (тезкор ва ташки хотира).
3. Ахборот хавфсизлиги усулига.
4. Тармоқ инфраструктурасига.

КАКТларида маълумотларни излаш учун сарфланадиган вақтни камайтириш учун мантиқий семантик излаш усулини таклиф қиламиз [59].

Фараз қиламиз, $q = \{q_i\}$ сўровлар тўплами ва q_i га мос $R^i = \{r_k^i\}$, $i = 1 \dots N_q$, $k = 1 \dots M_d$ ЭРлар тўплами берилган бўлсин. M_d – қаралаётган тизимдаги ЭРлар сони, N_q – тизимдаги сўровлар сони. Мантикий семантик излаш усулини бир жуфт сўров учун қуйидагича ёзамиз:

$$f(q_i, q_j) = \langle q_i | q_j \rangle = \frac{|R^j \cap R^i|}{|R^j|} \quad (2.13)$$

Бунда q сўровлар тўплами, $\langle a | b \rangle$ – a нинг b га семантик боғланганлиги ва ўхшашлигини билдиради ва $\frac{|R^j \cap R^i|}{|R^j|}$ амал билан ҳисобланади. $|R^j \cap R^i| \leq |R^j|$ шарт доим ўринли бўлиши лозим.

1-қоида. Агар $\langle q_i | q_j \rangle = 0$ бўлса, q_i , q_j га семантик боғланмаган ва бир бирига ўхшаш эмас, деб ҳисобланади.

Исботи: фараз қиламиз, $q = \{q_1, q_2\}$ сўровлар берилган. Маълумотларни излаш натижасида q_1 сўровга мос $R^1 = \{1, 45, 25\}$, q_2 сўровга мос $R^2 = \{100, 35, 15, 5\}$. Бунда $|R^1| = 3$, $|R^1 \cap R^2| = \emptyset$ лигидан $\frac{|R^1 \cap R^2|}{|R^1|} = \frac{\emptyset}{3} = 0$ келиб чиқади.

2-қоида. Агар $\langle q_i | q_j \rangle = 1$ бўлса, q_i , q_j га тўлиқ семантик боғланган ҳисобланади ва ўхшашлиги 1 (q_i айнан ўхшаш) деб ҳисобланади.

Исботи: фараз қиламиз, $q = \{q_1, q_2\}$ сўровлар берилган. Маълумотларни излаш натижасида q_1 сўровга мос $R^1 = \{1, 45, 25, 100\}$, q_2 сўровга мос $R^2 = \{100, 45, 25\}$. Бунда $|R^2| = 3$, $|R^1 \cap R^2| = 3$ лигидан $\frac{|R^1 \cap R^2|}{|R^2|} = \frac{3}{3} = 1$ келиб чиқади. q_2 сўров q_1 сўровга семантик боғланган ҳисобланади ва ўхшашлиги 1 (q_1 айнан ўхшаш) деб ҳисобланади.

3-қоида. Агар (3.24) ифодада $i = j$ га тенг бўлса, $\langle q_i | q_j \rangle = 1$ бўлади ва тўлиқ семантик боғланган ҳисобланади, ўхшашлиги 1 (айнан ўхшаш) деб ҳисобланади. $f(q)$ нинг диагонал элементлари ҳар доим 1 га тенг бўлади.

4-қоида. Агар $0 \leq \langle q_i | q_j \rangle \leq 1$ бўлса, q_i , q_j га семантик боғланган ҳисобланади ва ўхшашлиги $\frac{|R^j \cap R^i|}{|R^j|}$ ифода билан ҳисобланади.

(2.13) мантикий семантик излаш усулини барча сўровлар учун қуйидагича ёзиб оламиз:

$$f(q_1, q_2, \dots, q_{N_q}) = \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} \langle q_i | q_j \rangle = \sum_{i=1}^{N_q} \sum_{j=1}^{N_q} \frac{|R^j \cap R^i|}{|R^j|} \quad (2.14)$$

Бунда: \sum – тўплам элементлари бирлашмаси. (2.14) ифода МИТда семантик билимлар базасини яратиш имкониятини беради.

(2.13) усулнинг сонли ва графли тасвирини кўриш учун фараз қиламиз, $q = \{q_1, q_2, q_3, q_4, q_5\}$ сўровлар тўпламига мос $\{\{1,3,5,7,9\}, \{2,4,6,8\}, \{1,3,2,4,8\}, \{1,3,5,7,9,2\}\}$ маълумотлар тўплами берилган бўлсин. Бу рақамлар мос равишда маълумотларнинг ресурсдаги уникал номери.

(2.13) усул бўйича ҳисоблашни амалга оширамиз.

$$\frac{|R^1 \cap R^1|}{|R^1|} = \frac{5}{5} = 1, \quad \frac{|R^1 \cap R^2|}{|R^1|} = \frac{\emptyset}{5} = 0, \quad \frac{|R^1 \cap R^3|}{|R^1|} = \frac{2}{5} = 0.4, \quad \frac{|R^1 \cap R^4|}{|R^1|} = \frac{5}{5} = 1$$

$$\frac{|R^2 \cap R^1|}{|R^2|} = \frac{\emptyset}{4} = 0,$$

$$\frac{|R^2 \cap R^2|}{|R^2|} = \frac{4}{4} = 1, \quad \frac{|R^2 \cap R^3|}{|R^2|} = \frac{2}{4} = 0.5, \quad \frac{|R^2 \cap R^4|}{|R^2|} = \frac{1}{4} = 0.25$$

$$\frac{|R^3 \cap R^1|}{|R^3|} = \frac{2}{5} = 0.4,$$

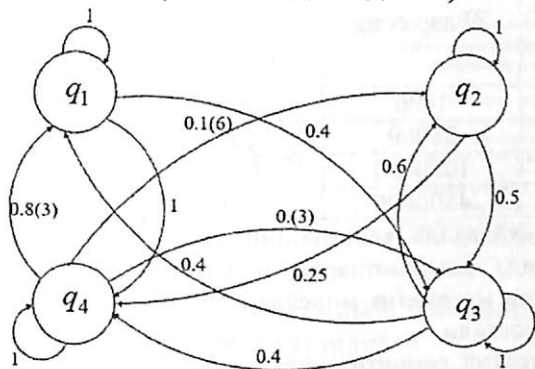
$$\frac{|R^3 \cap R^2|}{|R^3|} = \frac{3}{5} = 0.6, \quad \frac{|R^3 \cap R^3|}{|R^3|} = \frac{5}{5} = 1, \quad \frac{|R^3 \cap R^4|}{|R^3|} = \frac{2}{5} = 0.4$$

$$\frac{|R^4 \cap R^1|}{|R^4|} = \frac{5}{6} = 0.8(3), \quad \frac{|R^4 \cap R^2|}{|R^4|} = \frac{1}{6} = 0.1(6), \quad \frac{|R^4 \cap R^3|}{|R^4|} = \frac{2}{6} = 0.(3),$$

$$\frac{|R^4 \cap R^4|}{|R^4|} = \frac{6}{6} = 1$$

Натижада $q = \{q_1, q_2, q_3, q_4, q_5\}$ сўровларнинг мантикий семантик боғланиши (2.2-расм) ва ўхшашлигининг сонли кўриниши (2.15) да келтирилган.

$$f(q) = \begin{pmatrix} 1 & 0 & 0.4 & 1 \\ 0 & 1 & 0.5 & 0.25 \\ 0.4 & 0.6 & 1 & 0.4 \\ 0.8(3) & 0.1(6) & 0.(3) & 1 \end{pmatrix} \quad (2.15)$$



2.2-расм. Мантикий семантик боғланишнинг граф кўриниши

Умумий ҳолда мантикий семантик излаш моделида ўзаро семантик боғланиш ва ўхшашликни гуруҳлаштириш асосида фойдаланиш мумкин. Агар фойдаланувчи томонидан шакллантирилган q сўровни ахборот тизимидаги сўровлар орқали шакллантириш мумкин бўлса, яъни:

$$q_{new} = \bigcup q_{old}$$

кўринишида, шунда маълумотларни излаш тезлиги сезиларли ўзгаради. Чунки тизимдаги сўровлар учун мос маълумотларнинг мантикий семантик боғланиши ва ўхшашлиги ҳақидаги маълумотлар (2.13) ва (2.15) ифодалар орқали автостатик аниқлаштириб қўйилади.

Юқорида келтирилган (2.13) ва (2.15) мантикий семантик излаш усули ва анъанавий излаш ёрдамида олинган солиштирма натижаларни жадвал кўринишида келтирамиз (2.10-жадвал).

**Мантикий семантик излаш модели ёрдамида олинган
натижалар**

Тажриба рақами	ЭРлар сони	Анъанавий излаш тезлиги	Мантикий семантик излаш тезлиги	Нисбати
1	1000	5 с	0.5 с	10
2	25000	23 с	1 с	23
3	1000000	125 с	3 с	41.6
4	4500000	250 с	3.5 с	71.4

2.10-жадвалда келтирилган тажриба маълумотлари асосида КАКТларида маълумотларни излашни тезлаштириш учун таклиф қилинаётган математик моделдан фойдаланиш сезиларли даражада таъсир кўрсатади.

Мантикий семантик излаш моделининг афзаллик томонлари куйидагилардан иборат:

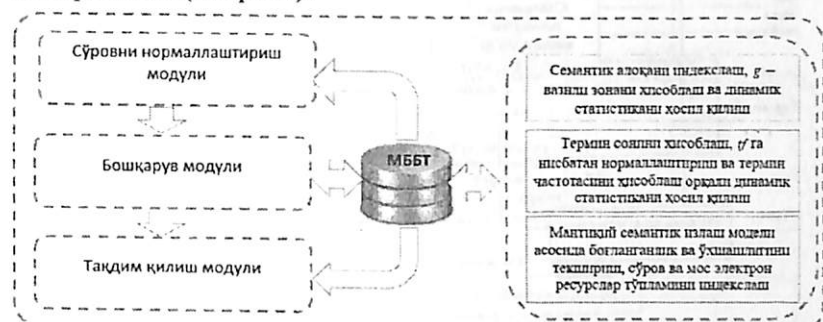
1. Сўровга мос маълумотлар рўйхатини тез шакллантириш.
2. Янги сўровни мавжуд сўровлар билан алмаштириш.
3. Сўровларнинг ўзаро мантикий боғланиши ҳақидаги маълумотларни сақлаш.
4. Сўровларнинг ўзаро мантикий ўхшашлиги ҳақидаги маълумотларни сақлаш.

Шундай қилиб, бугунги кунда яратилажак барча АТлари, қолаверса, КАКТларда, архивларда, ташкилотларнинг ЭР тизимларида мантикий семантик излаш моделидан фойдаланиш мақсадга мувофиқ.

2.5. Стохастик ахборот мухитларида маълумотларни излаш тизимининг дастурий модулини яратиш

Жорий бобнинг одинги параграфларда САМларда маълумотларни излаш ва қайта ишлаш учун маълумотларни излаш тизимларининг асосий элементлари ва моделлаштириш омиллари, ЭР частотаси ва векторли моделга асосланган долзарбликни ҳисоблаш усуллари тадқиқ қилинган ҳолда ишлаб чиқилди. Ушбу параграфда шуларга асосланган ҳолда САМда маълумотларни излаш тизимининг инструментал дастурий модулини лойиҳалаштириш, IDEF технологиялари асосида моделларини ишлаб чиқамиз.

САМда маълумотларни излаш тизимининг умумий функционал имкониятини куйидаги функционал тузилма билан тасвирлаймиз (2.3-расм):



2.3-расм. САМда маълумотларни излаш тизимининг умумий функционал тузилмаси

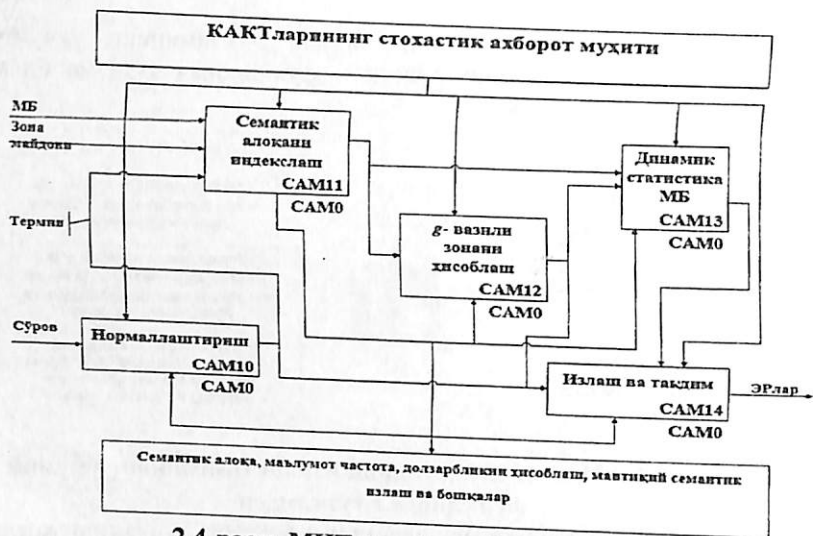
САМларда сўровга мос электрон ресурслар тўпламини ҳосил қилиш учун 3 та мустақил тоифали инструментал воситаларни функционал имкониятлари тасвирланган IDEF0 моделларини ишлаб чиқиш лозим. Булар:

1-тоифа. Семантик алоқани индекслаш, g – вазли зонани ҳисоблаш ва динамик статистика ҳосил қилиш (3.4-расм).

2-тоифа. Термин сонини ҳисоблаш, t^f га нисбатан нормаллаштириш ва термин частотасини ҳисоблаш орқали динамик статистика ҳосил қилиш (3.5-расм).

3-тоифа. Мантиқий семантик излаш модели асосида боғланганлик ва ўхшашлигини текшириш, сўров ва мос ЭРлар тўпламини индекслаш (3.6-расм).

Ушбу тоифалар САМда S сўровларни излаш учун статистик маълумотларни ҳосил қилиш усул ва алгоритмларига асосланади.



2.4-расм. МИТнинг 1-тоифа IDEF0 модели

САМда маълумотларни ишлаш тизимининг 1-тоифа IDEF0 моделига кирувчи параметрлар «сўров», «МБ», «Зона майдони» ва чикувчи параметрлар «ЭРлар» бўлиб ҳисобланади.

«Зона майдони» учун маълумот 2 хил махсус кенгайтирилган, яъни аниқ кўринишда (2.4-расм) ва олдиндан аниқланган қолиплар асосида шакллантирилиши мумкин.

Берилган сўров нормаллаштирилгандан сўнг терминларга ажратилади. Терминлар асосида зона майдонлари аниқланади ва қолиплар билан мослиги аниқлангандан сўнг «Зона майдони» маълумоти сифатида олинади. МБсидаги ЭРларнинг мос «Зона майдони» маълумоти асосида терминлар билан ЭРлар орасида семантик алоқа мавжудлиги аниқланиб, индексланади. ЭРга мос зона майдонларининг g қийматлари ҳисобланади ва динамик статистика МБсига езилади. Динамик статистика МБда термин ёки ЭР қўшилганда семантик алоқани янгилаш ва g қийматни қайта аниқлаш лозим. Биринчи тоифа берилган сўровга мос ЭРнинг долзарблигини зонали майдонлар асосида қайта ишлаш орқали ҳисоблаш учун хизмат қилади.



2.5-расм. SAMда маълумотларни излаш тизимининг 2-тоифа IDEF0 модели

SAMда маълумотларни излаш тизимининг 2-тоифа IDEF0 моделига кирувчи параметрлар «сўров», «МБ» ва чикувчи параметрлар «ЭРлар» бўлиб ҳисобланади.

2-тоифа IDEF0 моделида асосий ёндашув сўровда зона майдонларисиз ёки зона қолипларига мос келмаган ҳолларда самарали ҳисобланади. Бу сўров терминлар кетма-кетлиги асосида берилган бўлиб, нормаллаштиришдан сўнг фақат терминлар тўплами ҳосил қилинади. Ушбу терминларга ЭРларнинг долзарблигини термин частотани ҳисоблаш билан белгиланади. Бунинг учун ЭРдаги терминлар сонини ҳисоблаш ва статистик маълумотларни сақлаш кетма-кет амалга оширилади. Терминга мос ЭРларни тақдим қилишда термин частотаси ҳисобланади ва камайиш тартибида сараланади. Термин частотаси ўзгарувчан ҳисобланади, шунинг учун у ҳар доим ҳисобланади.

Биринчи ва иккинчи тоифадаги динамик маълумотларга эга ҳисоблаш SAMнинг маълумотларни излаш тизимларида маълумотларни излаш тезлигига ўз таъсирини кўрсатади. Бунинг учун инструментал дастурий модулини, яъни мантиқий семантик излаш модели асосида яратилган SAMда маълумотларни излаш тизимининг 3-тоифа IDEF0 модели, боғланганлик ва ўхшашлигини текшириш, сўровни ва мос ЭРлар тўпламини индекслашдан фойдаланиш мумкин.



2.6-расм. САМда маълумотларни излаш тизимининг 3-тоифа IDEF0 модели.

3-тоифада кирувчи параметрлар «сўров», «МБ» ва чиқувчи параметрлар «Электрон ресурслар» бўлиб ҳисобланади.

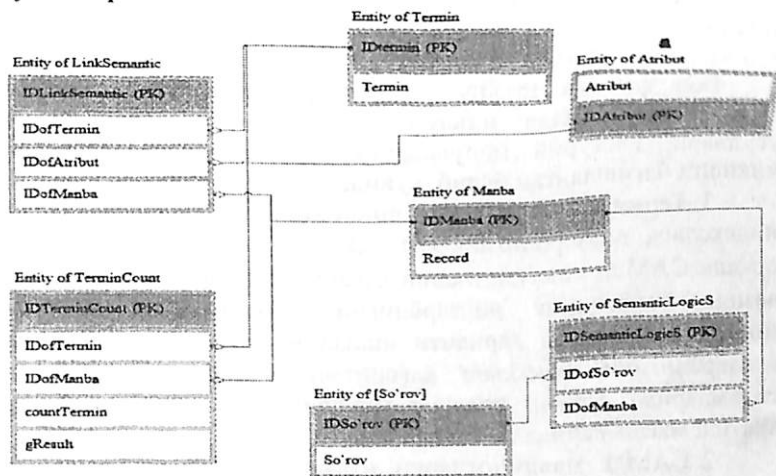
Сўров нормаллаштирилгандан сўнг, индексланади. Агар сўров аввалги сўровларнинг комбинацияларидан шакллантириш мумкин бўлса, сўровларнинг боғланганлиги аниқланади ва мос ЭРларни индекслаб, тақдим қилади. Агар сўров аввалги сўровларнинг комбинацияларидан шакллантириш мумкин бўлмаса, биринчи ва иккинчи босқичда инструментал дастурий модулларга муурожаат қилинади ва улар асосида ЭРлар тўплами ҳосил қилинади, сўров учун индекслангандан сўнг тақдим қилинади. Учинчи тоифа аввалдан амалга оширилган амаллар ҳисобига (ўқув мисоллари) бир неча марта МИТда маълумотларни излаш ва қайта ишлаш жараёнини тезлаштиради.

САМда маълумотларни излаш тизими учун ишлаб чиқилган 3 тоифали IDEF0 моделлари инструментал дастурий модулни яратиш имкониятини беради. Бу воситаларни ихтиёрий дастурий комплексга интеграция, мослаштириш, ўзаро биргаликда ишлашини таъминлаш учун улар билан бирга, яъни IDEF0 моделлар билан бирга IDEF1x модель ҳам ишлаб чиқилиши ва интеграция қилиниши керак. Чунки IDEF0 моделларини турли дастурий таъминотларга интеграция қилиш учун IDEF1x модели асосида ишлаб чиқилган маълумотлар тузилмалари ва уларнинг ўзаро моҳият-муносабат алоҳида муҳим ўрин тутади.

IDEF0 моделлар учун IDEF1x модели статик ва динамик маълумотларни сақлаш учун мўлжалланган реляцион маълумотлар

тузилмасига эга бўлган реляцион МБсининг моделидир. САМда маълумотларни излаш тизимлари учун IDEF1x моделда 7 та объект ва унинг хусусиятлари тасвирланган бўлиб, улар қуйидагилар(2.7-расм):

1. «Entity of Termin» – терминлар тўплами учун яратилган ва 2 та «IDTermin» ва «Termin» хусусиятларига эга.
2. «Entity of Atribut» – зонали майдонлар тўплами учун яратилган ва 2 та «IDAtribut» ва «Atribut» хусусиятларига эга.
3. «Entity of Manba» – ЭРлар тўплами учун яратилган ва 2 та «ID Manba» ва «Manba» хусусиятларига эга.
4. «Entity of SemanticLogicS» – сўровга мос ЭРлар тўпламини йиғиш учун яратилган ва 3 та «IDSemanticLogicS», «IDofSo'rov» ва «IDofManba» хусусиятларига эга.
5. «Entity of [So'rov]» сўровлар тўплами учун яратилган ва 2 та «IDSo'rov» ва «So'rov» хусусиятларига эга.
6. «Entity of TerminCount» ЭРда термин сони ва мос қийматни сақлаш учун яратилган ва 5 та «IDTerminCount», «IDofTermin», «IDofManba», «countTermin» ва «gResult» хусусиятларига эга.
7. «Entity of SemanticLink» – термин, зона майдони ва ЭРларни семантик алоқасини йиғиш учун яратилган ва 4 та «IDLinkSemantic», «IDofTermin», «IDofManba» ва «IDofAtribut» хусусиятларига эга.



2.7-расм. САМда маълумотларни излаш тизимининг IDEF1x модели

IDEF1x моделдаги объектлар орасида фақат бирга кўп (1:N) моҳият-муносабатли боғланишлар мавжуд (2.11-жадвал).

2.11-жадвал

IDEF0 моделлар учун IDEF1x моделидаги боғланишлар

№	ПК майдон номи	FK майдон номи	Боғланиши
1-тоифа IDEF0 модели учун			
1	IDTermin	IDofTermin	(1:N)
2	IDManba	IDofManba	(1:N)
3	IDAtribut	IDofAtribut	(1:N)
2-тоифа IDEF0 модели учун			
1	IDTermin	IDofTermin	(1:N)
2	IDManba	IDofManba	(1:N)
2-тоифа IDEF0 модели учун			
1	IDSo'rov	IDofSo'rov	(1:N)
2	IDManba	IDofManba	(1:N)

CAMда маълумотларни излаш тизими учун яратилган IDEF0 ва IDEF1x моделларини интеграция қилиш учун дастурий комплексга MVC технологияси асосида Code first тамойилига асосланган ёки моделларни қайта ишлаш имкониятига эга бўлиши лозим. Бу IDEF0 ва IDEF1x моделларни қайта ишлаб, уларнинг объектларни яратиш имкониятини беради. Яратилган объектлар дастурий комплексни объектлари билан мустақил ва ўзаро ахборотни бошқариш ҳамда алмашиш жараёнларида фойдаланилади.

Иккинчи боб бўйича хулосалар

Боб фан ва таълимга оид корпоратив ахборот-кутубхона тизимининг CAMда маълумотларни излаш ва қайта ишлаш усуллари, дастурий модуллари IDEF моделларини ишлаб чиқишга бағишланган бўлиб, қуйидаги хулосаларга олиб келади:

1. Термин ва зона майдонининг семантик алоқасини аниқлаш, индекслаш, долзарблигини ҳисоблашнинг махсус зонали хусусияти орқали CAMда маълумотларни излаш тизимлари учун терминнинг вазнли зонасининг долзарблигини ҳисоблаш алгоритмининг модификацияланган варианты ишлаб чиқилган. Вазнли зонанинг долзарблигини ҳисоблаш алгоритмлари маълумотларни излаш тизимларида термин ва зона майдонининг семантик алоқасини яратиш масалаларида фойдаланилади.

2. CAMда маълумотларни излаш тизимлари учун электрон ресурс ва термин частотасининг вазнини ҳисоблаш усуллари модификацияланган вариантлари ишлаб чиқилди. Электрон ресурс-

ларни терминлар асосида синфлаштириш, ресурсларни таҳлил қилиш масалаларини ечишда хизмат қилади.

3. Маълумотларни излаш тизимларини баҳолаш учун аниқлик ва тўлалик мезонлари, сарфланадиган вақтни камайтириш учун мантиқий семантик излаш усули ишлаб чиқилган. Мезонлар ва усул асосида кўп тармоқли ва босқичли ахборот-ресурсларида маълумотларни излаш ва қайта ишлаш учун сўров ва олдиндан аниқланган ЭР тўпламларининг индекслаш, уларнинг семантик алоқаларни яратиш, излаш тезлигини оптималлаштириш масалаларини тадқиқ қилишга хизмат қилади.

4. САМда маълумотларни излаш тизимлари учун инструментал дастурий модулнинг функционал имкониятининг тузилмаси ва лойиҳалаштириш учун 3 та IDEF0 моделлари ва реляцион маълумотлар тузилмасининг IDEF1х модели ишлаб чиқилган. IDEF моделлари асосида мураккаб тузилмали ахборот муҳитлари учун МИҚИнинг дастурий модуллари функционал тузилмасини ишлаш чиқиш ва маълумотлар базасини лойиҳалаштириш масалаларига оид тадқиқотларни амалга оширишда фойдаланилади.

III-боб. НОРАВШАН АХБОРОТ МУҲИТЛАРИДА МАЪЛУМОТЛАРНИ ИЗЛАШ ВА ҚАЙТА ИШЛАШ МОДЕЛЛАРИ, УСУЛЛАРИ ВА ДАСТУРИЙ МОДУЛИНИ ЯРАТИШ

3.1. Норовшан сўровларни қайта ишлаш усули

Компьютер техникаси ёрдамида объектларнинг нафақат сонли хусусиятларини ҳисоблаш, балки сонли характерга эга бўлмаган хусусиятлари билан ҳам ишлашга тўғри келмоқда. Бундай масалалардан гуманитар соҳаларда, фан ва таълимга оид КАКТларида маълумотларни излаш ҳамда қайта ишлашда фойдаланиш мумкин. Объектларнинг сонли характерга эга бўлмаган хусусиятлари билан ишлашда турли ёндашувлар, назариялар мавжуд [43, 100]. Тадқиқотимизда норовшан тўпламлар назарияси асосида МИҚИ усулларини тадқиқ қиламиз. Маълумки, КАКТда маълумотларни излашда асосан фойдаланувчи томонидан тузиладиган сўровлардек фойдаланилади. Бу сўров мазмун жиҳатидан сонли характерга эга бўлмаган лингвистик термлардан ташкил топади («Сўнгги янгликлар», «Қадимий асарлар», «Замонавий кино», «Янги Тошкент» ва бошқалар).

КАКТда турли маълумотлар аниқ характерли кўринишда МБларида сақланади ва қайта ишланади. Уларни қайта ишлаш ЭРларни излашда учун муҳим роль ўйнайди. Чунки қайта ишланиши лозим бўлгани ЭРни аввал МБдан излаб топиш лозим. Бунинг учун КАКТ фойдаланувчилари томонидан реляцион МБ (РМБ) учун ёзиладиган норовшан сўровларнинг қайта ишлаш усулини ишлаб чиқиш лозим.

РМБларида маълумотлар қисман аниқ ва сонли характерга эга бўлиб, зарур ЭРларни замонавий дастурий таъминотлар томонидан тузиладиган сўровларнинг модели, усули ва алгоритми ёрдамида тез, излаб топиш мумкин. Агар РМБда маълумотларни излаш учун фойдаланувчилар сўровларидан фойдаланиш лозим бўлсачи? Фойдаланувчининг сўрови аниқ ва сонли характерда бўлса, уни аниқ сўров ёки S сўровларга киритамиз, буларни РМБ учун лойиҳалаштириш муаммо ҳисобланмайди («XIX асрга тааллуқли қўлёзма манбалар»). Аммо, фойдаланувчи сўровида сифатий, сонли характер маънодаги сўзлардан фойдаланган бўлса, яъни норовшан терм, бунда сўровни норовшан сўровларга киритамиз, РМБ учун лойиҳалаштириш норовшан тўпламлар назариясига асосланади («XIX аср бошларига тааллуқли қўлёзма манбалар»).

Юкоридаги норавшан сўровлардан хар куни фойдаланилади. Шунинг учун ЭРлар тўпламида норавшан сўровларни лойихалаштириш ва қайта ишлаш масаласини кўриб чиқамиз.

Фараз қиламиз, фан ва таълимга оид корпоратив ахборот-кутубхона тизимининг РМБсида аниқ, сонли характерга эга маълумот тузилмаси берилган бўлсин (3.1-жадвал).

3.1-жадвал

«1-жадвал» номли РМБ жадвали

т.р.	ЭР номи	Аср	Йил	Баҳоси (бирликда)
1	Тарихий асар 1	17	30	1000
2	Тарихий асар 2	17	80	2500
3	Тарихий асар 3	17	50	10000
4	Тарихий асар 4	18	12	2000
5	Тарихий асар 5	17	10	2000
6	Тарихий асар 6	18	90	3000
7	Тарихий асар 7	18	25	100
8	Тарихий асар 8	18	55	9000
9	Тарихий асар 9	18	44	2500
10	Тарихий асар 10	19	10	500

КАКТда фойдаланувчилар томонидан «XVIII асрга тааллуқли баҳоси 3000 дан ортмаган асар» сўров ёзилган бўлсин. Бу аниқ сўров бўлиб, 3.1-жадвалдан маълумотларни танлаш усули асосида амалга ошириш имкони бор. РМБ учун ушбу сўровни SQL стандарти асосида қуйидагича лойихалаштириш мумкин:

«*SELECT`Тарихий асар`FROM`Жадвал1`WHERE`аср`=18 AND`Баҳоси`< 3000*»

КАКТда фойдаланувчилар томонидан қуйидагича «XVIII асрга тааллуқли, баҳоси қиммат асар» сўров ёзилган бўлсин. Бу норавшан сўров бўлиб, 3.1-жадвалдан тўғридан-тўғри маълумотларни танлаш имкони йўқ. Бунинг учун норавшан тўпламлар назариясидан фойдаланиб, қуйидаги босқичларни амалга оширамиз:

1) β – лингвистик ўзгарувчини аниқлаймиз (сўровда `Баҳоси` майдони);

2) $T(\beta)$ – лингвистик ўзгарувчига мос норавшан термлар тўпламини аниқлаймиз (сўров учун $T(\beta) = \{`арзон`,`ўртача`,`қиммат`}\}$);

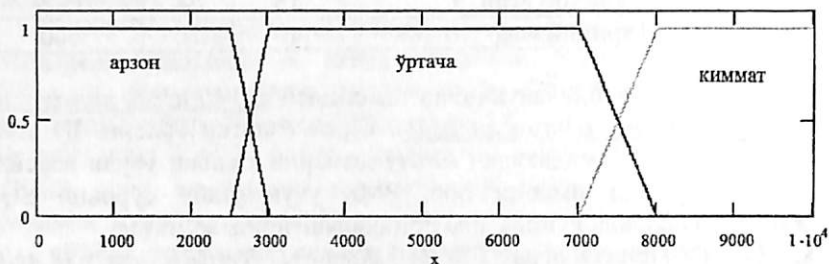
3) X – норавшан термлар тўплами учун кийматлар тўпламини аниқлаймиз (сўров учун $X = \{\min(\text{'Баҳоси' }), \max(\text{'Баҳоси' })\}$);

4) $G(\beta)$ – мантикий ва модификатор амаллари орқали таърифланадиган янги норавшан термлар тўпламини аниқлаймиз;

5) μ_x – X тўпламининг тегишлилик функциясини аниқлаймиз. Сўров учун A – 'арзон', B – 'ўртача', C – 'қиммат' ва $G(\beta)$ – учун норавшан мантиқ амалларидан фойдаланилади. $T(\beta)$ га мос тегишлилик функцияларини трапециясимон кўринишини танлаб, қуйидагича аниқлаб оламиз.

$$\mu_x(A) = \begin{cases} 1, & x < 2500 \\ \frac{3000-x}{500}, & x \in (2500, 3000] \\ 0, & x > 3000 \end{cases}, \quad \mu_x(B) = \begin{cases} 0, & x < 2500 \text{ ва } x > 8000 \\ \frac{x-2500}{500}, & x \in [2500, 3000] \\ 1, & x \in (3000, 7000) \\ \frac{8000-x}{1000}, & x \in [7000, 8000] \end{cases},$$

$$\mu_x(C) = \begin{cases} 0, & x \leq 7000 \\ \frac{x-7000}{1000}, & x \in (7000, 8000] \\ 1, & x > 8000 \end{cases}$$



3.1-расм. Тегишлилик функцияларининг трапециясимон график кўриниши

$G(\beta)$ – мантикий ва модификатор амаллари орқали таърифланадиган янги норавшан термлар тўпламини аниқлаймиз:

$A \wedge B$ яъни, $\mu_{A \wedge B} = \min(\mu_A, \mu_B)$, $A \vee B$ яъни, $\mu_{A \vee B} = \max(\mu_A, \mu_B)$, A эмас, яъни, $1 - \mu_x(A)$, B эмас, яъни, $1 - \mu_x(B)$, C эмас, яъни, $1 - \mu_x(C)$, жуда A , яъни, $\sqrt{\mu_x(A)}$, озгина A , яъни, $(\mu_x(A))^2$, жуда B , яъни, $\sqrt{\mu_x(B)}$, озгина B , яъни, $(\mu_x(B))^2$, жуда C , яъни, $\sqrt{\mu_x(C)}$, озгина C , яъни, $(\mu_x(C))^2$.

Юкоридагилардан фойдаланиб, ихтиёрий лингвистик ўзгарувчи учун (2.1) кўринишидаги математик модель ҳосил қилинади [94].

$$f = F(\beta, T(\beta), X, G(\beta), \mu_x) \quad (3.1)$$

(3.1) ёрдамида сўров лойиҳалаштирилади ва РМБсида SQL стандарти асосида қуйидагича лойиҳалаштирилади:

«SELECT `Тарихий асар` FROM `Жадвал1` WHERE `аср`=18 AND
`Баҳоси`=`қиммат`»

Фараз қиламиз, сўров сифатида «Қиммат асарлар» берилган бўлсин. (3.1) модель асосида сўров лойиҳалаштирилади, КАКТнинг РМБсидан қуйидаги маълумотларни танлаш мумкин (3.2-жадвал).

3.2-жадвал

Сўров натижаси

№	Тарихий асар
1	Тарихий асар 8
2	Тарихий асар 3

Битта РМБ учун бир неча лингвистик ўзгарувчилардан фойдаланилса, норавшан сўровларга мос маълумотларни излаш аниқлиги ошади [75].

Фараз қиламиз, фойдаланувчи томонидан «Яқин аср бошларидаги қиммат асар» кўринишидаги норавшан сўров ёзилган бўлсин. Бунда учта норавшан сўзлар ишлатилган. Шунинг учун РМБсига `аср` ва `йил` майдонлари учун лингвистик ўзгарувчилар киритилади.

`аср` учун:

1) β - `аср`;

2) $T(\beta) = \{`узок`, `тош`, `мис`, `ўрта`, `илм-фан`, `яқин`, `ахборот`\}$;

3) $X = \{x_i, x_i = i, i = 1 \dots 21\}$;

4) μ_x - X тўпламнинг тегишлилик функцияси. Мос равишда A, B, C, D, E, F, J ;

`йил` учун:

1) β - `йил`;

2) $T(\beta) = \{`bir`=>\{`бошларида`, `ўрталарида`, `охирларида`\}, `ikki`=>`а-ярмида`, `uch`=>`b-10 йилликда`, `to`rt`=>`с - чорак`\}$; бунда a ($a = \{0, 1, 2\}$), b ($b = \{i, i = 0 \dots 10\}$), c ($c = \{1, 2, 3, 4\}$) бутун типли параметрлар.

3) $X = \{x_i, x_i = i, i = 1 \dots 99\}$;

4) μ_x - X тўпламининг тегишлилик функцияси. Мос равишда $bir_A, bir_B, bir_C, ikki_a, uch_b, to'rt_c$; бунда $a (a=\{0, 1, 2\})$, $b (b=\{i\}, i=0...10)$, $c (c=\{1, 2, 3, 4\})$ бутун типли параметрлар.

(3.1) модель асосида лойиҳалаш бажарилгандан сўнг, сўров учун SQL стандарти асосида қуйидагича лойиҳалаштириш мумкин:
 «SELECT `Тарихий асар` FROM `Жадвал1` WHERE `аср` = `яқин`
 AND `йил` = `бошлари` AND `Баҳоси` = `қиммат`»

Ушбу сўров РМБсидаги 3.1-жадвалдан қуйидаги маълумотни чиқариб беради (3.3-жадвал).

3.3-жадвал

Сўров натижаси

№	Тарихий асар
1	Тарихий асар 10

Юқорида келтирилган (3.1) модель ёрдамида фан ва таълимга оид КАКТлардаги тарихий ЭРларнинг асрлар, йиллар ва баҳо (кий-мат)ларига боғлиқ бўлган сўровларни лойиҳалаштириш ҳамда танланган термларнинг норавшан тўпламини кенгайтириш фойдаланувчилар томонидан ёзиладиган ихтиёрий сўровларни лойиҳалаштириш имконини беради.

Ихтиёрий КАКТларнинг маълумотлари РМБсида жойлашганлигидан қуйидаги ғояларни илгари суриш мумкин:

1. Мавжуд РМБсида маълумотлар излашни интеллектуал амалга оширишда (3.1) модел ёрдамида лингвистик ўзгарувчи киритиш лозим.

2. Ушбу моделни фақат аниқ ва сонли характерга эга бўлган РМБ учун эмас, балки ихтиёрий РМБда фойдаланиш мумкин.

3. Ушбу моделни турли соҳадаги РМБда қўллаш учун махсус конструкторли дастурий таъминот яратиш мақсадга мувофиқ, деб ҳисоблаймиз.

4. Конструкторли дастурий модул яратиш универсал усул ва алгоритм, лингвистик ўзгарувчилар учун параметрик тегишлилик функцияларини лойиҳалаштириш ҳам талаб қилинади.

3.2. Параметрли тегишлилик функцияснни лойиҳалаш усули

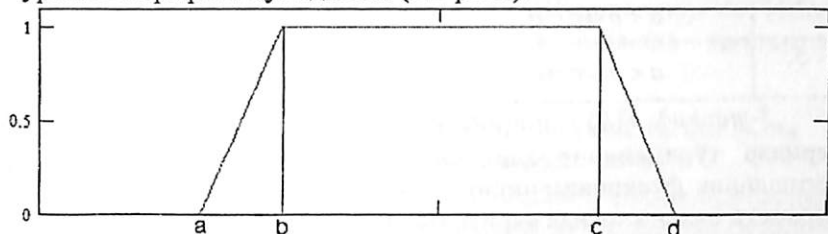
Юқорида келтирилган сўровларни қайта ишлаш усулида мос ЭРларни танлаш учун Л.Заде томонидан киритилган норавшан тўпламлар назариясидаги “норавшан ўзгарувчи” ва “лингвистик ўзгарувчи” тушунчалари [42; 4-6 -б.] ва (3.1) моделидан фойдаланилган.

(3.1) модель асосида МИҚИ муаммосини ўрганганимизда объектларнинг шундай хос хусусиятлари мавжуд бўлар эканки, улар динамик равишда маълумотлар кўпайишига (камайишига) қараб ўзгариб боравераркан (агар объектимизда “йил” хусусияти мавжуд бўлса, яратилган дастури модул учун ҳаёт даврининг ошиши натижасида унинг хусусияти динамик ўзгариб боришини кўриш мумкин).

Шунинг учун (3.1) модель μ_x - X тўпламининг β – лингвистик ўзгарувчиларга мос объект хос хусусиятини тегишлилик функцияси параметрли лойihalаштиришни амалга ошириш лозим. μ_x - функцияга 4 та параметрни киритамиз[121].

$$\mu_x(a, b, c, d) = \begin{cases} 0, & \text{агар } (x < a \ \& \ y > d) \\ \frac{x-a}{b-a}, & \text{агар } (x \in [a, b] \ \& \ a \neq b) \\ 1, & \text{агар } (x \in (b, c)) \\ \frac{d-x}{d-c}, & \text{агар } (x \in [c, d] \ \& \ c \neq d) \end{cases} \quad (3.2)$$

(3.2) параметрли тегишлилик функциясининг умумий кўриниши графиги қуйидагича (3.2-расм):



3.2-расм. (3.2) – параметрли тегишлилик функцияси

(3.2) – параметрли тегишлилик функцияси учун қуйидаги таъкидларни киритамиз:

1-таъкид. $\mu_x(a, b, c, d)$ функциядаги a, b, c, d параметрлар X – қийматлар тўпламидан олинган бўлса, $\mu_x(a, b, c, d)$ функция X – қийматлар тўпламининг тегишлилик функцияси бўлади.

2-таъкид. $\mu_x(a, b, c, d)$ тегишлилик функциясидаги a, b, c, d параметрлар ўзаро қатъий ўсиш тартибида тартибланган бўлиши шарт. Яъни, $a < b$ ёки $b < c$ ёки $c < d$ шартларидан бири ҳар доим бажарилиши керак.

(3.2) параметрли тегишлилик функциясининг алгоритмини SQL стандарти асосида лойиҳалаштирилган процедураси куйидагича:

```
PROCEDURE CalcTFunc (@lingVar[] int, @a int, @b int, @c
int, @d int) AS BEGIN IF ((@lingVar<@a) and ((@lingVar>@d))
RETURN 0;
```

```
IF ((@lingVar IN @a between @b) and (@a <> @b) RETURN
((@lingVar-@a)/(@b-@a));
```

```
IF ((@lingVar IN @c between @d) and (@c <> @d))
RETURN ((@d - @lingVar) / (@d-@c));
```

```
IF ((@lingVar IN ONLY(@b between @c)) RETURN 1; END.
```

(3.2) параметрли тегишлилик функцияси параметрлари ўзаро муносабатларининг мос шартлари ҳам муҳим. Ўзаро муносабатлар учун тегишлилик функциясининг барча ҳолларини 3.4-жадвалда келтирамыз.

3.4-жадвал

(3.2) параметрли тегишлилик функцияси параметрларининг ўзаро муносабатлари

№	ларнинг ўзаро муносабатлари	№	Параметрларнинг ўзаро муносабатлари
1.	$\min(X) = b = c < d$	2	$a = b = c < d$
3.	$a < b = c < d$	4.	$a = b < c = d$
5.	$a < b = c = d$	6.	$a < b = c = \max(X)$

3-таъкид. $T(\beta)$ – лингвистик ўзгарувчиларнинг мос норавшан термлар тўпламининг ҳар бир терми учун (3.2) параметрли тегишлилик функцияларининг параметрлари камида битта фаркли параметр билан алоҳида киритилиши шарт.

Агар “ихтиро” объектининг “йил” хусусияти учун $T(\beta) = \{ \text{'бошлари'}, \text{'ўрталари'}, \text{'охирлари'}, \text{'сўнги'}, \text{'замонавий'} \}$ норавшан термлар келтирилган бўлса, уларнинг (3.2) параметрли тегишлилик функцияси параметрларининг қийматларини РМБда куйидагича шакллантириш мумкин:

Норавшан термлар параметрлари

т.р.	Норавшан терм	a	b	c	d
1.	Бошлари	1	1	30	35
2.	Ўрталари	30	35	65	70
3.	Охирлари	65	70	100	100
4	Сўнгги	Жорий вақт -5	Жорий вақт	Жорий вақт	Жорий вақт
5.	Замонавий	Жорий вақт -25	Жорий вақт - 20	Жорий вақт	Жорий вақт

3.5-жадвал асосида параметрли тегишлилик функцияларни лойиҳалаштириш учун параметрлар учун “динамик метрика” тушунчасини ҳам киритиш мумкин. Бу метрика маълумотлар сони орқали бошқарилади ва РМБга олдиндан киритилади.

Умуман олганда, фойдаланувчилар томонидан яратиладиган сўровларнинг кўпларида сонли характерга эга бўлмаган термларни учратиш мумкин. Бундай термларни қайта ишлашда юқорида келтирилган таърифларга асосланган параметрли тегишлилик функцияларидан фойдаланиш муҳим. Сонли характерга эга бўлмаган термларга мос тегишлилик функциялари ва уларнинг синфини танлаш муҳим аҳамиятга эга. Шунинг учун навбатдаги параграф маълумотларни интеллектуал излашда тегишлилик функцияларининг синфларига бағишланади.

3.3. Маълумотларни интеллектуал излашда тегишлилик функцияларининг синфларини танлаш

Ахборот тизимларининг ривожланиши ахборот инфраструктураси ва унинг дастурий модулларига ўз таъсирини кўрсатиб келмоқда. Хусусан, АТларидаги барча тизим хусусияти ва вазифаларини интеграция ёки модернизация қилиш ёки янгисини яратишни такозо қилмоқда. Шунингдек, АТлари жамият турли соҳаларида қўлланилмоқда, таклиф ва талаб қилинаётган МБларини, дастурий қулайликлар билан таъминлаш чора-тадбирлари кенгайиб бормоқда [116].

Инсониятни энг ишончли ва қимматли маълумотлар билан таъминлаш мақсадида, фан ва таълимга оид КАКТларда ЭРларини бошқариш ва улардан фойдаланиш жараёнларини автоматлаштириш масалалари билан бирга, бу тизимларда маълумотларни

интеллектуал излаш, ишлов бериш, тақдим этиш каби масалалар ўз ечимини кутмоқда.

Табиий ёки сунъий тилда киритилган сўровни каноатлантирувчи объектлар тўпланини аниқлаш учун сўровни ўзгармас матн эмас, балки мазмун жиҳатдан компьютерда аниқлаштириш лозим. Бунинг учун норавшан тўпламлар назариясида фойдаланилган “лингвистик ўзгарувчи” (ЛЎ) тушунчасидан фойдаланиш ва уларнинг норавшан термлар тўплами учун тегишлилик функцияларининг асосий синфларини тадқиқ қилиш масаласи пайдо бўлади.

Фан ва таълимга оид КАКТларининг норавшан ахборот муҳитларида маълумотларни интеллектуал излаш тизимини $\langle q, r \rangle$ жуфтлик билан белгилаймиз. Жуфтликда q – сўров, яъни фойдаланувчининг ахборотга бўлган эҳтиёжи, r – сўровнинг объектлари, яъни сўровга мос аниқланган ЭРлар тўплами.

Айнан, q сўровни каноатлантирувчи r сўров объектлар тўпланини аниқлаш учун норавшан тўпламлар назариясининг математик аппарати асосида амалга ошириш мумкин.

Фараз қилайлик, $X = \{x_1, x_2, \dots, x_N\}$ базавий тўплам асосида L ЛЎ учун $L = \{l_1, l_2, \dots, l_N\}$ термлар тўплами берилган бўлсин. У ҳолда T термлар тўпланининг мос равишда l_i элементи учун A_i норавшан тўпламни куйидагича аниқлаймиз:

$$A_i = \left\{ \langle x_1, \mu_{A_i}(x_1, l_i) \rangle, \langle x_2, \mu_{A_i}(x_2, l_i) \rangle, \dots, \langle x_n, \mu_{A_i}(x_n, l_i) \rangle \right\}, \quad i = 1..M \quad (3.3)$$

Агар x_i лар дискрет қийматлар бўлса, (3.3) ни куйидагича ёзиш мумкин:

$$A_i = \sum_{j=1}^N \frac{\mu_{A_i}(x_j, l_i)}{x_j}, \quad i = 1..M \quad (3.4)$$

Агар x_i лар узлуксиз қийматлар ($X \subseteq R$) бўлса, (3.4) ни куйидагича ёзиш мумкин:

$$A_i = \int \frac{\mu_{A_i}(x, l_i)}{x}, \quad i = 1..M \quad (3.5)$$

(3.4) ифодада \sum , (3.5) ифодада \int тўплам элементлари бирлашмаси. $\mu_{A_i}(x, l_i)$ эса мос равишда l_i лингвистик ўзгарувчи учун A_i норавшан тўпламнинг X – базавий тўпламга тегишлилик функцияси.

НАМларида маълумотларни интеллектуал излаш ва қайта ишлашда ЛЎлар учун тегишлилик функцияларини қандай танлаш

ёки қайси синфларнинг қайси норавшан терм учун фойдаланиш кераклигини аниқлашни кўриб чиқамиз.

Ҳар хил объектли масалаларни ечишда тегишлилик функцияларини турлича танлаш ёки киритиш мумкин. Аммо, шу синфларни маълумотларни интеллектуал излашда, яъни аниқ ЛЎ учун базавий тўплам ва норавшан тўплам термлар тўпламидан иборат бўлганда тегишлилик функцияларини асосий синфларидан қайси синфни қандай терм ёки ЛЎ учунлигини аниқлаш лозим. Бунинг учун куйидаги тегишлилик функцияларининг синфларидан фойдаланилади.

Чизикли тегишлилик функциялари, одатда, учбурчак ва трапецияли тегишлилик функцияларининг синфларидан ташкил топди. Кўп адабиётларда бу синф тегишлилик функциялари t (учбурчак) ва T (трапеция) синф тегишлилик функциялари деб юритилади [106; 145-б., 116; 73-б.]. Умумий ҳолда t синф тегишлилик функцияларини ЛЎдан фойдаланган ҳолда МИҚИДа куйидагича танлаб олиш мумкин (3.6):

$$\mu_x^t(x, l_x, a, b, c) = \begin{cases} 0, & \text{агар } x \leq a \text{ ёки } c \leq x \\ \frac{x-a}{b-a}, & \text{агар } a < x < b \\ \frac{c-x}{c-b}, & \text{агар } b \leq x < c \end{cases} \quad (3.6)$$

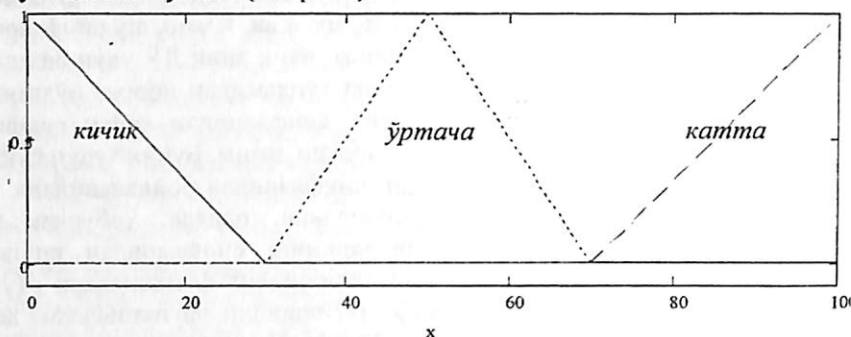
Бунда, X базавий тўплам, l_x – лингвистик ўзгарувчи учун D термлар тўплами аниқланади ва ҳар бир терм учун a, b, c сонли катталиклар киритилади ($a, b, c \in X = \{x_i\}, i = 1 \dots m$) ва 1) $a < b < c$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак (3.7-жадвал асосида). 1 – шартни чап чегарада $a \leq b < c$, ўнг чегарада $a < b \leq c$ шарт билан алмаштириш ҳам мумкин.

3.6-жадвал

Термларга a, b, c сонли катталикларни мослаштириш

X	l_x – Лингвистик ўзгарувчи			
	t_1	t_2	...	t_n
a	a_1	a_2	...	a_n
b	b_1	b_2	...	b_n
c	c_1	c_2	...	c_n

t синф $\mu_x^t(x, \text{Кичик}, 0, 0, 30)$, $\mu_x^t(x, \text{Ўртача}, 30, 50, 70)$, $\mu_x^t(x, \text{Катта}, 70, 100, 100)$ тегишлилик функцияларининг графиги куйидагича бўлади (3.3-расм):



3.3-расм. t синф тегишлилик функцияларининг графиги

T синф тегишлилик функцияларини лингвистик ўзгарувчидан фойдаланган ҳолда МИҚИда куйидагича танлаб олиш мумкин (3.7):

$$\mu_x^T(x, l_x, a, b, c, d) = \begin{cases} 0, & \text{агар } x \leq a \text{ ёки } d \leq x \\ \frac{x-a}{b-a}, & \text{агар } a < x \leq b \\ 1, & \text{агар } b < x \leq c \\ \frac{d-x}{d-c}, & \text{агар } c < x < d \end{cases} \quad (3.7)$$

Бунда, X базавий тўплам, l_x – ЛЎ учун D термлар тўплами аниқланди ва ҳар бир терм учун a, b, c, d сонли катталиклар киритилади ($a, b, c, d \in X = \{x_i\}$, $i=1..n$) ва 1) $a < b \leq c < d$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак (4.7-жадвал асосида). 1 – шартни чап чегарада $a \leq b \leq c < d$, ўнг чегарада $a < b \leq c \leq d$ шарт билан алмаштириш ҳам мумкин.

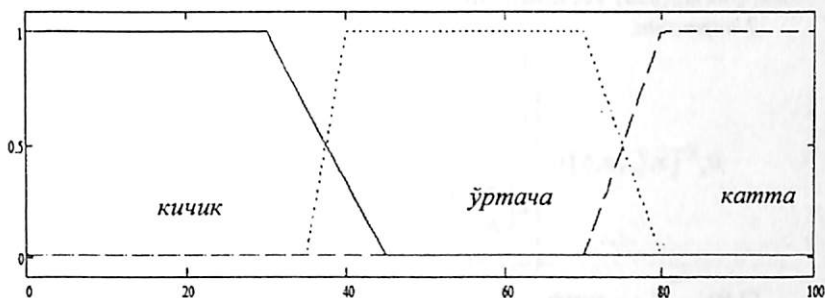
3.7-жадвал

Термларга a, b, c, d сонли катталикларни мослаштириш

X	l_x - Лингвистик ўзгарувчи			
	t_1	t_2	...	t_n
a	a_1	a_2	...	a_n
b	b_1	b_2	...	b_n
c	c_1	c_2	...	c_n
d	d_1	d_2	...	d_n

T синф $\mu_x^T(x, \text{Кичик}, 0, 0, 30, 45)$, $\mu_x^T(x, \text{Ўртача}, 35, 40, 70, 80)$, $\mu_x^T(x, \text{Катта}, 70, 80, 100, 100)$ тегишлилик функцияларининг a, b, c, d сонли параметрлар асосидаги графиклари 3.4-расмда тасвирланган.

Юқорида келтирилган t синф (3.6) ва T синф (3.7) тегишлилик функциялари маълумотларини интеллектуал излаш ва қайта ишлашда «тахминан тенг», «ўртача», «қиймат», «оралиқ», «жойлашган», «ўхшаш», «мумкин», «тегишлилик даражаси» каби характердаги норавшан термларнинг лингвистик ўзгарувчиси учун фойдаланиш тавсия қилинади.



3.4-расм. T синф тегишлилик функцияларининг графиклари

Z -тасвирли тегишлилик функциялари (Z синф) ҳам турлича берилиши мумкин [116; 74-б.]. Бу функцияларнинг тасвирлари “ Z ” ҳарфининг ёзилиши каби бўлиши керак. Одатда, бу синф функцияларини қуйидаги вариантларда танлаш мумкин:

1-вариант.

$$\mu_x^Z(x, l_x, a, b) = \begin{cases} 1, & \text{агар } x < a \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{x-a}{b-a} \pi\right), & \text{агар } a \leq x \leq b \\ 0, & \text{агар } x > b \end{cases} \quad (3.8)$$

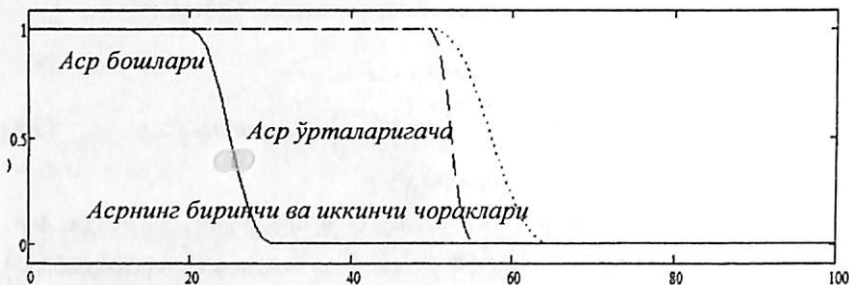
(3.8) Z синф $\mu_x^Z(x, \text{Аср бошлари}, 20, 30)$, $\mu_x^Z(x, \text{Аср ўрталаригача}, 50, 65)$, $\mu_x^Z(x, \text{Асрнинг биринчи ва иккинчи чораклари}, 50, 55)$ тегишлилик функциясининг графикли тасвири 3.5-расмда тасвирланган.



3.5-расм. (3.8) тегишлилик функцияларининг графиклари 2-вариант.

$$\mu_x^Z(x, l_x, a, b) = \begin{cases} 1, & \text{агар } x < a \\ 1 - 2\left(\frac{x-a}{b-a}\right)^2, & \text{агар } a < x \leq \frac{a+b}{2} \\ 2\left(\frac{b-x}{b-a}\right)^2, & \text{агар } \frac{a+b}{2} < x < b \\ 0, & \text{агар } x \geq b \end{cases} \quad (3.9)$$

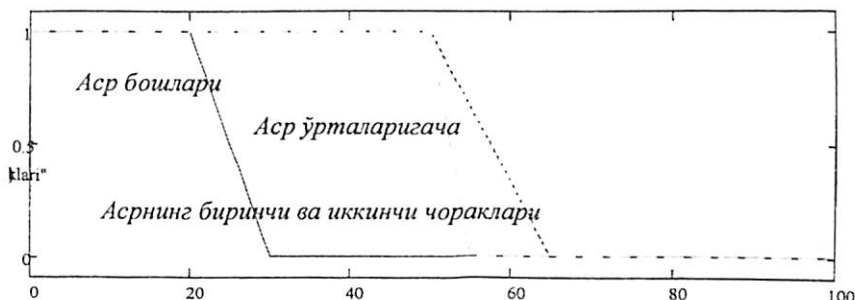
(3.9) Z синф $\mu_x^Z(x, \text{Аср бошлари}, 20, 30)$, $\mu_x^Z(x, \text{Аср ўрталаригача}, 50, 65)$, $\mu_x^Z(x, \text{Асрнинг биринчи ва иккинчи чорақлари}, 50, 55)$ тегишлилик функциясининг графикли тасвири 3.6-расмда тасвирланган.



3.6-расм. (3.9) тегишлилик функцияларининг графиклари 3-вариант. Z синф тегишлилик функциялари ҳам чизикли кўринишда ифодаланиши мумкин.

$$\mu_x^Z(x, l_x, a, b) = \begin{cases} 1, & \text{агар } x \leq a \\ \left(\frac{b-x}{b-a}\right), & \text{агар } a \leq x \leq b \\ 0, & \text{агар } x \geq b \end{cases} \quad (3.10)$$

(3.10) Z синф $\mu_x^Z(x, \text{Аср бошлари}, 20, 30)$, $\mu_x^Z(x, \text{Аср ўрталаригача}, 50, 65)$, $\mu_x^Z(x, \text{Асрнинг биринчи ва иккинчи чорақлари}, 50, 55)$ тегишлилик функциясининг графикли тасвири 3.7-расмда тасвирланган.



3.7-расм. (3.10) тегишлилик функцияларининг графиклари

Юқорида келтирилган Z синфига кирувчи (3.8), (3.9) ва (3.10) тегишлилик функцияларида X базавий тўплам, l_x – лингвистик ўзгарувчи учун D термлар тўплами аниқланди ва ҳар бир терм учун a, b сонли катталиқлар киритилади ($a, b \in X = \{x_i\}$, $i = 1 \dots n$) ва 1) $a < b$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак. a ва b лар ҳар бир термнинг аниқланганлик оралиғи бўлиб ҳисобланади.

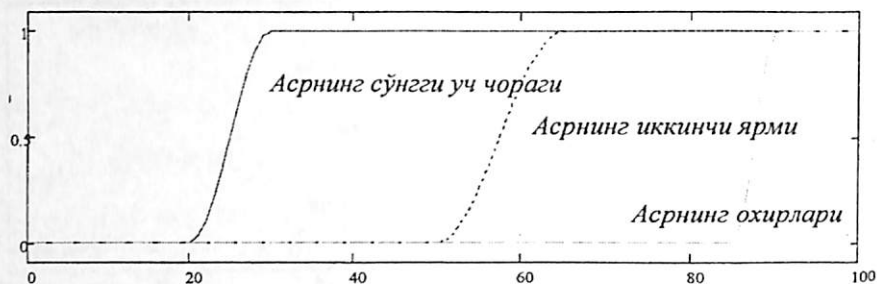
Z синфига кирувчи (3.8), (3.11) ва (3.10) функцияларни маълумотларни интеллектуал излаш ва қайта ишлашда «объект сони», «қиймат», «қиймат чегараси», «оралиқ катталиқ (қиймат)», «нарх», «мақсад», «тааллуқлилиқ», «тегишлилик даражаси оралиғи» каби характердаги норавшан термларнинг «кичик, ... гача» мазмунидаги лингвистик ўзгарувчисининг модификаторлари учун фойдаланиш тавсия қилинади.

S-тасвирли тегишлилик функциялари (S синф) ҳам турлича берилиши мумкин [116; 75-б.]. Бу функцияларнинг тасвирлари "S" ҳарфининг ёзилиши каби бўлиши керак. Одатда, бу синф функцияларини қуйидагича танлаш мумкин (Z синф тескари, $S \leftrightarrow 1 - Z$).

1-вариант.

$$\mu_x^S(x, l_x, a, b) = \begin{cases} 1, & \text{агар } x < a \\ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{x-b}{b-a} \pi\right), & \text{агар } a \leq x \leq b \\ 0, & \text{агар } x > b \end{cases} \quad (3.11)$$

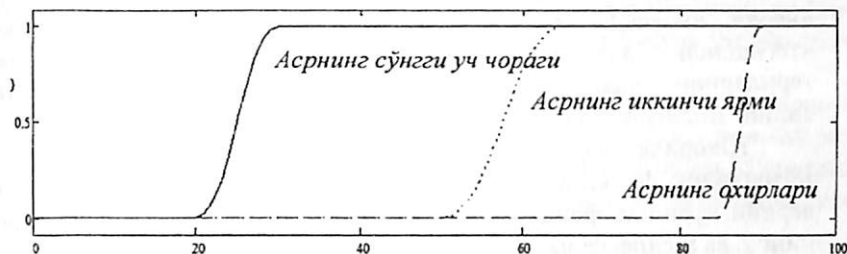
(3.11) S синф $\mu_x^S(x, \text{Асрнинг сўнги уч чораги}, 20, 30)$, $\mu_x^Z(x, \text{Асрнинг иккинчи ярми}, 50, 65)$, $\mu_x^Z(x, \text{Асрнинг охирлари}, 85, 90)$ тегишлилик функциясининг графикли тасвири 3.8-расмда тасвирланган.



3.8-расм. (3.11) тегишлилик функциясининг графикли тасвири 2-вариант.

$$\mu_x^S(x, l_x, a, b) = \begin{cases} 1, & \text{агар } x \leq a \\ 2\left(\frac{x-a}{b-a}\right)^2, & \text{агар } a < x \leq \frac{a+b}{2} \\ 1 - 2\left(\frac{b-x}{b-a}\right)^2, & \text{агар } \frac{a+b}{2} < x < b \\ 0, & \text{агар } x \geq b \end{cases} \quad (3.12)$$

(3.12) S синф $\mu_x^S(x, \text{Асрнинг сўнги уч чораги}, 20, 30)$, $\mu_x^Z(x, \text{Асрнинг иккинчи ярми}, 50, 65)$, $\mu_x^Z(x, \text{Асрнинг охирлари}, 85, 90)$ тегишлилик функциясининг графикли тасвири 3.9-расмда тасвирланган.

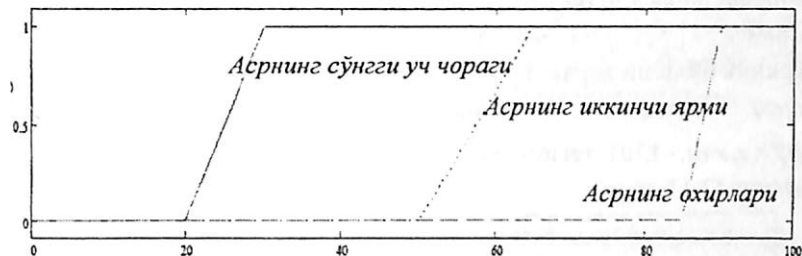


3.9-расм. (3.12) тегишлилик функциясининг графиги

3-вариант. S синф тегишлилик функцияларини ҳам чизиқли кўринишда ифодалаш мумкин.

$$\mu_x^S(x, l_x, a, b) = \begin{cases} 1, & \text{агар } x \leq a \\ \frac{x-a}{b-a}, & \text{агар } a \leq x \leq b \\ 0, & \text{агар } x \geq b \end{cases} \quad (3.13)$$

(3.13) S синф $\mu_x^S(x, \text{Asrning sўnggi uch choraqi}, 20, 30)$, $\mu_x^Z(x, \text{Asrning ikkinchi yarmi}, 50, 65)$, $\mu_x^Z(x, \text{Asrning oxirlari}, 85, 90)$ тегишлилик функциясининг графикли тасвири 3.10-расмда тасвирланган.



3.10-расм. (3.13) тегишлилик функциясининг графиги

Юкорида келтирилган S синфига кирувчи (3.11), (3.12) ва (3.13) тегишлилик функцияларида X базавий тўплам, l_x – лингвистик ўзгарувчи учун D термлар тўплами аниқланди, ҳар бир терм учун a, b сонли катталиклар киритилади ($a, b \in X = \{x_i\}$, $i = 1 \dots n$) ва 1) $a < b$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак. a ва b лар ҳар бир термнинг аниқланганлик оралиғи бўлиб ҳисобланади.

S синфига кирувчи (3.11), (3.12) ва (3.13) функцияларни маълумотларни интеллектуал излаш ва қайта ишлашда «объект

сони», «киймат», «қиймат чегараси», «оралиқ катталиқ (киймат)», «нарх», «мақсад», «хизмат», «сифат», «шакл», «тааллуқлилик», «тегишлилик даражаси оралиғи» каби характердаги норавшан термларнинг «дан, катта» мазмунидаги лингвистик ўзгарувчисининг модификаторлари учун фойдаланиш тавсия қилинади.

Юқоридагилардан шуни фарқлаш мумкинки, Z ва S синф тегишлилик функцияларидан t ва T синф тегишлилик функцияларини қуришда фойдаланиш мумкин. t ва T синф функцияларининг Z ва S синф функциялари орқали тасвирланиши қуйидагича:

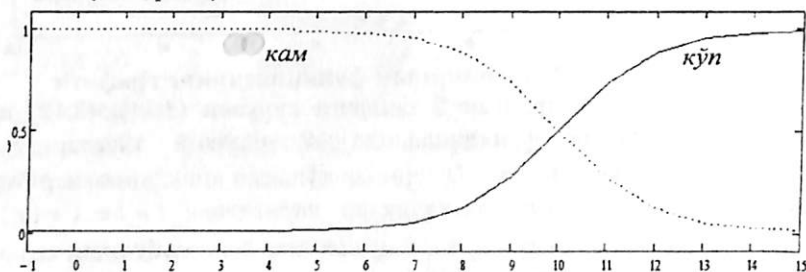
$$\mu_x^t(x, l_x, a, b, c) = \left\{ \min_{x \in X} \{ \mu_x^Z(x, l_x, a, b), \mu_x^S(x, l_x, b, c) \}, l_x \equiv \{ l_x^Z \cup l_x^S \} \right\},$$

$$\mu_x^T(x, l_x, a, b, c, d) = \left\{ \min_{x \in X} \{ \mu_x^Z(x, l_x, a, b), \mu_x^S(x, l_x, c, d) \}, l_x \equiv \{ l_x^Z \cup l_x^S \} \right\}$$

σ – тасвирли тегишлилик функцияларида (σ синф) X базавий тўплам симметрик характерда бўлганда фойдаланиш мақсадга мувофиқ [116; 76-б.]. Бу синф Z ва S синфларга тааллуқли синф бўлиб ҳисобланади ва қуйидагича ифодаланеди:

$$\mu_x^\sigma(x, l_x, b, c) = \frac{1}{1 + e^{-a(x-b)}} \quad (3.14)$$

σ синфига кирувчи (3.14) тегишлилик функцияларида X симметрик базавий тўплам, l_x – ЛЎ учун D термлар тўплами аниқланди ва ҳар бир терм учун a, b сонли катталиқлар киритилади ($a, b \in X = \{x_i\}, i=1 \dots n$) ва 1) $a < b$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак. Бу σ синф функцияси $a > 0$, бўлганда S синф, $a < 0$ бўлганда Z синфни беради ва $\mu_x^\sigma(x, \text{кўп}, 1, 10)$, $\mu_x^\sigma(x, \text{кам}, -1, 10)$ тегишлилик функцияларининг графиклари қуйидагича (3.11-рasm):



3.11-рasm. (3.14) тегишлилик функциясининг графиклари σ синфига кирувчи (3.14) тегишлилик функцияси маълумотларни интеллектуал излаш ва қайта ишлашда бир-бирига тескари

(кўп-кам, аччиқ-ширин, ок-кора, янги-эски, оддий-мураккаб, кийин-осон) характердаги норавшан термларнинг лингвистик ўзгарувчиси учун фойдаланиш тавсия қилинади.

П-тасвирли тегишлилик функцияларини (П синф) танлашнинг турли ёндашувлари мавжуд [116. 76-б.]. П синф тегишлилик функциясида $a = b$ ва $c = d$ сонли қийматлари тенг бўлса, П синфни беради. Шунингдек, бу П синф тегишлилик функцияларини қуйидагича бериш мумкин:

1-вариант. S ва Z синфларнинг тегишлилик функцияларидан фойдаланилган ҳолда

$$\mu_x^{\Pi}(x, l_x, a, b, c, d) = \mu_x^Z(x, l_x, a, b) \cdot \mu_x^S(x, l_x, c, d), \quad l_x^{\Pi} \equiv l_x^Z \cup l_x^S$$

Бу ерда X базавий тўпلام, l_x – лингвистик ўзгарувчи учун D термлар тўплами аникланди, ҳар бир терм учун a, b, c, d сонли катталиклар киритилади ($a, b, c, d \in X = \{x_i\}$, $i = 1 \dots n$) ва 1) $a \leq |b| \leq |c| \leq d$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак.

2-вариант. σ синфнинг тегишлилик функциясидан фойдаланилган ҳолда,

$$\mu_x^{\Pi}(x, l_x, a, b, c, d) = \mu_x^{\sigma}(x, l_x, a, b) \cdot \mu_x^{\sigma}(x, l_x, c, d), \quad l_x^{\Pi} \equiv l_x^{\sigma}$$

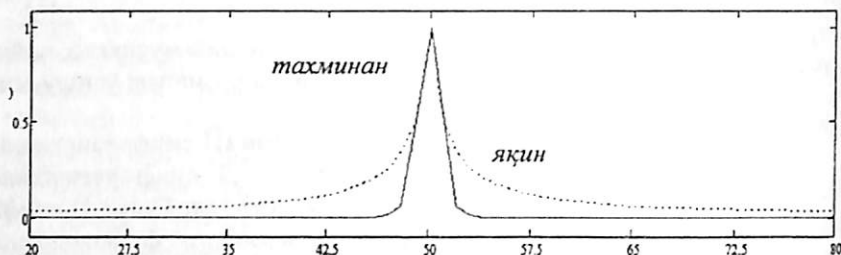
Бу ерда X симметрик базавий тўпلام, l_x – лингвистик ўзгарувчи учун D термлар тўплами аникланди, ҳар бир терм учун a, b, c, d сонли катталиклар киритилади ($a, b, c, d \in X = \{x_i\}$, $i = 1 \dots n$) ва 1) $a \leq |b| \leq |c| \leq d$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак.

3-вариант. Чизикли функциядан фойдаланилган ҳолда:

$$\mu_x^{\Pi}(x, l_x, a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (3.15)$$

Бу ерда X базавий тўпلام, l_x – лингвистик ўзгарувчи учун D термлар тўплами аникланди, ҳар бир терм учун a, b, c сонли катталиклар киритилади ($a, b, c \in X = \{x_i\}$, $i = 1 \dots n$) ва 1) $a \ll b \ll c$, 2) $a_1 \leq a_2 \leq \dots \leq a_n$ шартлар ўринли бўлиши керак.

П синф $\mu_x^{\Pi}(x, 50\text{га тахминан тенг}, 1, 2, 50)$, $\mu_x^{\Pi}\left(x, 50\text{га яқин}, 1, \frac{1}{2}, 50\right)$ тегишлилик функцияларининг графиклари қуйидагича:



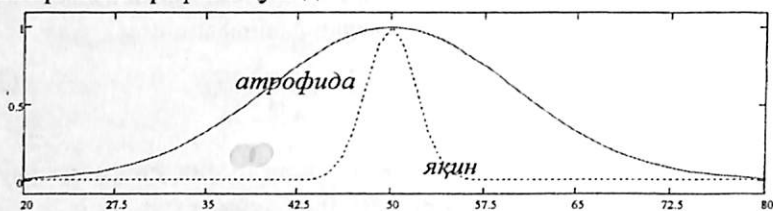
3.12-расм. (3.15) тегишлилик функциясининг графиги

П синфига кирувчи (3.15) тегишлилик функциялар маълумотларни интеллектуал излаш ва қайта ишлашда айнан термга яқин характердаги норавшан термларнинг ЛЎси учун фойдаланиш тавсия қилинади.

4-вариант. Экспонентал функциядан фойдаланилган ҳолда

$$\mu_x^{\Pi}(x, l_x, a, c) = e^{-\frac{(x-c)^2}{2a^2}} \quad (3.16)$$

Бу ерда X базавий тўплам, l_x – лингвистик ўзгарувчи учун D термлар тўплами аниқланди, ҳар бир терм учун a, c (a^2 – термнинг ноаниқлилик қиймати, c – термни қаноатлантирувчи максимал қиймат) сонли катталиқлар киритилади ($a, c \in X = \{x_i\}, i = 1 \dots n$) ва $1) a_1 \leq a_2 \leq \dots \leq a_n$ шарт ўринли бўлиши керак. Бу Π синф $\mu_x^{\Pi}(x, 50 \text{ атрофида}, 10, 50)$, $\mu_x^{\Pi}(x, 50 \text{ га яқин}, 2, 50)$ тегишлилик функцияларининг графиги куйидагича:



3.13-расм. (3.16) тегишлилик функциясининг графиги

П синфига кирувчи (3.16) тегишлилик функциясининг маълумотларни интеллектуал излаш ва қайта ишлашда айнан ва қисман ноаниқлик асосида термга яқин характердаги норавшан термларнинг лингвистик ўзгарувчиси учун фойдаланиш тавсия қилинади.

Юқорида келтирилган тегишлилик функцияларининг синфларидан маълумотларни излаш учун норавшан қоидалар ва ББсини ишлаб чиқиш лозим. Бунинг учун эса фан ва таълимга оид КАКТларининг НАМида семантик ядрони яратиш керак.

3.4. Фан ва таълимга оид маълумотларни интеллектуал излашда семантик ядро модели ва алгоритмларини ишлаб чиқиш

Фан ва таълимга оид корпоратив ахборот-кутубхона тизимларининг НАМда маълумотларни интеллектуал излашда асосий муаммо семантик ядро яратиш ва бошқаришдир. Семантик ядро жуда кўп соҳалар учун ишлаб чиқилиши мумкин. Шунинг учун НАМда маълумотларни излаш ва қайта ишлаш учун семантик ядрога таъриф келтирамиз:

Семантик ядро – нормаллаштирилган сўровлар ва электрон ресурсга мазмун-моҳиятан долзарб, мос деб ҳисобланган муносабатлар тўплами.

Ушбу таърифга асосланган ҳолда НАМда маълумотларни излаш тизимларининг семантик ядросини 3 та усул билан ишлаб чиқишни таклиф қиламиз. Булар:

- излаш ва қайта ишлаш натижасида автоматик тарзда семантик ядрони яратиш усули;
- ЭРларга берилган ҳаволалар цитатасини қайта ишлаш билан семантик ядро яратиш усули;
- эксперт гуруҳлари ёрдамида семантик ядро яратиш усули.

Биринчи усул. Бунда МИҚИ натижасида автоматик тарзда семантик ядро яратиш усули фойдаланувчиларнинг фаолиятларига боғлиқ бўлади. Фойдаланувчи q сўровни муайян термлар асосида яратади, МИТ сўровга мос d ЭРлар тўпламини тақдим қилади. Фойдаланувчи тақдим қилинган тўпламни таҳлил қилиб, сўровга энг долзарб деб ҳисоблаган ЭРларга мурожаат қилади. Мана шу мурожаат сўров (сўров терминлари) ва танланган ЭРлар орасида семантик алоқани таъминлайди. МИТларида бундай ёндашув кейинчалик киритиладиган сўров (сўров термларига) фойдаланувчилар орқали яратилган семантик алоқа ҳисобига интеллектуал излаш тамойилига ўтади.

3.1-Алгоритм. Автоматик тарзда семантик ядрони яратиш.

1-қadam. Берилган сўровни нормаллаштириш $[57] q \rightarrow q^*$.

2-қadam. Сўров, термин ва термларни ажратиш $q^* \rightarrow (q_1^*, q_2^*, q_3^*)$.

3-қadam. (q_1^*, q_2^*, q_3^*) га мос D тўпламни танлаш.

4-қadam. Сўров ва D электрон ресурслар тўплами орасида семантик муносабатни яратиш.

5-қadam. Сўровга долзарб деб топилган D^* электрон ресурсларни танлаш.

6-қadam. Сўровга долзарб деб топилган электрон ресурслар орасида семантик муносабатни яратиш.

7-қadam. Яратилган семантик алокаларни МИТ тизими модулига қайд қилиш.

Бу алгоритм маълумотларни интеллектуал излаш учун семантик ядрони фойдаланувчиларнинг интеллекллари доирасида яратади.

Иккинчи усул. Бунда ЭРга берилган ҳавола цитатасини қайта ишлаш билан семантик ядрони яратиш усулидир. Ҳар бир ЭР яратишда бошқа бир ЭРда ҳаволалар келтирилади. Ҳавола цитата орқали келтирилади ва ЭР аниқ кўрсатилади. Демак, ЭР келтирилган цитатага долзарб ҳисобланади ёки ЭРнинг асосий мазмун-моҳиятига боғлиқ бўлади. Келтирилган цитатани сўров сифатида қарайдиган бўлсак, сўров ва ЭР ҳавола орқали боғланганлигини МИТнинг семантик ядросини яратишда фойдаланиш мумкин. Бу усулнинг алгоритми қуйидагича:

3.2-алгоритм. ЭРга берилган ҳавола цитатасини қайта ишлаш билан семантик ядро яратиш.

1-қadam. Берилган t цитатани нормаллаштириш ва сўровга айлантириш $t \rightarrow q^*$.

2-қadam. Сўров, термин ва термларни ҳосил қилиш $q^* \rightarrow (q_1^*, q_2^*, q_3^*)$.

3-қadam. Цитатага берилган ҳавола бўйича электрон ресурсни аниқлаш.

4-қadam. Сўров ва электрон ресурс орасида семантик муносабатни яратиш.

5-қadam. Яратилган семантик алокани МИТ тизимининг модулига қайд қилиш.

Алгоритм маълумотларни интеллектуал излаш учун семантик ядро ЭРларнинг муаллиф интеллекллари доирасида яратади.

Учинчи усул. Эксперт гуруҳлари ёрдамида семантик ядро яратиш бўлиб, буни 2 вариантда амалга ошириш мумкин.

Биринчи вариантда эксперт электрон ресурсни танлаб, унга мос сўровларни (термларни) яратиш орқали семантик ядро яратилади. Бу вариантнинг алгоритми қуйидагича:

3.3-алгоритм. Эксперт гуруҳлари ёрдамида семантик ядро яратиш.

1-қadam. Ресурсдан d электрон ресурсни танлаш.

2-қadam. d электрон ресурсга мос q_i сўровларни киритиш.

3-қadam. d электрон ресурс ва q_i сўровлар орасида семантик алоқа яратиш.

4-қadam. Яратилган семантик муносабатни МИТ тизимининг модулига қайд қилиш.

Бу вариантда маълумотларни интеллектуал излаш учун семантик ядрони экспертнинг интеллекти доирасида яратади.

Иккинчи вариантда эксперт мавжуд семантик ядрони қайта ишлаш, зарур ҳолларда сўров ва электрон ресурсларнинг муносабатини узади, яратади, ўзгартиради. Бу вариант юкорида келтирилган алгоритмларнинг натижасини янада интеллектуаллаштиришга олиб келиш орқали семантик ядрони мустақкамлайди. Бу вариантнинг алгоритми қуйидагича:

3.4-алгоритм. Эксперт мавжуд семантик ядрони қайта ишлаш орқали семантик ядрони яратиш.

1-қadam. D электрон ресурслар тўпламини танлаш.

2-қadam. Навбатдаги d_i электрон ресурсни танлаш, агар биринчи марта бўлса, биринчи d_i электрон ресурсни танлаш.

3-қadam. Электрон ресурсга мос q сўровлар тўпламини аниқлаш.

4-қadam. Навбатдаги сўровни q_i танлаш, агар биринчи марта бўлса, q_i биринчи сўровни танлаш.

5-қadam. Сўровни таҳлил қилиш.

6-қadam. Зарур бўлса, таҳрирлаш, муносабатни узиш.

7-қadam. Охирги сўров бўлмаса, 4-қadamга ўтиш.

8-қadam. Электрон ресурсга мос янги сўровларни киритиш.

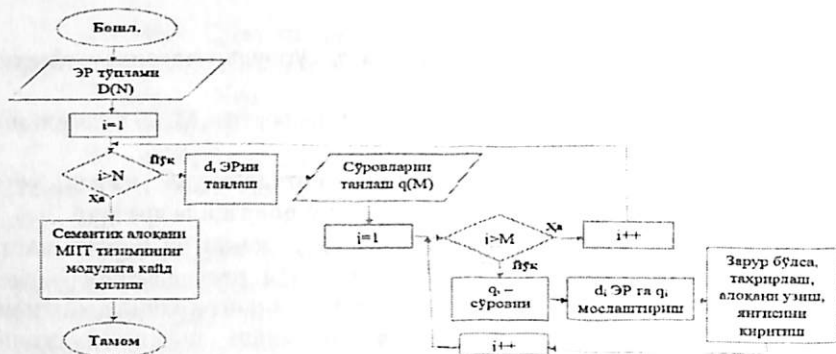
9-қadam. Электрон ресурс ва янги сўровлар орасида семантик алоқа яратиш.

10-қadam. Охирги электрон ресурс бўлмаса, 2-қadamга ўтиш.

11-қadam. Семантик алоқани МИТ тизимининг модулига қайд қилиш.

Бу вариант маълумотларни интеллектуал излаш учун семантик ядрони фойдаланувчи, электрон ресурс муаллифи, экспертлар интеллекллари асосида яратади.

Учинчи усулнинг блок-схемада тасвирланиши қуйидагича:



3.14-расм. Эксперт гурухлари ёрдамида семантик ядрони яратиш алгоритми

Бу усулда сўровни нормаллаштириш деганда, берилган бошланғич сўровдан термлар тўпламини ҳосил қилиш назарда тутилиб, қуйидаги мезонлар асосида амалга оширилади:

1- мезон. *Бир хиллик* – сўровдаги сўз ва белигилар бир алфавитга тегишли бўлишлиги.

2- мезон. *Мослашувчанлик* – таклиф қилган сўров синфларининг бирига мансублиги.

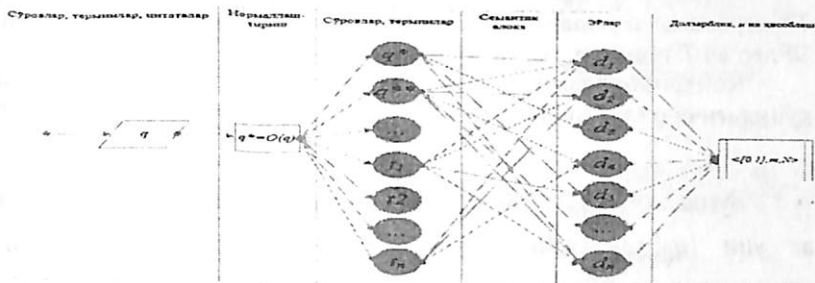
3- мезон. *Мавжудлик* – сўровда камида битта термин бўлиши.

4- мезон. *Позицияни сақлаб қолиш* – сўровдаги ҳар бир терм ва мантикий амалларнинг ўринларини сақлаб қолиш.

Бу вариантда битта сўров N та ЭРга семантик алоқага эга бўлиши мумкин ёки тескариси. Бу ҳолда эксперт бунинг қолдириши, N та ЭРга мослаб таҳрирлаши ёки жорий ЭР ва сўров алоқасини узиши мумкин.

Бу ҳолдан сўнг яна бир вазият, яъни семантик алоқага эга бўлмаган сўровлар ҳосил бўлиши мумкин. Бунинг учун МИТ модулида семантик ядро назоратчиси модули киритилади. Семантик ядронинг назоратчиси семантик алоқага эга бўлмаган сўровларни аниқлаш ва доимий экспертларга тақдим қилиш, семантик ядрога фаоллашмаган ЭРларни аниқлаш ва экспертларга тақдим қилиш, семантик ядрога таҳлиллар олиб бориш, сўровларнинг ўзаро семантик алоқасини, ЭРларнинг ўзаро семантик алоқасини аниқлаш ва экспертларга тақдим қилиш орқали интеллектуал семантик ядрони ривожлантиришдан иборат.

Юқорида келтирилган семантик ядрони яратишнинг 4 та усули ва семантик ядро назоратчисининг ижросига асосланган семантик ядро моделини қуйидагича ишлаб чиқдик:



3.15-расм. НАМда маълумотлар излаш тизими учун семантик ядро модели

Ушбу моделда долзарбликни аниқлаш учун семантик ядрога 4 та хусусият киритилган. 1-хусусият, сўров ва ЭР семантик алоқага эга бўлса 1, акс холда 0 ни қабул қилувчи хусусият, 2-хусусият, N – жами алоқалар сони, 3-хусусият, m – мурожаат қилинган алоқалар сони, 4-хусусият μ тегишлилик функциясини фаоллаштириш.

НАМнинг маълумотларни излаш тизимларида семантик ядрони яратиш модели асосида маълумотларни интеллектуал излаш ва қайта ишлашни амалга ошириш, семантик ядросини интеллектуал ривожлантиришга эришилади.

3.5. Маълумотларни интеллектуал излаш тизимида норавшан излаш модели ва қоидаларни ишлаб чиқиш

Фан ва таълимга оид корпоратив ахборот-кутубхона тизимларини НАМнинг маълумотларни излаш тизимларида маълумотларни интеллектуал излаш ва қайта ишлаш учун норавшан ББ ва қоидалар муҳим ўринга эга [73, 106]. Шунинг учун ушбу параграфда Мамдани алгоритмидаги қоидаларга асосланиб [106; 70-б.], уларни яратиш орқали ББсини ишлаб чиқишга бағишлаймиз.

Норавшан ББ – ўрганилаётган объектнинг кириш ва чиқиш параметрлари ўртасидаги ўзаро боғлиқликни аниқловчи «Агар X бўлса, у ҳолда Y » норавшан қоидалар мажмуига айтилади [106, 66].

Бунда «Агар» қоиданинг ўзаги, « X » қоиданинг шarti, «у ҳолда» қоиданинг боғловчиси, « Y » қоиданинг хулосаси ҳисобланади.

Кириш ва чиқиш параметрлари кўп параметрли бўлганлиги учун «ВА» ва «ЁКИ» мантиқий амалларидан фойдаланиб, қоидаларни ишлаб чиқиш мумкин.

НАМнинг маълумотларни излаш тизимларида маълумотларни интеллектуал излашда фойдаланиш мумкин бўлган қоидадаги шарт ва хулосасини аниқлаштириб олишимиз лозим.

Қоидадаги шарт – q норавшан сўров, T термлар тўплами мос T^* норавшан термлар тўплами, МИТнинг семантик ядроси, яъни D ЭРлар ва T термлар тўпламининг алоқасига боғлиқ бўлади.

Қоиданинг шартини шакллантириш учун q норавшан сўровни қуйидагича q^* норавшан сўровлар тўпламига акслантирамиз (3.17):

$$f: q \rightarrow q^* = \{q_1^*, q_2^*, \dots, q_n^*\} \quad (3.17)$$

Бунда $q_i^* - q_{i1}^*$ терм, q_{i2}^* норавшан терм жуфтлигидан иборат ва уни $\langle q_{i1}^* | q_{i2}^* \rangle$ каби ёзамиз, $i = 1 \dots n$, $n \leq 2 \cdot |q_{i2}^*|$, $|q_{i2}^*| - q_{i2}^*$ норавшан термлар сони.

«ВА» мантикий амали ва (3.17)нинг q_1^* терм жуфтлиги учун қоиданинг шартини қуйидагича ёзиш мумкин:

$$(q_1^* = q) \text{ ВА } (q_{11}^* = d) \text{ ВА } (q_{12}^* = t^*) \text{ ВА } (a_1 = \mu_{1,x_1}(q_{12}^*)) \quad (3.18)$$

Бунда $(q_1^* = q) - q_1^* - q$ сўровининг термларидан ташкил топган термлар жуфтлиги ёки жуфтлиги эмаслиги, $(q_{11}^* = d) - d$ ЭР ва q_{11}^* термнинг семантик алоқаси бор ёки йўқлиги, $(q_{12}^* = t^*) - q_{12}^*$ норавшан терм МБсидаги t^* норавшан термлар тўпламида мавжуд ёки мавжуд эмаслиги, $(a = \mu_x(q_{12}^*)) - (q_{11}^* = d)$ нинг a вазнининг q_{12}^* норавшан терм тегишлилик функциясининг қийматларига тегишлилиги ва тегишли эмаслиги.

(3.18) қоида битта q_1^* термин жуфтлигидан иборат сўров ва битта d ЭР учун ишлаб чиқилди. Энди битта q_i^* терм жуфтликлари ёки q сўров ва битта d ЭР учун «ЁКИ» мантикий амали ва (3.18) га асосланиб, қоиданинг шартини қуйидагича ёзиш мумкин:

$$\begin{aligned} & ((q_1^* = q) \text{ ВА } (q_{11}^* = d) \text{ ВА } (q_{12}^* = t^*) \text{ ВА } (a_1 = \mu_{1,x_1}(q_{12}^*))) \\ & \text{ЁКИ } ((q_2^* = q) \text{ ВА } (q_{21}^* = d) \text{ ВА } (q_{22}^* = t^*) \text{ ВА } (a_2 = \mu_{2,x_2}(q_{22}^*))) \\ & \text{ЁКИ } ((q_3^* = q) \text{ ВА } (q_{31}^* = d) \text{ ВА } (q_{32}^* = t^*) \text{ ВА } (a_3 = \mu_{3,x_3}(q_{32}^*))) \\ & \dots \\ & \text{ЁКИ } ((q_n^* = q) \text{ ВА } (q_{n1}^* = d) \text{ ВА } (q_{n2}^* = t^*) \text{ ВА } (a_n = \mu_{n,x_n}(q_{n2}^*))) \end{aligned}$$

Бу шарт орқали МИТларда маълумотларни интеллектуал излашда норавшан қоидани қуйидагича ёзиш мумкин:

$$\begin{aligned} & \text{АГАР } ((q_1^* = q) \text{ ВА } (q_{11}^* = d) \text{ ВА } (q_{12}^* = t^*) \text{ ВА } (a_1 = \mu_{1,x_1}(q_{12}^*))) \\ & \text{ЁКИ } ((q_2^* = q) \text{ ВА } (q_{21}^* = d) \text{ ВА } (q_{22}^* = t^*) \text{ ВА } (a_2 = \mu_{2,x_2}(q_{22}^*))) \\ & \text{ЁКИ } ((q_3^* = q) \text{ ВА } (q_{31}^* = d) \text{ ВА } (q_{32}^* = t^*) \text{ ВА } (a_3 = \mu_{3,x_3}(q_{32}^*))) \\ & \dots \end{aligned}$$

ЁКИ $((q_n^* = q) \text{ BA } (q_{n1}^* = d) \text{ BA } (q_{n2}^* = t^*) \text{ BA } (a_n = \mu_{n,x_n}(q_{n2}^*)))$
 БЎЛСА, У ҲОЛДА R ;

Ушбу норавшан коидани ихчам кўринишда ёзиб оламиз.

$$\bigcup_{i=1}^n ((q_i^* = q) \text{ BA } (q_{i1}^* = d) \text{ BA } (q_{i2}^* = t^*) \text{ BA } (a_i = \mu_{i,x_i}(q_{i2}^*))) \rightarrow R \quad (3.19)$$

(3.19) норавшан коидани МБсидаги D ЭРлар тўплами учун ёзиб олсак, (3.20) норавшан коидалар тизимига эга бўламиз. Бунда

$$D = \{d_1, d_2, \dots, d_m\}, m - \text{МБсидаги ЭРлар сони ва } \bigcup_{i=1}^m R_i \subseteq D.$$

$$\begin{cases} \bigcup_{i=1}^n ((q_i^* = q) \text{ BA } (q_{i1}^* = d_1) \text{ BA } (q_{i2}^* = t^*) \text{ BA } (a_i = \mu_{i,x_i}(q_{i2}^*))) \rightarrow R_1 \\ \bigcup_{i=1}^n ((q_i^* = q) \text{ BA } (q_{i1}^* = d_2) \text{ BA } (q_{i2}^* = t^*) \text{ BA } (a_i = \mu_{i,x_i}(q_{i2}^*))) \rightarrow R_2 \\ \dots \\ \bigcup_{i=1}^n ((q_i^* = q) \text{ BA } (q_{i1}^* = d_m) \text{ BA } (q_{i2}^* = t^*) \text{ BA } (a_i = \mu_{i,x_i}(q_{i2}^*))) \rightarrow R_m \end{cases} \quad (3.20)$$

(3.20) норавшан коидалар тизими норавшан ББсини ташкил қилади. Бу ББси маълумотларни интеллектуал излаш ва қайта ишлаш билан бирга семантик ядрони ривожлантиришга хизмат қилади.

(3.20) ББдан фойдаланиш учун доимий эксперт мутахассислар томонидан ривожлантирилиб, янгиланиб турувчи тегишлилик функцияларининг мос параметрлари, мос норавшан терм ва мос модификатори ва ҳисоблаш усули мавжуд бўлиши керак (3.8-жадвал). Бу ББнинг ядроси бўлиб ҳисобланади. Агар ББсининг ядроси ривожлантирилмаса, ББ ривожланмайди. Чунки улар ўзаро боғлиқ яхлит тизим ҳисобланади.

ББси ядросини ривожлантириш қуйидаги алгоритм асосида шаклланади.

3.5. Алгоритм. ББси ядросини ривожлантириш алгоритми.

1-қadam. ЭРни танлаш;

2-қadam. ЭРнинг семантик ядрогаги алоқаларини таҳлил қилиш;

3-қadam. ЭР учун норавшан термлар билан учрайдиган термлар тўпламини баҳолаш;

4-қadam. Ҳар бир терм учун тегишлилик функцияларининг синфини киритиш;

5-қadam. Мос норавшан терм, модификаторлар ва уларни ҳисоблаш усулини киритиш.

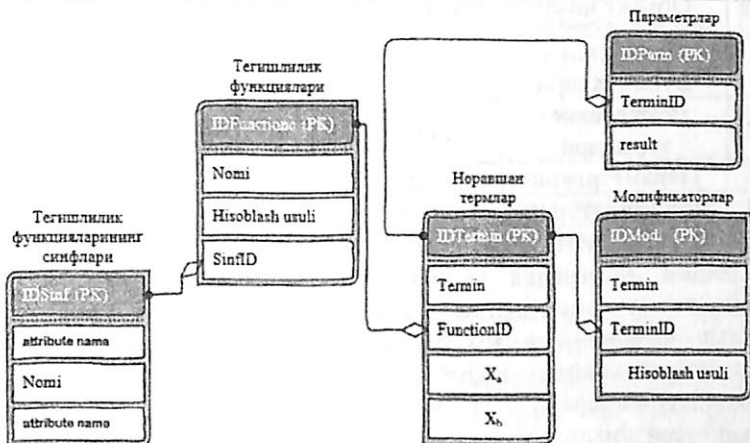
Билимлар базасининг ядроси

Тегшиллик функцияларнинг синфи	Норавшан терм ва параметрлар					Термининг chegarалари		Модификатор ва ҳисоблаш усули	
	Норавшан терм	a	b	c	d	x_a	x_b	Модифика- торлар	Ҳисоблаш усули
I синф	кичик	0	0	30	0	100	100	Жуда	$\mu_x = \mu_x^2$
								-роқ	$\mu_x = \mu_x - \frac{\mu_x}{c}$
							
II синф	катта	70	80	100	100	0	100	Модифика- торлар	Ҳисоблаш усули
								Жуда	$\mu_x = \sqrt{\mu_x}$
								-роқ	$\mu_x = \mu_x + \frac{\mu_x}{a}$
							

Норавшан терм	a	b	x_a	x_b	Модификаторлар	Ҳисоблаш усул
Z сифф	ўртача	70	0	100	Жуда	$\mu_x = \begin{cases} \mu_x - \frac{\square \mu}{10}, \text{ агар } \mu_x > \mu_x^* \\ \mu_x + \frac{\square \mu}{10}, \text{ агар } \mu_x < \mu_x^* \end{cases}$
					-роқ	$\mu_x = \begin{cases} \mu_x - \frac{\square \mu}{50}, \text{ агар } \mu_x > \mu_x^* \\ \mu_x + \frac{\square \mu}{50}, \text{ агар } \mu_x < \mu_x^* \end{cases}$
...						
Норавшан терм	a	b	x_a	x_b	Модификатор	Ҳисоблаш усул
S сифф	сифат	90	0	100	-ли	$\mu_x = \mu_x + \frac{\mu_x}{10}$
					-лироқ	$\mu_x = \sqrt{\mu_x + \frac{\mu_x}{10}}$
...						

3.8-жадвал давомли										
σ синф	Норавшан терм		a	b	x_5	x_3	Модификатор	Ҳисоблаш усул		
	оддий							-1	10	0
-роқ			$\mu_x = \mu_x - \frac{\mu_x}{10}$							
							...			
Π синф	Норавшан терм		a	b	c	d	x_5	x_3	Модификатор	Ҳисоблаш усул
	яхин	1							2	50
-роқ			$\mu_x = \mu_x + \frac{\mu_x}{10}$							
							...			

3.8-жадвалга келтирилган ББ ядросининг ўзгариши ББсига таъсир қилмайди, аммо вариантлар сони ошишига хизмат қилади.



3.16-расм. ББ ядросининг IDEF1x модели

IDEF1x моделда 5 та объект ва унинг хусусиятлари тасвирланган бўлиб, булар қуйидагилар:

1. «Тегишлилик функцияларининг синфи» – тегишлилик функция-ларининг синф номларини рўйхати учун яратилган ва 2 та «IDStaf» ва «Nomi» хусусиятларига эга.

2. «Тегишлилик функциялари» – тегишлилик функциялар номининг рўйхати, ҳисоблаш усули, функцияси ва синфи ҳақидаги маълумотни сақлаш учун яратилган, 4 та «IDFuncions», «Nomi», «Hisoblash usuli» ва «SinfID» хусусиятларига эга.

3. «Норавшан термлар» – норавшан термлар рўйхати, мос тегишли-лик функцияси, қуйи ва юкори чегаралари учун яратилган ва 5 та «IDTermin», «Termin», «FunctionID» ва « X_a », « X_b » хусусиятларига эга.

4. «Параметрлари» – бу объект норавшан термга мос параметрлар-нинг қийматлари учун яратилган бўлиб, 3 та «IDParm», «TerminID» ва «result» хусусиятларига эга.

5. «Модификаторлар» – норавшан термнинг модификаторлари рўй-хати ва уларни ҳисоблаш функциясининг номи учун яратилган бўлиб, 4 та «IDModi», «Termin» «TerminID» ва «Hisoblash usullari» хусусиятларига эга.

IDEF1x моделдаги объектлар орасида фақат бирга кўп (1:N) муносабатли боғланишлар мавжуд.

ББсининг ядроси учун IDEF1x моделидаги боғланишлар

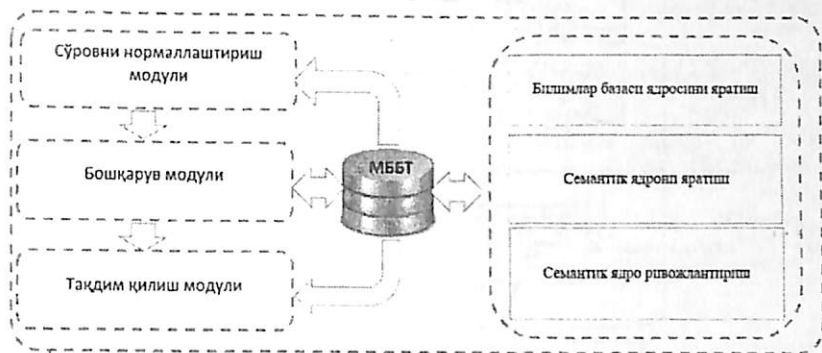
№	Объект номи	FK майдон номи	PK майдон номи	Боғланиши
1	Тегишлилик функциялари	SinfID	IDSinf	(1:N)
2	Норавшан термлари	FunctionID	IDFunctions	(1:N)
3	Параметрлари	TerminID	IDTermin	(1:N)
4	Модификаторлар	TerminID	IDTermin	(1:N)

НАМда МИТларнинг ББ ядроси учун яратилган IDEF1x моделини интеграция қилиш учун дастурий таъминотда XML технологияси тамойилига асосланган ёки моделларини қайта ишлаш имкониятига эга бўлиши лозим. Бу имконият IDEF1x моделларни қайта ишлаб, уларнинг объектларини яратиш имкониятини беради. Яратилган объектлар дастурий комплекснинг объектлари билан мустақил ва ўзаро ахборотни бошқариш ҳамда алмашиш жараёнларида ББси ядросини яратишда фойдаланилади.

3.6. Норавшан ахборот муҳитларида маълумотларни излаш тизимининг дастурий модулини яратиш

Фан ва таълимга оид корпоратив ахборот-кутубхона тизимларининг НАМларида маълумотларни интеллектуал излаш ва қайта ишлаш учун маълумотларни излаш тизимларининг асосий элементлари ва моделлаштириш омиллари, норавшан сўровларни қайта ишлаш модели, параметрик тегишлилик функцияларини лойиҳалаштириш, тегишлилик функциялари синфларини танлаш, семантик ядрони яратиш усуллари, ББсини ишлаб чиқиш ва уларнинг элементларига асосланган ҳолда НАМда МИТ инструментал дастурий модулини лойиҳалаштириш IDEF технологиялари асосида моделларини ишлаб чиқамиз.

НАМда маълумотларни излаш тизимининг умумий функционал имконияти куйидаги функционал тузилма билан изоҳланади (3.17-расм):



3.17-расм. НАМда маълумотларни излаш тизимининг умумий функционал тузилмаси.

НАМларда сўровга мос электрон ресурслар тўпламини ҳосил қилиш учун норавшан сўровларни қайта ишлаш, семантик ядрони, ББсининг ядросига асосланган ББсини яратиш бир-бири билан боғлиқ бўлган 3 та инструментал воситалар тоифасининг функционал имкониятлари тасвирланган IDEF0 моделларини ишлаб чиқиши лозим. Булар:

1-тоифа. Норавшан сўровларни қайта ишлаш орқали норавшан терм ва унинг модификаторини аниқлаш (3.18-расм).

2-тоифа. Норавшан сўровга асосланган ҳолда маълумотларни интеллектуал ишлаш учун семантик ядро яратишнинг 3 босқичи (3.19-расм).

3-тоифа. ББси ядросини яратиш орқали маълумотларни интеллектуал излашда ББсининг имконияти (3.20-расм).

Биринчи тоифада норавшан сўровларни қайта ишлаш орқали норавшан терм ва унинг модификаторини аниқлаш учун инструментал дастурий модулини лойиҳалаштириш IDEF0 технологиялари асосида келтирилган (3.18-расм).



3.18-расм. НАМда маълумотларни излаш тизимининг 1-тоифа IDEF0 модели

НАМда маълумотларни излаш тизимининг 1-тоифа IDEF0 моделига кировчи параметрлар «сўров», «МБ» ва чиқувчи параметр МБсида бажариш учун тайёр сўров бўлиб ҳисобланади.

Унда 7 та функционаллик мавжуд:

1 - *функционаллик*. Сўровни нормаллаштириш бўлиб, норавшан сўровни вақтинчалик индекслаш амалга оширилади [57].

2 - *функционаллик*. Сўровни зона шаблонларига мослигини аниқлаш бўлиб, вақтинчалик индексда зона шаблонларига мос келувчи индекслар аниқланади ва ажратиб зона индексга ўтказилади.

3 - *функционаллик*. Сўровни мантикий амалларини аниқлаш бўлиб, вақтинчалик индексда мантикий амалларга мос келувчи индекслар аниқланади, ажратиб мантикий индексга ўтказилади.

4 - *функционаллик*. Индексда қолган терминлар бўйича норавшан термларни аниқлаб олиш.

5 - *функционаллик*. Индексда аниқланган норавшан термларнинг мос тегишлилик функциясини танлаш амалга оширилади. Агар норавшан терм бир нечта синфда учраса, модификатори орқали танлаш амалга оширилади. Агар норавшан термнинг модификатори бир нечта учраса, энг кўп ишлатиладиган оддий тегишлилик функцияси танланади.

6 - *функционаллик*. Тайёрланган зона, мантикий индекс ва норавшан термнинг тегишлилик функциясининг қиймати асосида, сўров SQL стандартига ўгирилади.

7 - функционаллик. МБси учун бажаришга тайёр SQL стандартига асосланган сўров.

НАМда маълумотларни излаш тизимининг 2-тоифа IDEF0 моделида норавшан сўровга асосланган ҳолда маълумотларни интеллектуал ишлаш учун семантик ядрони яратишни мос алгоритмларига асосланган инструментал дастурий модулининг функционал имконияти келтирилган (3.19-расм).



3.19-расм. НАМда маълумотларни излаш тизимининг 2-тоифа IDEF0 модели

Ушбу моделда 5 та функционаллик мавжуд:

1 - функционаллик. Семантик ядро ишлаб чиқиш биринчи босқичининг алгоритмига асосланади (3.1-алгоритм).

2 - функционаллик. Семантик ядрони ишлаб чиқиш иккинчи босқичининг алгоритмига асосланади (3.2-алгоритм).

3 - функционаллик. Семантик ядро ишлаб чиқиш учинчи босқичи биринчи усулининг алгоритмига асосланади (3.3-алгоритм).

4 - функционаллик. Семантик ядро ишлаб чиқиш учинчи босқичининг иккинчи усулининг алгоритмига асосланади (3.4-алгоритм).

5 - функционаллик. Семантик ядро ишлаб чиқиш ҳамма алгоритмларининг охириги қадами, яъни яратилган семантик алоқаларни маълумотлар излаш тизимининг модулига қайд қилиш.

НАМда маълумотларни излаш тизимининг 2-тоифа IDEF0 моделига кирувчи параметрлар «сўров», «МБ» ва чикувчи параметр семантик алоқалар модулига қайд қилинганлиги тўғрисидаги маълумот бўлиб ҳисобланади.



3.20-расм. НАМда маълумотларни излаш тизимининг 3-тоифа IDEF0 модели

НАМда маълумотларни излаш тизимининг 3-тоифа IDEF0 моделида норавшан сўровга асосланган ҳолда маълумотларни интеллектуал ишлаш учун ББси ва унинг ядросидан фойдаланиш учун мўлжалланган инструментал дастурий модулининг функционал имконияти келтирилган (3.20-расм).

Бу моделда 5 та функционаллик мавжуд:

1 - *функционаллик*. Сўровдаги норавшан сўровларни аниқлаш бўлиб, НАМда МИТнинг 1-тоифа IDEF0 модели асосида норавшан терм ва модификаторини аниқлаб олиш.

2 - *функционаллик*. Норавшан сўровларнинг элементларини аниқлашда норавшан терм ва модификатори жойлашган ўрни асосида вариантларнинг энг яхшиси ва кўп учрайдигани яратилади.

3 - *функционаллик*. НАМда 1 маълумотларни излаш тизимининг 2-тоифа IDEF0 модели асосида яратиладиган семантик ядродан фойдаланиш.

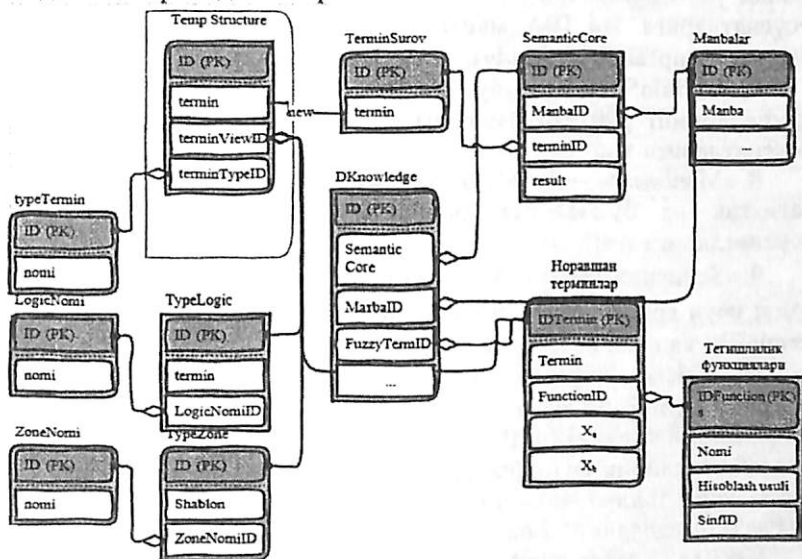
4 - *функционаллик*. Экспертлар томонидан яратилган ББсининг ядросидан фойдаланиш (3.8-жадвал ва 3.15-расмда тасвирланган IDEF1x моделга асосланади.

5 - *функционаллик*. (3.21) моделда тасвирланган ББ асосида ЭРларнинг тўпламини аниқлаш.

НАМда маълумотларни излаш тизимининг 3-тоифа IDEF0 моделига кирувчи параметрлар «сўров», «МБ» ва чикувчи параметр эса ББси асосидаги манбалар бўлиб ҳисобланади.

НАМда маълумотларни излаш тизими учун ишлаб чиқилган 3 та ўзаро боғлиқ IDEF0 моделлари инструментал дастурий таъминотни яратиш имкониятини беради. Бу воситаларни ихтиёрий дастурий комплексга интеграция, мослаштириш, ўзаро биргаликда ишлашни таъминлаш учун IDEF0 моделлар билан бирга IDEF1x модел ҳам ишлаб чиқилиши ва интеграция қилиниши керак. Чунки IDEF0 моделларини турли дастурий таъминотларга интеграция қилиш учун IDEF1x модели асосида ишлаб чиқилган маълумотлар тузилмалари муҳим ўрин тутди.

Юқорида келтирилган IDEF0 моделлар учун IDEF1x модели статик ва динамик маълумотларни сақлаш учун мўлжалланган реляцион маълумотлар тузилмасига эга бўлган реляцион МБсининг моделидир. НАМда маълумотларни излаш тизимлари учун IDEF1x модели 3.21-расмда тасвирланган.



3.21-расм. НАМда маълумотларни излаш тизимининг IDEF1x модели

IDEF1x моделда 12 та объект ва унинг хусусиятлари тасвирланган бўлиб, булар куйидагилар:

1. «TempStructure» – норавшан сўровни индекслаш учун яратилган. Уни қайта ишлаб бўлгандан сўнг ўзини ўзи (триггер асосида) холи қилади, яъни барча қиймат ва индексларни ўчиради,

бошидан бошлаш учун созланади. Бу объект 4 та – «ID», «termin», «terminViewID» ва «terminTypeID» хусусиятларига эга.

2. «typeTermin» – термин турларининг рўйхати учун яратилган ва 2 та «ID» ва «Nomi» хусусиятларига эга (зона, мантикий ва норавшан).

3. «ZoneNomi» – бу зона номларининг рўйхати учун яратилган ва 2 та «ID» ва «Nomi» хусусиятларига эга (сана, вақт, ФИО).

4. «TypeZone» – бу зона майдонларига мос шаблонларни сақлаш учун яратилган ва 3 та «ID», «Shablon» ва «ZoneNomiID» хусусиятларига эга (вақт зона майдони учун шаблонлар: «уууу-мм-dd», «уу-мм-dd», «уууу/мм/dd», «уууу.мм.dd», «уууу-dd-мм»).

5. «LogicNomi» – бу мантикий номлар рўйхати учун яратилган ва 2 та «ID» ва «Nomi» хусусиятларига эга.

6. «TypeLogic» – бу мантикий номларга мос шаблонларни сақлаш учун яратилган ва 3 та «ID», «termin» ва «LogicNomiID» хусусиятларига эга (ВА мантикий номи учун терминлар: «ва», «билан», «бирга», «и», «and»).

7. «TerminSurov» - бу МИТ бўйича барча сўровлар ва терминларнинг рўйхати учун яратилган ва 2 та «ID» ва «termin» хусусиятларига эга.

8. «Manbalar» – бу МИТ бўйича барча ЭРлар рўйхати учун яратилган ва бу IDEF1x моделида 2 та «ID» ва «Manba» хусусиятларига эга.

9. «SemanticCore» – МИТ учун семантик ядрони шакллантириш учун яратилган, бу IDEF1x моделида 4 та «ID», «ManbaID», «terminID» ва «result» хусусиятларига эга.

10. «DKnowledge» – МИТ учун ББси шакллантириш учун яратилган бўлиб, бу IDEF1x моделида 4 та «ID», «SemanticCore», «manbaID» ва «FuzzyTermID» хусусиятларига эга.

«Тегишлилик функциялари» ва «Норавшан термлар» ББнинг ядроси учун ишлаб чиқилган IDEF1x модели ҳам мавжуд (3.15-расм ва объектларнинг 2 ва 3-объектларга қаранг).

IDEF1x моделдаги объектлар орасида муносабатли боғланишлар мавжуд (3.10-жадвал).

**НАМда маълумотларни излаш тизимининг IDEF0 моделлар
учун IDEF1x моделидаги боғланишлар**

№	Объект номи	FK майдон номи	PK майдон номи	Боғланиши
1-тоифа IDEF0 модели учун				
1	Temp Structure	terminTypeID	typeTermin.ID	(1:N)
2	Temp Structure	terminViewID	TypeLogic.ID	(1:N)
3	Temp Structure	terminViewID	TypeZone.ID	(1:N)
4	Temp Structure	terminViewID	[Норавшан термлар].IDTermin	(1:N)
5	TerminSurov	termin	[Temp Structure].termin	(1:1)
6	TypeLogic	LogicNomiID	LogicNomi.ID	(1:N)
7	TypeZone	ZoneNomiID	ZoneNomi.ID	(1:N)
2-тоифа IDEF0 модели учун				
1	SemanticCore	terminID	TerminSurov.ID	(1:N)
2	SemanticCore	manbaID	Manbalar.ID	(1:N)
3-тоифа IDEF0 модели учун				
1	DKnowledge	SemanticCore	SemanticCore.ID	(1:N)
2	DKnowledge	ManbaID	Manbalar.ID	(1:N)
3	DKnowledge	FuzzyTermID	[Норавшан термлар].IDTermin	(1:N)
4	[Норавшан термлар]	FunctionID	[Тегишлилик функциялари].IDFunctions	(1:N)

Корпоратив ахборот-кутубхона тизимларининг НАМда маълумотларни излаш тизимлари учун яратилган IDEF0 ва IDEF1x моделларини интеграция қилиш учун дастурий таъминот MVC технологиясининг тамойилига асосланган ёки моделларни қайта ишлаш имкониятига эга бўлиши лозим. Бу, ўз навбатида, IDEF0 ва IDEF1x моделларни қайта ишлаб, уларнинг объектларини яратиш имкониятини беради. Яратилган объектлар дастурий таъминотнинг объектлари билан маълумотларни интеллектуал излашда семантик ядро ва ББсини шакллантириш жараёнларида фойдаланилади.

Учинчи боб бўйича хулосалар

Боб НАМда маълумотларни интеллектуал излаш ва қайта ишлаш учун норавшан сўровларни қайта ишлаш, лойиҳалаштириш,

семантик ядро ва билимлар базасини ишлаб чиқишга бағишланган бўлиб қуйидаги хулосаларга олиб келади:

1. НАМнинг маълумотлар базасида маълумотларни излашни интеллектуал амалга оширишда лингвистик ўзгарувчи учун (3.1) модель ёрдамида норавшан сўровларни қайта ишлаш алгоритмини лойиҳалаштириш усули ишлаб чиқилган. Ишлаб чиқилган усулдан табиий тилда берилган норавшан сўровларни қайта ишлаш, матнларнинг ўхшашлигини аниқлашда, таҳлил қилиш масалаларида ечишда фойдаланилади.

2. Маълумотларни интеллектуал излаш ва қайта ишлаш усулларида тегишлилик функцияларининг синфларига мос норавшан термларни танлаш, параметрлари орасидаги ўзаро муносабатлари, параметрик тегишлилик функцияларини лойиҳалаштириш усули ва дастурий модули ишлаб чиқилган. Бу турли тузилмалли ахборот-ресурслари, тизимлари учун билимлар базасининг норавшан термларга асосланган ядросини яратишга хизмат қилади.

3. Корпоратив ахборот-кутубхона тизимларининг НАМда маълумотларни излаш тизимлари учун семантик ядро модели ва семантик ядро яратишни 3 та усул ва алгоритмлари ишлаб чиқилган. Катта ҳажмли ахборот-ресурсларида маълумотларни интеллектуал излаш ва қайта ишлаш, модификациялаш учун хос семантик ядро яратиш, алгоритмлар асосида дастурий модулларнинг семантик ядросини ишлаб чиқиш, ривожлантириш ва бошқариш масалаларини тадқиқ қилишда асос бўлади.

4. Маълумотларни излаш тизимида норавшан билимлар базаси ва қоидаларини ишлаб чиқиш механизми, билимлар базасининг ядроси ва уни ривожлантириш алгоритми ишлаб чиқилган ҳамда билимлар базасининг ядроси учун IDEF1x модели лойиҳалаштирилган. Норавшан билимлар базаси ва қоидалари асосида норавшан сўровларни қайта ишлаш, матнларни таҳлил қилиш учун билимлар базасини яратиш, ривожлантириш алгоритмларини ишлаб чиқиш имконини беради.

5. Корпоратив ахборот-кутубхона тизимларининг НАМда маълумотларни излаш тизимларининг ўзаро бир-бири билан боғлиқ 3 та IDEF моделлари ишлаб чиқилган. Матнларни таҳлил қилиш, қайта ишлаш инструментал дастурий модулининг функционал имкониятини лойиҳалаштириш ва IDEF моделларини ишлаб чиқиш масалаларида, НАМда индекслаштириш, сиқиш, ахборот хавфсизлиги, техник воситаларни танлашга оид тадқиқотларда фойдаланилади.

IV-боб. КОРПОРАТИВ ТАРМОҚДА МАЪЛУМОТЛАРНИ ИЗЛАШ ВА ҚАЙТА ИШЛАШ ДАСТУРИЙ ТАЪМИНОТНИ ЯРАТИШ ҲАМДА ЖОРИЙ ҚИЛИШ

4.1. Корпоратив тармоқда маълумотларни излаш ва қайта ишлашнинг FSV технологияси

Кўп жиҳатли корпоратив тармоқнинг фан ва таълимга оид КАКТларини ахборот муҳитларининг улкан ҳажми керакли маълумотларни излаш масаласини сифат жиҳатидан янги даражага олиб чиқишни тақозо қилмоқда. Бу масаланинг ечими ЭРларнинг ўзгарувчанлиги ва уларнинг миқдори доимий ўсиб бориш омилларини мураккаблаштириб бориш билан боғлиқ [79, 120, 122].

Корпоратив ахборот муҳитларининг асосий ЭР тизимларда излаш модулларининг самарали ишлаши учун бир нечта базавий босқичлар бажарилиши лозим. Улардан энг муҳимлари қуйидагилар деб ҳисоблаймиз:

- излаш сўровига ишлов бериш;
- излаш натижаларининг сўровга мослиги;
- топилган ЭРларни тўғри ва оқилона даражалаш имконияти.

Google, Yahoo, Bing, Yandex, Rambler каби излаш модуллари катта қувватли ва энг машҳур бўлган маълумотларни излаш тизимлари миллиардлаб веб-саҳифаларни қамраб олади. Бундай тизимлар сифатли ва тез излашни таъминлашга имкон берадиган махсус алгоритмлари билан бир-биридан ажралиб туради. Лекин бу алгоритмларнинг барчаси асосий ёндашувлар – ИМларининг модификациялари ҳисобланади [36].

ИМ – реалликнинг соддалаштирилиши бўлиб, унинг асосида математик формула ва уни ЭРларга татбиқ этишнинг қоидалари асосида ҳосил қилинадиган дастурий модул. Бу формула ва қоидалар тизимга, қандай ЭРни излаш сўровига мос бўлган деб ҳисоблаш ҳамда топилган ЭРлар тўпламини қандай даражалаш масаласини ҳал этишга имкон беради. Аънанавий ва замонавий усулларнинг негизида учта асосий ёндашув мавжуд [40, 108].

Биринчи ёндашув – тўпламлар назариясига асосланади. Унинг ҳар хил кўринишлари сифатида қуйидаги турларга ажратиш мумкин:

- Буль модель (мантикий модель);
- кенгайтирилган Буль модели.

Иккинчи ёндашув – векторли алгебрага асосланади. Ушбу ёндашув

- векторли;

- умумлаштирилган векторли;
- семантик ва нейротармокли моделлар шаклида тақдим этилиши мумкин.

Учинчи ёндашув – экспертлар гуруҳи, машинали ўқув, интеллектуал таҳлил, норавшан тўпламлар назарияси ва мантиғи, *k* қийматли мантик, эҳтимоллар каби назарияларга асосланиб:

- эҳтимоллик модели;
- норавшан моделлар;
- семантик боғланишни ҳисобловчи моделларга ажратиш мумкин.

Турли муҳитларда МИҚИнинг классик моделлари ЭРларни тақдим этувчи калит сўзлар тўпламларидан фойдаланилади, бундай сўзлар терминлар деб юритилади.

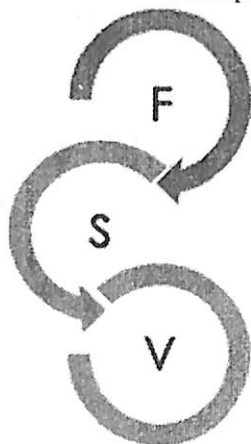
Тадқиқот ишининг олдинги параграфларида аниқланган, модификацияланган ва ишлаб чиқилган математик моделлар, усул ва алгоритмлар асосида МИТнинг инструментал дастурий таъминотини комплекс қамраб олган, интеграциялаш учун ягона тизим сифатида инструментал платформани ишлаб чиқиш зарурияти бор. Бу инструментал платформа замонавий дастурлаш технологиялари DOM, XML, ORM, MVC ва турли Framework технологиялари каби архитектураси, математик асослари, маълумотлар тузилмаси, IDEF моделлари ва кутубхоналари, бошқаришнинг инструментал воситасига эга бўлиши лозим.

Фикримизча, МИТнинг ҳар қандай математик модели ёки IDEF моделлари қуйидаги таркибий қисмлардан иборат бўлиши керак:

1. (F) - Сўровни тақдим этиш усули – тизим фойдаланувчисининг ахборот эҳтиёжларини ифодалашнинг шакллаштириш усули (F – forming).

2. (S) - ЭРнинг сўровга мувофиқлик функцияси – сўровнинг ва топилган ЭРнинг мувофиқлиги даражаси (долзарблиги, мослиги) (S – searching).

3. (V) - ЭРларни тақдим этиш усули (V – viewing).



4.1-расм.
FSV технологияси

Буни уч таркибий қисмни бирлаш-тириб, КАМларида МИТлар учун FSV технологияси (FSV платформаси, FSV Framework) деб номлаймиз (4.1-расм).

FSV технологияси асосини корпоратив ахборот муҳитларнинг НСАМларида маълумотларни излаш ва қайта ишлаш учун мазкур тадқиқот бўйича яратилган моделлар, усул ва алгоритмлар ҳамда дастурий модулларнинг функционал тузилмаси, IDEF моделлар асосида сервер иловали мижоз-сервер архитектурасига асосланган ҳолда архитектурасини ишлаб чиқамиз (4.2-расм).

Сервер иловали мижоз-сервер архитектураси фойдаланувчи ва сервер иловаларини бошқариш ва қайта ишлаш учун қулай ҳамда асосий имкониятлари қуйидагича:

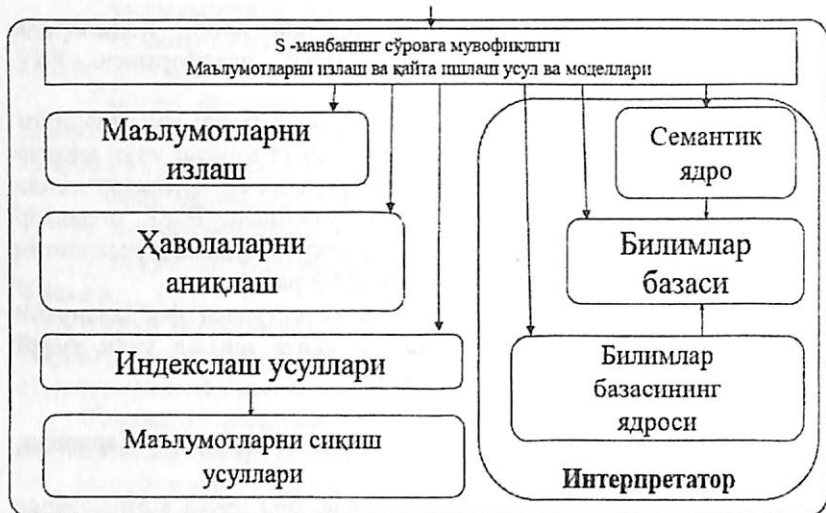
1. Нозик мижоз.

2. Мижоз ва сервер орасида минимум маълумот берилади, параметр ва натижалар.

Сервер иловалари бир неча нусхада, бир нечта компьютерда ишга туширилиши мумкин.

FSV технологияси – корпоратив ахборот муҳитларида МИҚИ моделлари, усул ва алгоритмларини интеграция ҳамда модификацияловчи, сервер иловали мижоз-сервер ахтитектурасига асосланган инструментал дастурий платформа.





4.2-расм. FSV технологиясининг архитектураси

FSV технологиясидан турли корпоратив тармоқларнинг ахборот тизимларида маълумотларни излаш ва қайта ишлаш учун фойдаланиш мумкин. Бунинг учун ахборот тизимлари қуйидаги талабларга жавоб бериши лозим:

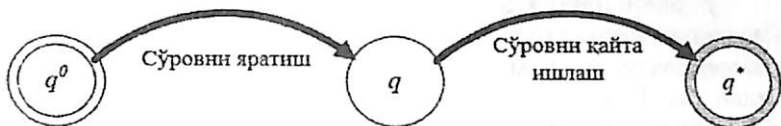
1. Сервер иловали мижоз-сервер архитектураси.
2. Уч босқичли тамойил, яъни фойдаланувчи интерфейси, бизнес-логика, маълумотларни бошқаришга асосланган.
3. MVC ва Framework технологияси асосида ишлаб чиқилган [77].
4. XML тузилмасини таҳлил қилиш ва МББТга боғлаш мавжуд бўлиши.

5. Дастурлаш тили объектга йўналтирилган бўлиши.

Ушбу талабларга жавоб берган АТларида FSV технологиясини жорий қилиш ва унинг функционал имкониятларидан МИҚИда фойдаланиш мумкин.

4.2. Корпоратив тармоқда сўровларни қайта ишлаш инструментал дастурий модулни яратиш

Сўровларни қайта ишлаш FSV технологиясининг муҳим қисми ҳисобланади ва F – сўровларни шакллантириш деб айтилади. Бу жараёни қуйидаги схема билан тасвирлаймиз [89]:



4.3-расм. Сўровларни қайта ишлаш схемаси

4.3-расмда q^0 – фойдаланувчининг ахборотга бўлган эҳтиёжи, q – фойдаланувчининг ахборот эҳтиёжининг формал кўриниши – сўров, q^* – нормаллаштирилган сўров. Сўровларни қайта ишлашни 2 та босқичга ажратамиз, яъни:

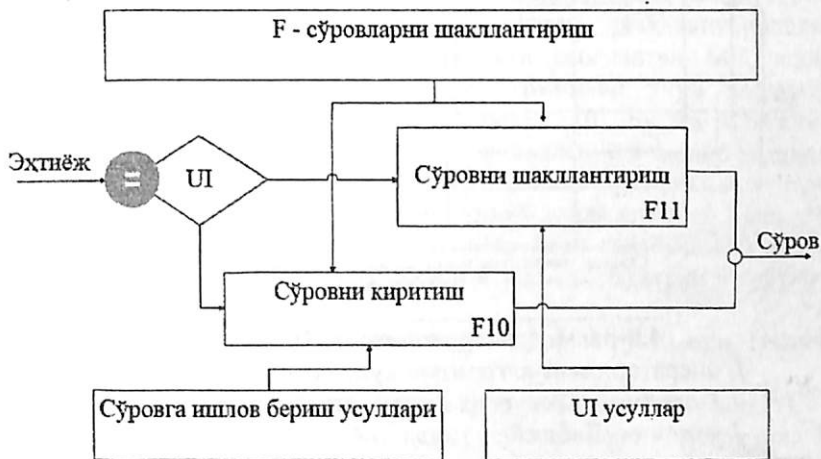
1-босқич. q сўровни яратиш $f : q^0 \rightarrow q$.

2-босқич. q^* нормаллашган сўровни яратиш $f_1 : q \rightarrow q^*$.

Бу иккала босқич алоҳида инструментал дастурий модул бўлиб, уларни мос равишда f ва f_1 операторлар билан белгилаб оламиз.

f оператори вариант сўровлар, ўрин алмаштириш, масофали таҳрирлаш, масофали таҳрирлашда k – граммли индекс, сўров хатоларини тузатиш, контекст нуқтаи назар асосида хатоларни тузатиш, фонетика асосида хатоларни тузатиш, тайёр сўровларни танлаш каби сўровга ишлов бериш усуллари мажмуасидан фойдаланилади. Фойдаланувчининг бошланғич сўровини шакллантиришни бошлаши биланок бу усуллар сўровни яратиш учун таъсир кўрсатади.

f операторининг IDEF0 моделини қуйидагича ишлаб чиқамиз:



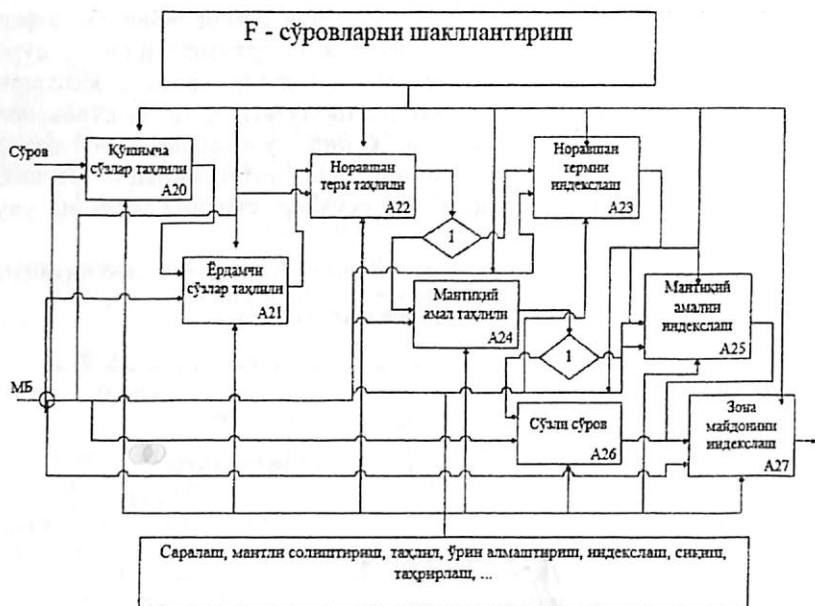
4.4-расм. f операторининг IDEF0 модели

f оператори UIга асосланган аклий ҳужум, синов, тест, масалани ечиш, муаммо, савол-жавоб каби фойдаланувчининг ахборот эҳтиёжини аниқлашга мўлжалланган UI усуллар мажмуасидан ҳам фойдаланади. Бунда фойдаланувчи сўровни яратиш учун UI усуллардан бирини танлаши керак. UI усул фойдаланувчининг эҳтиёжини аниқлайди ва мос сўровни яратиб беради.

f операторининг сўровга ишлов бериш усуллари асосида яратилган сўровлар табиий тилда тузилган сўров деб айтамыз.

f операторининг UI усуллар асосида яратиладиган сўров сунъий (машина) сўрови дейилади.

f_1 оператори эса рекурент кетма-кетликда бўлиб, кўшимча ва ёрдамчи терминлардан тозалаш, сўровни турини аниқлаш, зона майдон, мантикий ва норавшан термларнинг индексларини яратиш орқали амалга оширилади. f_1 операторининг IDEF0 моделини келтирамыз.



4.5-расм. f_1 операторининг IDEF0 модели

f_1 операторининг алгоритми қуйидагича амалга оширилади:

4.1-алгоритм. Рекурент кетма-кетликда сўровни қайта ишлаш.

1-қadam. Табиий тилда ёзилган сўровда кўшимча терминлардан тозалаш.

2-қadam. Сўровни ёрдамчи терминлардан тозалаш.

3-қadam. Сўровда норавшан терм мавжудлигини текшириш.

4-қadam. Агар сўровда норавшан терм бўлса, сўров норавшан сўров ҳисобланиб, 8-қadamга ўтиш.

5-қadam. Сўровда мантикий амал мавжудлигини текшириш.

6-қadam. Агар сўровда мантикий амал бўлса, сўров мантикий сўров ҳисобланиб, 9-қadamга ўтиш.

7-қadam. Сўров сўзли сўров ҳисобланиб, 10-қadamга ўтиш.

8-қadam. Норавшан термларни индекслаш.

9-қadam. Мантикий амалларни индекслаш.

10-қadam. Зона майдонларини аниқлаш ва индекслаш.

11-қadam. Сўров семантикасини аниқлаш [10].

12-қadam. Нормаллаштирилган сўров.

f ва f_1 операторлар ёрдамида фойдаланувчининг ахборотга бўлган эҳтиёжини нормаллаштирилган сўровга олиб келинади. Шунинг ҳам таъкидлаш лозимки, UI усуллари ёрдамида тайёрланган сўровлар жуда тез оптималлашган сўровга ўтади.

F – сўровларни шакллантириш жараёнида бошланғич маълумотлар, яъни кўшимча терминлар тўплами, ёрдамчи сўзлар тўплами, норавшан термлар тўплами, мантикий амаллар тўплами, зона майдонлари тўплами FSV технологияси учун олдиндан тайёрланган статик маълумотномалар бўлиб ҳисобланади. Буларни кейинчалик ҳам тўлдириб, бойитиб бориш мумкин.

F – сўровларни шакллантириш жараёнининг натижасида FSV технологиясида сўровлар ва терминларнинг тўплами ҳамда сўров терминининг тахминий семантик алоқаси шакллантирилади. Буларнинг асосида FSV технологиясига асосланган МИТларда тайёр, олдиндан маълум сўров ва терминлар учун электрон ресурсларнинг семантик алоқаси яратилади [10]. Бу эса МИТнинг сезиларли даражада вақтдан самарали фойдаланишига олиб келади.

Биз таклиф бераётган UI усуллари фойдаланувчига ҳеч қандай қийинчилик туғдирмасдан сўровни сунъий тилда яратади. Буни эса машина махсус тилда тузилган сўровлар каби қайта ишлайди. UI усуллари фикрларни излаш усулларига таянади, буларни 3 гуруҳга ажратиш мумкин:

1. Психологик фикрлашни фаоллаштириш усуллари (аклий ҳужум, тескари аклий ҳужум).

2. Фикрни тизимли излаш усуллари (саволлар рўйхати, тест).

3. Фикрни излашга қаратилган усуллар (венн диаграмма, 7 бошқарув элементи).

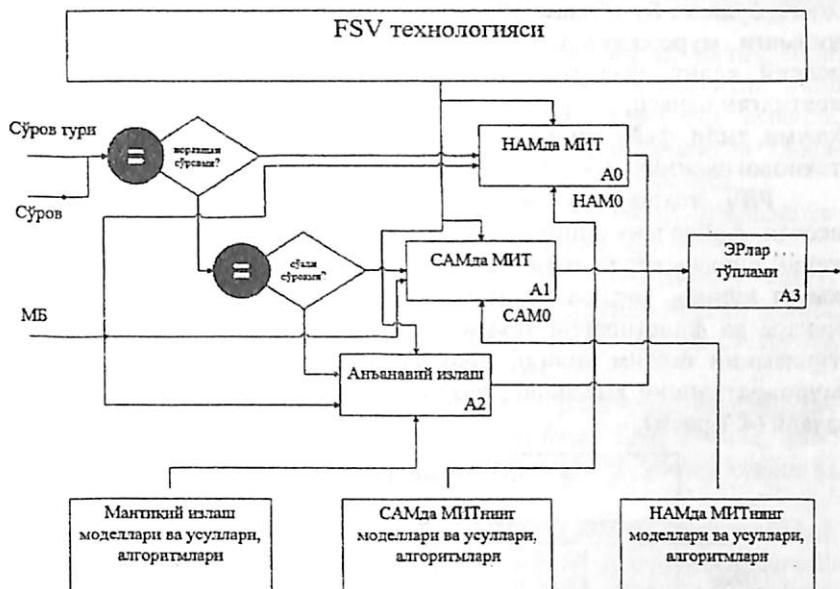
Бундай усуллар кўп бўлиб, булардан МИТлар учун сўровларни шакллантиришда самарали фойдаланиш мумкин. Аммо буларнинг инструментал дастурий модулини ишлаб чиқиш ва МИТда жорий қилиш лозим. Шунингдек, сўровларни шакллантиришда инсон-компьютер мулоқот усулларида ҳам фойдаланиш мумкин. FSV технологиясида UI усулларнинг 3 тасини, яъни аклий хужум (5 босқич), саволлар рўйхати, 7 элемент бошқаруви учун инструментал дастурий модуллар яратилган.

4.3. Корпоратив тармоқда маълумотларни излаш ва такдим қилиш инструментал дастурий модулини яратиш

FSV технологияси архитектураси 3 та асосий қисмдан иборат (4.2-расм). Бу қисмлардан бирининг функционал имкониятини олдинги параграфда ишлаб чиқдик. Бу параграфда S – ЭРнинг сўровга мувофиқлигини ҳисоблаш жараёнлари ва V- ЭРлар такдим этишнинг усули – қисмларнинг функционал имкониятларини ишлаб чиқамиз. Бу иккита қисм бир-биридан алоҳида инструментал дастурий модул сифатида қаралади.

S – ЭРнинг сўровга мувофиқлигини ҳисоблаш жараёнлари ҳаволалар асосида электрон ресурснинг долзарблигини ҳисоблаш, САМда маълумотларни излаш тизими дастурий модулининг IDEF0 моделлари ва НАМда маълумотларни излаш тизимининг дастурий модулининг IDEF0 моделлари ҳамда F-сўровларни шакллантириш қисмининг сўровлар ва терминлар тўплами, V- ЭРларни такдим этишнинг ҳисобчи-назоратчи функцияларини статистик маълумотларга асосланади (4.6-расм).

Қуйида келтирилган IDEF0 модели FSV технологиясининг ядроси бўлиб унга кирувчи параметрлар «сўров», «сўров тури», «МБ» ва чиқиш параметрлари «ЭРлар тўплами»дир.



4.6-расм. S – ЭРнинг сўровга мувофиқлигини ҳисоблаш жараёнларининг IDEF0 модели

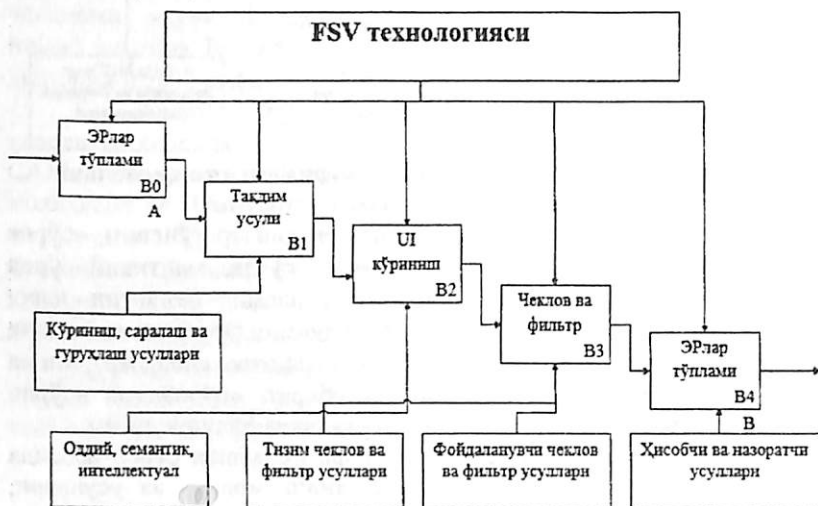
Кирувчи параметрларда «сўров» терминлар тўплами, «сўров тури» сўровнинг тури, яъни норавшан, сўзли, мантикий сўров бўлиши мумкин, «МБ» ЭРларнинг тўплами, семантик ядро, норавшан ББси ва ядроси, сўровлар ва терминлар тўплами, статик маълумот ва ҳисоботлар, мантикий семантик алоқалар, йиғма динамик ва статик маълумотлардан иборат. «сўров» ва «сўров тури» жуфтлик ҳисобланади, доим биргаликда бўлиши лозим.

IDEF0 моделида «НАМда МИТ»да норавшан сўров асосида НАМда маълумотларни излаш тизимининг модель ва усуллари, алгоритмларига асосланади. «САМда МИТ»да сўзли сўров (зона майдонли сўров) асосида САМда маълумотларни излаш тизимининг моделлари ва усуллари, алгоритмларига асосланади. «Анъанавий излаш» да мантикий, махсус лингвистик восита асосида яратилган сўровлар мантикий излаш моделларининг индекслаш усуллари [14], термин-ЭР матрицаси модели, маълумотларни излашда синфлаш ва кластеризациянинг усуллари, матнли излаш, саралаш каби усулларидан фойдаланилади.

IDEF0 моделида чиқиш параметри «Манбалар тўплами», яъни сўровга мос долзарб ҳисобланган ЭРларнинг сараланмаган тўплами

хосил бўлади. Бу тўплам ЭР номи, ЭРнинг уникал коди, долзарблик киймати, мурожаатлар сони, ЭРнинг бошқа ЭРлар билан алоқаси, асосий калит сўзлари (термин, сўров кўринишида), муаллифи, яратилган санаси, охириги мурожаат санаси, ёрлиғи, тўлик матннинг ҳажми, тили, файл типи каби маълумотлар тузилмасига эга. Бу FSV технологиясининг V- ЭРларни тақдим этиш қисми учун зарур.

FSV технологиясининг V- ЭРларни тақдим этиш қисми, асосан, фойдаланувчининг хоҳиши ва тизимнинг имкониятларидан келиб чиққан ҳолда амалга оширилади. Бу қисм ЭРлар тўпламини қабул қилиш, тақдим усулини танлаш, UI кўринишни аниқлаш, чеклов ва филтрларни таъминлаш ва шуларга асосланган ЭРлар тўпламини тақдим қилиш, фойдаланувчининг ЭРлар тўпламидаги мурожаатларини ҳисоблаш, назорат қилиш жараёнларини ўз ичига олади (4.7-расм).



4.7-расм. V- ЭРларни тақдим этишнинг IDEF0 модели

Бу моделда кирувчи ва чиқувчи параметрлар ЭРлар тўплами бўлиб, улар бир-биридан мазмун жиҳатдан фарқ қилмайдиган, фойдаланувчига қулай бўлган шакли билан фарқланади.

Моделда «Тақдим усули» 3та гуруҳга ажратилади. Булар:

- кўриниш усуллари (жадвал, ёрлик асосида, рўйхат, тартибсиз);
- саралаш усуллари (ЭР коди, долзарблиғи, мурожаатлар сони, санаси, охириги мурожаат санаси, матн тили, матн ҳажми);

- гуруҳлаш усуллари (муружаатлар сони, санаси, охири муружаат санаси, матн тили, матн ҳажми).

Моделда «UI кўриниши» ҳам 3 та кўринишда: оддий, яъни фақат ЭР номи ва бир нечта параметрлари билан, семантик, яъни ЭРнинг бошқа ЭРлар билан алоқаси, муружаатлар асосида, интеллектуал, яъни ББсининг ядроси асосидаги семантик алоқалари, ҳаволаларининг рейтинги асосида амалга оширилади.

Моделда «Чеклов ва филтър», асосан, тизим ва фойдаланувчининг талаби асосида аниқланади ва 2 гуруҳга ажратилади:

- тизимнинг чеклов ва филтърлари. Бунда тизимнинг махсус талаблари сўровга мос маълумотларнинг тўпламини тақдим қилиш, муружаатлар ЭРларни тўпламга киритиш, янги ЭРларни олдинроққа жойлаштириш ва бошқалар кўйилади;

- фойдаланувчининг чеклов ва филтърлари. Бунда фойдаланувчи йил бўйича чеклов, нашр тури бўйича, соҳа бўйича, файл типи, матн тили, файл ҳажми, махсус калит сўзлар бўйича чеклов ва филтърларни ўрнатиши мумкин.

Моделда, асосан, созлашлар фойдаланувчининг шахсий кабинетидан амалга оширилади. Агар созлаш амалга оширилмаса одатий (жорий) тизимнинг созланмалари билан тақдим қилади. V-ЭРларни тақдим этишнинг жорий созланма асосида тақдим қилган маълумотлар тўплами 4.8-расмда тасвирланган.

Элементар 1 - 25 дан 31

Кейинги санифага ўтиш

№	И	КОД	Луи	Тўплам
1	☐	44	<p>Муратов, А. К. Усиление качества продукции на предприятии в целях роста его конкурентоспособности :ВКР [002] / А. К. Муратов. - Ташкент: 2014г. - . - 85 сд.-</p>	☐
2	☐	199	<p>Рахмонов, А. М. Универсальный экзаменатор излучает и проверяет исправность поставленного тока :ВКР [002] / А. М. Рахмонов. - Ташкент: 2014г. - . - 61 сд.-</p>	☐
3	☐	465	<p>Масмонов, К. Ш. Использование ресурсов сетей для изучения кредитной активности в ИКС :ВКР [002] / К. Ш. Масмонов. - Ташкент: 2014г. - . - 72 сд.-</p>	☐
4	☐	490	<p>Халимуллоев, М. Х. Мультиязычные термины в транслитерации координатных базисов :БМИ [002] / М. Х. Халимуллоев. - Ташкент: 2014г. - . - 72 сд.-</p>	☐
5	☐	834	<p>Рахмонов, Б. Н. Построение систем защиты информации для программных комплексов, используемых в информационном доступе :ВКР [002] / Б. Н. Рахмонов. - Ташкент: 2014г. - . - 67 сд.-</p>	☐

4.8-расм. Электрон ресурслар рўйхати

Маълумотларни тақдим қилиш жараёни маълумотларни излаш тизимлари учун муҳим ҳисобланади. Чунки сўровга мос аниқланган ЭРларни тақдим қилиш фойдаланувчига ҳеч қандай шубҳа туғдирмаслиги ёки чарчатмаслиги, интерактивлиги билан фойдаланувчига таассурот қолдириши лозим. Бу уни яна шу тизимдан фойдаланишга чорлайди, тизим эса фойдаланувчидан

семантик ядро, сўровлар тўплами, уларнинг тахминий семантик алоқаси, ББсини ривожлантиришда фойдаланади.

Маълумотларни излаш тизимининг фойдаланувчилар сони қанча кўп бўлса, интеллектуал излаш ва қайта ишлаш аниқлиги шунча юқори бўлади.

4.4. FSV технологиясини жорий қилишда маълумотлар базасини ишлаб чиқиш

FSV технологиясини бирор дастурий комплексга жорий қилишда МБсини лойиҳалаштириш муҳим ўрин тутди. Чунки FSV технологиясининг ўз МБси ва жорий қилиш керак бўлган дастурий комплекснинг ўз МБси мавжуд. Жорий қилишда тизим фойдаланувчиларининг талаб ва таклифлари асосида янги бирлашувчи, ажратилувчи маълумот тузилмалари ҳосил бўлиши мумкин. Бунинг учун, аввало, FSV технологияси ва жорий қилиш керак бўлган дастурий комплекс МБсининг маълумот тузилмалари ҳамда IDEF1x моделини билиш талаб қилинади. FSV технологияси кўп дастурий комплексларга жорий қилиш олдиндан аниқ бўлганлиги учун дастурий комплексда асосий ЭРлар тўплами ва унинг тўлиқ матнлари учун маълумотлар тузилмаси берилган бўлиши керак. Умумийлик учун ушбу маълумот тузилмасини жадвал асосида келтирамиз (4.1 ва 4.2-жадваллар).

4.1-жадвал

ЭРлар тўплами учун маълумот тузилмаси

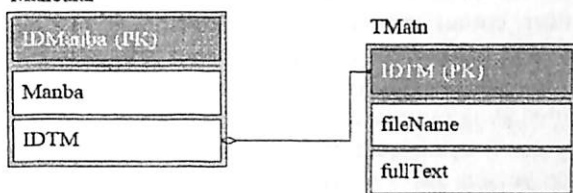
Т.р	Майдон номи	Майдон вазифаси	Типи	Null	PK/FK	триггер
1	IDmanba	ЭРнинг уникал коди	Int	NN	PK	inc
2	manba	ЭРнинг библиографик таснифи	Nvarchar(1024)	NN		MARC21
3	TMID	ЭРнинг тўлиқ матнининг уникал коди	Int	NN	FK	

ЭРларнинг тўлик матн тўплами учун маълумот тузилмаси

Т.р	Майдон номи	Майдон вазифаси	Типи	Null	PK/FK	Триггер
1	IDTM	Тўлик матннинг уникал коди	Int	NN	PK	Inc
2	fileName	Тўлик матннинг номи	Nvarchar(250)	NN		
3	fullText	Тўлик матн	objectofTypeFile	NN		fileSystem

4.1 ва 4.2-жадвалларда келтирилган маълумотлар тузилмаси учун куйидагича IDEF1x модель бўлиши керак.

Manbalar



4.9-расм. Дастурий комплекснинг IDEF1x модели

Келтирилган дастурий комплекснинг IDEF1x модели ва FSV технологиясининг IDEF1x моделларини ўзаро бир нечта маълумот тузилмасига келтириш мумкин. Бунинг учун FSV технологиясидаги асосий маълумот тузилма ва вазифаларини билиш керак (4.3-жадвал).

FSV технологиясининг маълумот тузилмалари рўйхати

Т.р.	Маълумот тузилмаси	Маълумот тузилмасининг вазифаси
1	Терминлар индекси	Термин ва сўровларнинг тўпламлари ҳамда индекси
2	1-семантик алоқа	Термин, атрибут ва ЭРларнинг алоқаси
3	Терминлар сони	ЭРдаги терминлар сони ва g қиймати
4	2-семантик алоқа	Сўров ва ЭРнинг семантик алоқаси
5	Норавшан термлар индекси	Норавшан термлар тўплами ва боғланишлари
6	И-семантик ядро	Интеллектуал семантик ядро – сўров ва ЭРларни НАМда яратиш асосида

7	Вақтинчалик индекс	Сўровларни нормаллашда терминлар индексини яратиш учун
8	Зона ва шаблонлар	Сўровнинг мос зоналари ва уларнинг шаблонлар тўплами
9	Мантиқий амаллар	Мантиқий амаллар ва уларнинг ёзилиши
10	ЭРлар	МИТдаги ЭРлар рўйхати

FSV технологиясидаги асосий маълумот тузилмалари ва дастурий комплекс маълумот тузилмаларини интеграциялаш учун статик маълумотларни йиғиш учун «Статик маълумотлар», статик ва динамик ҳисобот учун «Йиғма ҳисоботлар», маълумотлар алмашиши ва бошқарилиши учун «Қоидалар», маълумот тузилмаларини боғлаш учун «янги алоқалар», вақтинчалик маълумотлар учун «Вақтинчалик» маълумот тузилмаларини яратиш лозим. Ҳар бир маълумот тузилмаси ўзининг бирламчи калитига эга бўлиши шарт. Бирламчи калит деганда, маълумот объектининг нусхаси ягоналигини таъминловчи майдон ёки майдонлар гуруҳи тушунилади, унинг сонли типдан ташкил топиши мақсадга мувофиқ. Маълумотлар тузилмасида бирламчи калитни киритишдан мақсад маълумотларни излашда ягоналик, тартиблаштириш, саралаш ва танлаб олишда қулайликни таъминлашдан иборат.

Маълумот тузилмасидан фойдаланишда, яъни ундаги маълумотлар устида ҳар хил амаллар бажаришда бир неча усуллари бўлиб, улар маълумотлардан фойдаланишга мурожаат усуллари дейилади [52; 70-48-б.]. FSV технологиясида маълумотлардан фойдаланишга мурожаат усуллари қуйидагиларни киритамиз:

Кетма-кет мурожаат усули. Кетма-кет мурожаат усули маълумот тузилмасига сўровларнинг бажарилиши учун маълумот тузилмасидаги барча ёзувларни таҳлил қилади, биринчи ёзувдан охириги ёзувгача керакли ёзувларни берилган сўров бўйича таҳлил қилиб, сўровга мосини тақдим қилади.

Индексли кетма-кет мурожаат усули. Индексли кетма-кет мурожаат усули маълумот тузилмасига берилган сўров бажарилиши учун ёзув кўрсаткичини сўров шартини бажарувчи ёзувнинг биринчисига олиб қилиб қўяди. Кейин ёзув кўрсаткичи сўров шартини қаноатлантирувчи кейинги қаторга ўтади. Шундай қилиб, сўров шартини қаноатлантирувчи ҳамма қатор аниқланади [14].

Тўғридан-тўғри мурожаат усули. Индексли кетма-кет мурожаат усулида маълумот тузилмасидан бир ёки бир неча майдонлар гуруҳи қиймати бўйича тўғридан-тўғри мос келувчи ёзувлар танлаб олинади. Бу усул иккинчи усулнинг хусусий ҳоли

деб караш мумкин. Чунки индексли кетма-кет мурожаат усули тўғридан-тўғри усулни ишлатади. Ёзув кўрсаткичи сўров шартини каноатлантирадиган индекснинг биринчи каторига қўйилади ва керакли шартни каноатлантирувчи индексли ёзувларни аниқлаб бўлганидан сўнг кетма-кет мурожаат усули ва катор индекси бўйича силжийди [24].

Мурожаат усулларига асосланиб, FSV технологиясини дастурий комплексга жорий қилишда асосий зарур маълумот тузилмаларини келтирамиз (4.4-4.6-жадваллар).

4.4-жадвал

Статик маълумотларни йиғиш учун «Статик маълумотлар» маълумот тузилмаси

Т.р	Майдон номи	Майдон вазифаси	Типи	Null	PK/FK	Триггер
1	ID	уникал коди	Int	NN	PK	Inc
2	staticParm	Ихтиёрий параметр	Nvarchar(500)	NN		
3	Type	Параметрнинг нима эканлиги	Nvarchar(50)	NN		
4	staticID	Мавжуд бўлса алоқаси	Int	N	FK	
5	isStatic	Статик ёки динамиклигини аниқлаш учун	Bin	NN		

4.5-жадвал

Маълумотлар алмашиши ва бошқарилиши учун «Қондалар» маълумот тузилмаси

Т.р	Майдон номи	Майдон вазифаси	Типи	Null	PK/FK	Триггер
1	ID	уникал коди	Int	NN	PK	Inc
2	ruleIf	Қоида	Nvarchar(50)	NN		
3	A	True учун амал	Nvarchar(100)	NN		
4	B	False учун амал	Nvarchar(100)	N		
5	pram	Параметр борлиги	Bin	NN		

**Маълумот тузилмаларини боғлаш учун «Янги алоқалар»
маълумот тузилмаси**

Т.р	Майдон номи	Майдон вазифаси	Типи	Null	PK/ FK	Триггер
1	ID	Тўлик матннинг уникал коди	Int	NN	PK	Inc
2	OneID	Биринчи объект	Int	NN	FK	
3	typeID1	Объект тури	Nvarchar(100)	NN		
4	TwoID	Иккинчи объект	Int	NN	FK	
5	typeID2	Объект тури	Nvarchar(100)	NN		

FSV технологиясини дастурий комплексга жорий қилишда маълумот тузилмалари маълумот алмашиш кўприги бўлиб ҳисобланади. Умуман олганда, жорий қилиш учун дастурий комплексда 2 та ўзаро боғланган маълумот тузилмаси ва юқорида келтирилган 3 та маълумот тузилмаси керак. Буларни яратиш FSV технологиясида тайёрланган XML тузилмани импорт қилиш ва дастурий модулнинг битта асосий тузилмаси номи ва уникал майдонини киритиш талаб қилинади.

4.5. FSV технологиясини дастурий комплексларга жорий қилиш ва унинг самарадорлиги

FSV технологиясини дастурий комплексларга жорий қилиш учун маълумотлар базасини лойҳалаштиришни олдинги параграфда келтирдик. Бу параграфда FSV технологиясини дастурий комплексларнинг сервер дастурлари ва UI интерфейсларига жорий қилишни келтирамиз. Олдинги параграфларда FSV технологиясини интеграция қилиш учун талаблар асосида дастурий таъминот MVC технологияси асосида яратилган бўлиши кераклиги таъкидланган эди. MVC технологияси модель, контроль (бошқарув) ва кўриниш (UI) элементларидан иборат бўлади.

Модель – ORM технологиясига асосланган ҳолда МБсидаги маълумотлар тузилмаси ва уларнинг ўзаро мантиқий боғлиқлигининг синф кўриниши асосида яратади. FSV технологиясини дастурий комплексларга жорий қилишда моделларни дастур яратиш учун танланган framework инструментлари орқали яратилади.

Кўриниш (UI) – фойдаланувчининг интерфейси ҳисобланади. FSV технологиясида маълумотларни излаш учун иккита интерфейс

мавжуд (4.10-расм). Булар корпоратив тармоқда фойдаланилган-лиги учун DOM технологияси асосида ишлаб чиқилган. Бу икки интерфейс файлларини дастурий комплекснинг UI лар жойлашган манзилга нусхалаш керак. Агар дастурий комплексда бир неча босқичли бошқарувга таянилган бўлса, ҳар бир бошқарув учун алоҳида нусхаланади.



4.10-расм. Маълумотларни излаш UIнинг кўриниши

Интерфейсларда модель ва бошқарув элементи билан боғланиш мавжуд. Бу боғланишларни жорий қилиш вақтида қўл билан ўзгартириш керак.

Контроль (бошқарув) – дастурий комплекснинг асосий элементи бўлиб, модель ва UI орасидаги амалларни бажариш учун мўлжалланган. FSV технологиясининг бошқарувида асосий элементлар кутубхона шаклида яратилган, уларга тўғридан-тўғри муурожаат қилиш мумкин. Шунинг учун бу кутубхоналар дастурий комплексга импорт қилинади. UIлар билан боғлиқ бошқарувлар алоҳида сошлаб чиқилади.

FSV технологияси жорий қилинган дастурий комплекслар ЭРларни йиғиш, қайта ишлаш, тарқатиш, фойдаланишнинг турли даражали имкониятларига эга бўлади. Бундай тизимлар жамиятнинг ҳар бир соҳасида учрайди. Аммо бу жараён айнан АРМлари билан боғлиқ. FSV технологияси жорий қилинган дастурий таъминотлар рўйхатини келтирамыз:

1. Армаг++ – Ахборот-ресурс марказларининг автоматлашган тизими АРМАТнинг янги MVC технологиясига асосланиб, яратилган варианты.

2. МБИ ва МД МБ – Тошкент ахборот технологиялари университети ва унинг филиаллари талабаларининг малакавий битирув иши ва магистрларининг магистрлик диссертацияларининг МБсини яратиш ва фойдаланиш.

3. InterActiveNEW (IS-ICT) – фан ва таълимга оид энг сўнги янгиликлар МБсини яратиш ва жорий қилиш, фойдаланиш дастурий таъминот[51].

4. Smart-sinov – талабалар билимларини текширишга мўлжалланган, саволлар ва жавоблар МБсига эга дастурий таъминот.

5. Mobile-Library – кутубхона ресурсларидан мобиль алоқа дастурий таъминоти.

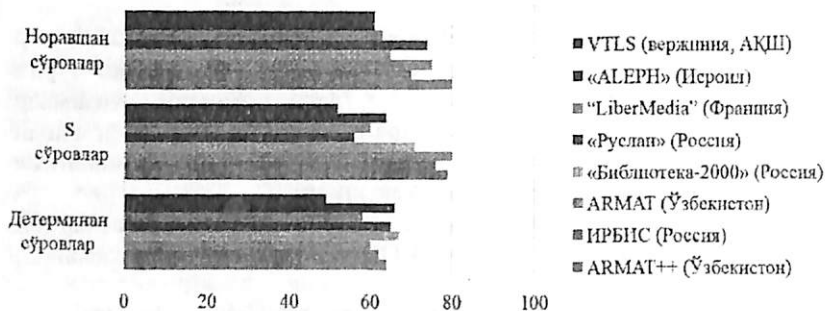
6. IS-ICT – БМИ ва МД натижалари ҳамда ютуқлари билан алмашиш имкониятини берувчи виртуал биржа платформаси.

7. vLibrary – XIX асрдаги Ўрта Осиёлик турклар ёзма маданий меросларининг виртуал кутубхонаси (Қозоғистон).

FSV технологияси интеграция қилинган тизимларнинг таққослаш ва таҳлил қилиш тажриба синов ташкил қилинди.

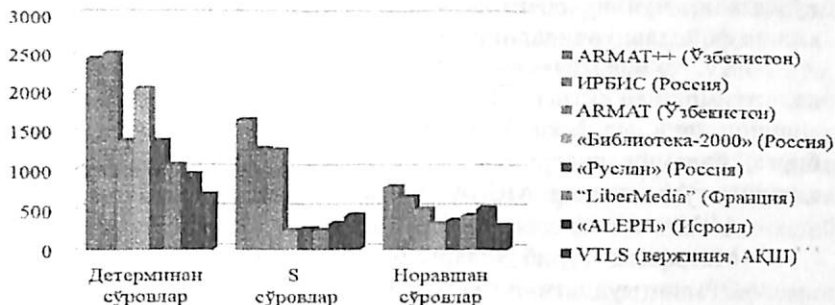
Тажриба синов учун FVS технологияси интеграция қилинган ва энг кўп жорий қилинган дастурий таъминот АРМАТ++ танланди. Унда tag21 форматида маълумотлар шакллан-тирилганлиги учун КАКТларида маълумотларни алмашишнинг ISO2709 стандарти фойдаланиш мумкин. Шунинг учун 50000 та ҳар хил номдаги электрон ресурс ва тўлиқ мантлар танлаб олинди. Таққослаш ва таҳлил натижалари учун «Тезлик», «Топилган маълумотлар сони», «Долзарбликнинг ўрта қиймати» каби параметрлар танланди. КАКТларида ушбу параметрлар маълумотларни излаш модулини таққослаш етарли деб ўйлаймиз.

КАКТларида маълумотларни излаш ва қайта ишлашининг «Тезлик» параметри бўйича натижалари қуйидаги 4.11-расмда келтирилган.



4.11-расм. Маълумотларни излаш ва қайта ишлашининг «Тезлик» параметри бўйича натижалари

«КАКТларида маълумотларни излаш ва қайта ишлашининг «Топилган маълумотлар сони» параметри бўйича натижалари қуйидаги 4.12-расмда келтирилган.



4.12-расм. Маълумотларни излаш ва қайта ишлашнинг «Топилган маълумотлар сони» параметри бўйича натижалари «КАКТларида маълумотларни излаш ва қайта ишлашнинг «Долзарблиликнинг ўрта қиймати» параметри бўйича натижалари куйидаги 4.13-расмда келтирилган.



4.13-расм. Маълумотларни излаш ва қайта ишлашнинг «Долзарбликнинг ўрта қиймати» параметри бўйича натижалари

Тажриба синов натижалари асосида олинган натижалар кўрсатдики, танланган КАКТларининг ичида FSV технологияси жорий қилинган АРМАТ++ дастурий таъминоти тезлик параметри 6% пасайгани, топилган маълумотлар сони бўйича 10% ошганлиги ва экспертлар томонидан берилган долзарбликнинг ўрта қиймати бўйича 25% ошганлигини кўрсатди. Тезликнинг юқорилиги, биринчидан, фойдаланувчиларга ортиқча маълумот такдим қилади, иккинчидан, ахборотга бўлган эҳтиёжларни кондирити мушкул, яъни маълумотлар тўпламини таҳлил қилиши керак. Бу жараён кўшимча вақт талаб қилади. Кўп вақт сарфлаб бўлса ҳам, керакли маълумотларни излаб топиш МИТларнинг самарали эканлигини

кўрсатади, чунки, ортиқча маълумотлардан чеклаш ва таҳлил қилиш фойдаланувчиларнинг вақтларини тежайди.

FSV технологиясини жорий қилиш асосида, биринчи навбатда, ижтимоий ва иқтисодий самарадорликка эришиш мумкин. Бунга эришиш учун эса фақат FSV технологиясини эмас, балки, у билан бирга, ёрдамчи инструментал дастурий модулларни ҳам жорий қилишга тўғри келади. ARMAТ++ дастурий комплексида:

- АРМга масофадан туриб рўйхатдан ўтиш.
- Масофадан туриб ЭРларга буюртма бериш.
- ЭРнинг муддатини узайтириш.
- Интернет ЭРларидан эмас, АРМ ЭРларидан фойдаланиш.
- Бир вақтда бир нечта ЭРни бериш ва қабул қилиш.
- Ягона йиғма электрон каталогдан фойдаланиш.
- Кутубхоналараро ЭРларни электрон олиб бериш.
- ЭРларни бир тизимдан иккинчи тизимга ISO2709 стандарти асосида амалга ошириш.

- Китобхонларни қарздорлиги ҳақида хабардор қилиш.
- Кутубхона ходимлари учун доимий ҳисоботларни ҳосил қилиш.

- Статик ва динамик ҳисоботлар яратиш.
- Mobile қурилмаларидан фойдаланиш.

Ёрдамчи инструментал дастурий модулларнинг ҳар бирида FSV техногиясининг имкониятларидан фойдаланилади.

FSV технологиясининг ижтимоий самарадорлиги қуйидагилардан иборат:

1. Инсоннинг вақтини тежаш.
2. Кутубхоналарда ишончли ЭРларни ўқиш.
3. Кутубхона фаолиятини фаоллаштириш.
4. Китобхонларнинг китобга бўлган эҳтиёжларини қондириш.

5. Ягона йиғма электрон каталог яратиш.
6. Китобдан фойдаланиш даражасини ошириш.
7. Дунё стандарти асосида ЭРларни алмашиш.
8. Mobile қурилмасидан максимал фойдаланиш.
9. ISO2709 стандарти асосида маълумот алмашиш.

10. Китобхонларга китобхонлик қарздорлиги олдини олиш ёки муддатни узайтириш имкониятлари бериш.

11. Кутубхоналараро ЭРларни электрон айирбошлаш хизмати.

12. Кутубхона учун керакли статик ва динамик ҳисоботлар тайёрлаш.

FSV технологиясининг иктисодий самарадорлигини ҳисоблаш битта муассаса учун куйидаги омиллар асосида амалга оширилади:

- Бир кутубхоннинг обуна бўлиши учун минимум 2400 сўм, махсус буюртма бериши минимум 2400 сўм, китобхоннинг китоб муддатини узайтириш учун 2400 сўм йўл кираси ишлатади. Бир китобхон 1 марта обуна бўлиши, бир ўқув йилида ўртача 30 марта махсус буюртма берса ва ўртача 10 марта китоб муддатини ўзайтирса, жами $T_1=2400+2400*30+2400*10=98400$ сўм иктисодий харажат қилган бўлади.

- Шунингдек, курс иши, реферат, мустақил таълим, дарсларга тайёрланиш учун бир ўқув йилда Интернетдан ўртача $T_2=24000$ сўмлик хизматдан фойдаланиш ўрнига кутубхонадан фойдаланиш.

- Бир кутубхоначи битта китобни бериш ёки қабул қилиш учун ўртача 5 минут вақти кетади. АрMAT++ тизими ёрдамида бир кутубхоначи битта китобни бериш ёки қабул қилиш учун ўртача 1 минут вақти кетади. Буларнинг нисбатидан битта китобни бериш ёки қабул қилиш 5 марта вақтни тежаш имконияти пайдо бўлади ёки 5 та кутубхоначининг ишини биттаси қилади. Қолган 4 та кутубхоначи ЭРларни каталоглаштириш билан шуғулланади, бир кунда ўртача 10 та библиографик ёзувни яратади. Битта библиографик ёзувни яратиш дунё тажрибасида 8 евро ($8*3500$ сўм) туради. Бир кунда 9 та ходим $n=4*10*8*3500$ сўмлик кўшимча иш қилади. Бир ўқув йилда ўртача 250 кун хизмат кўрсатилса, $T_3=250*n=280\ 000\ 000$ сўм бўлади.

- Ягона йиғма электрон каталогдан фойдаланиш орқали муассаса бир ўқув йилида ўртача 500 библиографик ёзувни ўзига нухаласа, $T_4=500*8*3500=14\ 000\ 000$ сўм иктисод қилинади.

ARMAT++ бир ўқув йилидаги иктисодий самарадорлигини T билан белгиласак, 4 та фактор ҳисобига $T=T_1+T_2+T_3+E_4$ га тенг, яъни $T=294122400$ сўмни ташкил қилади. Бунда FSV технологиясининг 20% хизмати бор деб ҳисобласак, унда FSV технологиясининг бир ўқув йилидаги иктисодий самарадорлиги ўртача 58824480 сўмни ташкил қилади.

Бугунги кунда жамиятимизда ARMAT++, МБИ ва МД МБ дастурий модуллари оммалашиб, амалиётда жорий қилинмоқда.

Ихтиёрий дастурий комплекс яратилгандан сўнг маълумотларни йиғади, маълумотлар йиғилгандан сўнг маълумотларни излаш ва қайта ишлашни самарали ташкил қилиш керак. Бунинг учун FSV технологиясини дастурий комплексини жорий қилиш орқали унинг фойдаланиш даврини ошириш мумкин.

Тўртинчи боб бўйича хулосалар

Бобда корпоратив ахборот муҳитларида маълумотларни излаш ва қайта ишлаш учун маълумотларни излаш тизимини яратувчи дастурий таъминот – FSV технологиясини ишлаб чиқиш ва жорий қилишга бағишланган бўлиб қуйидаги хулосаларга олиб келади:

1. Корпоратив ахборот муҳитларида маълумотларни излаш ва қайта ишлаш учун маълумотларни излаш тизимини яратувчи дастурий таъминот – FSV технологияси, архитектураси яратилган. FSV технологияси, архитектураси кенг камровли дастурий таъминотларда маълумотларни излаш ва қайта ишлаш учун маълумотларни излаш тизимининг дастурий модулини ишлаб чиқиш масалаларини тадқиқ қилиш ва амалиётга жорий қилишда хизмат қилади.

2. FSV технологиясининг архитектураси асосида корпоратив тармоқда сўровларни шакллантириш усуллари, алгоритмлари ва инструментал дастурий модулининг функционал тузилмаси, IDEF0 моделлари ишлаб чиқилган. Сўровларни шакллантириш усуллари, алгоритмлари ва инструментал дастурий модулини инсон фикрини аниқлаш, таҳлил қилиш, табиий тилда баён қилинган маълумотларга ишлов бериш масалаларида фойдаланилади.

3. Корпоратив тармоқда тақдим қилиш усул ва воситалари, маълумотларни излаш ва тақдим қилиш инструментал дастурий модулнинг функционал тузилмаси ва IDEF0 моделлари ишлаб чиқилган. Тақдим қилиш усул ва воситалари, инструментал дастурий модулнинг функционал тузилмаси ва IDEF0 моделлари асосида турли муҳитли ахборот тизимларида маълумотларни тақдим қилишда интеллектуал фойдаланувчининг интерфейсларини яратиш, тақдим қилишда чекловларни аниқлаш масалаларини ҳал қилишда хизмат қилади.

4. Дастурий таъминотларга FSV технологиясини жорий қилиш учун маълумотлар базасининг лойиҳаси, жорий қилиш усули ишлаб чиқилган. Дастурий таъминотларнинг ўзаро интеграция масалаларида маълумотлар базасини лойиҳалашда, жорий қилиш усулини ишлаб чиқишда фойдаланилади.

5. FSV технологиясининг самарадорлигини аниқлаш мақсадида ISO2907 стандартини қўллаб-қувватловчи дастурий таъминотлар асосида тажриба синови ишлари ўтказилди. Натижага кўра, FSV технологияси жорий қилинган КАКТ АРМАТ++ бўйича «Топилган маълумотлар сони» 5%, «Долзарбликнинг ўрта қиймати» 10% ошган.

ХУЛОСА

«Норавшан ва стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари» мавзудаги иш бўйича асосий хулоса ва натижалар куйидагилардан иборат:

1. Корпоратив ахборот муҳитларида маълумотларни интеллектуал излаш ва қайта ишлаш рекурент муносабатли саклаш усули такомиллаштирилган, дастурий модули ишлаб чиқилган. Корпоратив ахборот муҳитининг турли хусусиятли ва тузилмали, катта ҳажмли ресурсларда маълумотларни излаш ва қайта ишлаш масалаларида ахборот-ресурсларидан фойдаланиш моделлари, усул ва технологияларини ривожлантириш истикболларини аниқлаб беришда, маълумотларни киритиш, таҳрирлаш ва саклаш усули ҳамда дастурий модули асосий элемент сифатида хизмат қилади.

2. Корпоратив тармоқда кирувчи ва чиқувчи ҳаволалар асосида электрон ресурснинг дастлабки рейтингини ҳисоблаш усули таклиф қилинган. Электрон ресурснинг рейтингини ҳисоблаш самарадорлигини 50% га оширган. Ушбу математик усулдан электрон ресурслар тўпламининг импакт фактори ва рейтингини ҳисоблаш масалаларида фойдаланиш имконини беради.

3. Маълумотларни излаш тизимларини баҳолаш учун аниқлик ва тўлалик мезонлари, сарфланадиган вақтни камайтириш учун мантиқий семантик излашнинг усули ишлаб чиқилган ва маълумотларни излаш самарадорлигини 5% га оширган. Мезонлар ва математик усул асосида кўп тармоқли ҳамда босқичли ахборот-ресурсларида маълумотларни излаш ва қайта ишлаш учун сўров ва олдиндан аниқланган ЭР тўпламларининг индекслаш, уларнинг семантик алоқаларни яратиш, излаш тезлигини оптималлаштириш масалаларини тадқиқ қилишга хизмат қилади.

4. Маълумотларни интеллектуал излаш ва қайта ишлаш усулларида тегишлилик функцияларнинг синфларига мос норавшан термларни танлаш, параметрлари орасидаги ўзаро муносабатлари, параметрик тегишлилик функцияларини лойиҳалаштириш алгоритмлари ва дастурий модули ишлаб чиқилган. Норавшан сўровларни қайта ишлаш самарадорлигини 22%га оширган. Бу турли тузилмали ахборот-ресурслари, тизимлари учун билимлар базасининг норавшан термларга асосланган ядро яратиш имконини беради.

5. Корпоратив ахборот-кутубхона тизимларининг норавшан ахборот муҳитида маълумотларни излаш тизимлари учун семантик ядро модели ва семантик ядро яратиш бўйича 3 та усул,

алгоритмлар ва дастурий модули ишлаб чиқилган ҳамда сўровга мос маълумотлар сонини 10% мартага оширган. Катта ҳажмли ахборот-ресурсларида маълумотларни интеллектуал излаш ва қайта ишлаш, модификациялаш учун хос семантик ядро яратиш, ривожлантириш ва бошқариш масалаларини тадқиқ қилишда хизмат қилади.

6. Маълумотларни излаш тизимида норавшан билимлар базаси ва қоидаларини ишлаб чиқиш механизми, алгоритми ва дастурий модули ишлаб чиқилган ҳамда билимлар базасининг ядроси учун IDEF1x модели лойиҳалаштирилган. Норавшан билимлар базаси ва қоидалари маълумотларни интеллектуал излаш ва қайта ишлаш самарадорлигини 20%га ошириш излаш натижасида тақдим қилинган маълумотларни таҳлил қилиш вақтини тежаш имконини беради.

7. Корпоратив ахборот-кутубхона тизимларида маълумотларни излаш тизимларининг дастурий модуллари учун IDEF0 ва IDEF1x модели ишлаб чиқилган. Матнларни таҳлил қилиш, қайта ишлаш инструментал дастурий модулининг функционал имкониятини лойиҳалаштириш ва IDEF моделларини ишлаб чиқиш масалаларида, индекслаштириш, сиқиш, техник воситаларни танлашга оид тадқиқотларни амалга оширишга хизмат қилади.

8. Корпоратив ахборот-кутубхона тизимларида маълумотларни излаш ва қайта ишлаш учун маълумотларни излаш тизими яратувчи дастурий таъминот – FSV технологияси, архитектураси яратилди. FSV технологияси жорий қилинган корпоратив ахборот-кутубхона тизими АРМАТ++ бўйича «Топилган маълумотлар сони» – 7%, «Долзарбликнинг ўрта қиймати» – 10%, «Излаш» самарадорлигини 5% га ошириш сифатли хизмат кўрсатиш имконини беради.

Фойдаланилган адабиётлар рўйхати

1. Altingovde I.S., Demir E.Can F. Incremental Cluster-Based Retrieval Using Compressed Cluster-Skipping Inverted Files. *ACM Transactions on Information Systems*. Jun2008, Vol.26 Issue 3. P.1-36.
2. Anagnostopoulos A., Andrei Z., Kunal P. Effective and efficient classification on a search-engine model. 2006. In *Proc. CIKM*. P.208-217. ACM Press. DOI: doi.acm.org/10.1145/1183614.1183648.
3. Arlitsch K., Obrien P., Rossmann B. Managing Search Engine Optimization: An Introduction for Library Administrators. *Journal of Library Administration*. Feb2013, Vol. 53 Issue 2-3. P.177-188. DOI: 10.1080/01930826.2013.853499.
4. Azevedo C., Iacob M., Almeida J., Van S., Pires L., Guizzardi G. Modeling resources and capabilities in enterprise architecture: A well-founded ontology-based proposal for ArchiMate. *Information Systems*. Dec2015, Vol.54. P.235-262. DOI: 10.1016/j.is.2014.04.008.
5. Back G., Bailey A. Hacking Summon 2.0 The Elegant Way. *Code4Lib Journal*. 2014, Issue 26. P.1-8.
6. Bhat N.A., Ganaie Sh.A. E-resources: use and search strategies adopted by users of Dr.Y.S.Parmar University of Horticulture and Forestry. *Collection Building*. 2016, Vol. 35 Issue 1. P.16-21. DOI: 10.1108/CB-08-2015-0015.
7. Bourne C.P., Donald F.F. A study of methods for systematically abbreviating English words and names. 1961. *JACM* 8 (4): P.538-552. DOI: doi.acm.org/10.1145/321088.321094.
8. Brin S., Page L. The anatomy of a large-scale hyper textual Web search engine. In: *Proceedings of the 7th International World Wide Web Conference, Computer Networks* 30(1-7). 1998. P.107-117.
9. Brock E.O. Declarative semantics of transactions in ORM. *Information Systems*. Aug2016, Vol.60. P.85-94. DOI: 10.1016/j.is.2016.03.005.
10. Cai F., Rijke M. Learning from homologous queries and semantically related terms for query auto completion. *Information Processing & Management*. Jul2016, Vol.52. Issue 4. P.628-643. DOI:10.1016/j.ipm.2015.12.008.
11. Cheang B., Chu S. OR/MS journals evaluation based on a refined PageRank method: an updated and more comprehensive review. *Scientometrics*. Aug2014, Vol.100. Issue 2. P.339-361. DOI: 10.1007/s11192-014-1272-0.

12. Cleary J.G., Witten I.H. "Data Compression Using Adaptive Coding and Partial String Matching," IEEE Transactions on Communications COM-32(4): April 2005. P.396-402.

13. Cleverdon C.W. The significance of the Cranfield tests on index languages. In Proc ACM Press. SIGIR, 1991. P.3-12.

14. Daoud C.M. Fast top-k preserving query processing using two-tier indexes / Daoud C.M. and all. Information Processing & Management. Sep2016, Vol.52 Issue 5. P.855-872. DOI: 10.1016/j.ipm.2016.03.005

15. Datta A., VanderMeer D., Ramamritham K. Parallel Star Join + DataIndexes: Efficient Query Processing in Data Warehouses and OLAP. IEEE Transactions on Knowledge & Data Engineering. Nov/Dec2002, Vol.14. Issue 6. P.1299-1317.

16. David A. Google Semantic Search: Search Engine Optimization (SEO). July 10, 2013. P.1054.

17. Day W.H., Herbert E. Efficient algorithms for agglomerative hierarchical clustering methods. 1984. Journal of Classification. P.1-24.

18. Demigha S. Mining Knowledge of the Patient Record: "The Bayesian Classification to Predict and Detect Anomalies in Breast Cancer". Electronic Journal of Knowledge Management. Aug2016, Vol.14 Issue 3. P.128-139.

19. Ester N., Alistair M., Justin Z. Fast on-line index construction by geometric partitioning. In Proc. CIKM, 2005. P.776-783. ACM Press. DOI: doi.acm.org/10.1145/1099554.1099739.

20. Fenichel C. Editing the permuterm subject index. In North, Jeanne B., Ed. Proceedings Of The American Society For Information Science. Volume 8. 34th Annual Meeting, Denver, November 7 To 11, 2005. Greenwood Press, Westport, Connecticut. P.349-353.

21. Garfield E. The Permuterm Subject Index: An Autobiographical Review. Journal of the American Society for Information Science. Sep-Oct1976, Vol.27. Issue 5/6. P.288-291.

22. Gray J., Chaudhuri S., Bosworth A. Data Cube: A Relational Aggregation Operator Generalizing Group-By, Cross-Tab and Sub-Totals // Data Mining and Knowledge Discovery. 1997. № 1. P.29-53.

23. Haider J. The structuring of information through search: sorting waste with Google. Aslib Journal of Information Management. 2016. Vol.68 Issue 4. P.390-406. DOI: 10.1108/AJIM-12-2015-0189.

24. Hakimov M.X., Gaynazarov S.M. Berilganlar bazasini boshqarish tizimlari. –T.: Fan va texnologiya. 2013. 648 b.

25. Harman D., Ricardo B., Edward F., Lee W. Inverted files. In Frakes and Baeza-Yates (1992), 1992. P.28-43.
26. Judith P., Westport C. Information storage and retrieval. Essential Guide to Multifunction Optical Storage, 1991. P.134.
27. Kalnis P., Papadias D. Multi-query optimization for on-line analytical processing. Information Systems. Jul2003, Vol.28 Issue 5. P.457-463. DOI:10.1016/S0306-4379(02)00026-1
28. Karimov U.F., Mo'minov B.B., Karimov O'U. Axborot-kutubxona tizimlari uchun yangi dasturiy ta'minot // Infocom.uz – информационные технологии Узбекистана. 2015. №8. 65-67 б.
29. Karimov U.F., Mo'minov B.B., Karimov O'U. Korporativ tarmoqda ochiq resurslar elektron kutubxonasini yaratish usullari // Infocom.uz – информационные технологии Узбекистана. 2015. №8. 40-43 б.
30. Kukich K. Techniques for automatically correcting words in text. 1992. ACM Computing Surveys 24 (4): P.377-439. DOI: doi.acm.org/10.1145/146370.146380.
31. Lee S., Jang W., Lee E., Oh S. Search engine optimization. Library Hi Tech. 2016, Vol.34 Issue 2. P.197-206. DOI: 10.1108/LHT-02-2016-0014
32. Levkowitz H., Introduction to information retrieval. Choice: Current Reviews for Academic Libraries. 2009. Vol.46 Issue 5. P.943.
33. Massis B. The digital librarian. New Library World. 2014, Vol. 115 Issue 9/10. P.496-499. DOI: 10.1108/NLW-11-2013-0088
34. Mo'minov B.B. Informatika. –T.: 2014. 344 b.
35. Mo'minov B.B. Ma'lumotlarni izlash modullarida mantiqiy semantik izlash modeli. "Актуальные проблемы прикладной математики и информационных технологий – аль-Хорезми 2016": Труды международной конференции (9-10/11/2016г., Ташкент). 94-97 б.
36. Mo'minov B.B. Ma'lumotlarni izlash uchun elektron resursning reytingini hisoblash modeli. "Актуальные проблемы прикладной математики и информационных технологий – аль-Хорезми 2016": Труды международной конференции (9-10/11/2016г.). 91-94 б.
37. Mo'minov B.B. Ma'lumotlarni intellektual izlashda indekslash, siqish algoritmlari va usullari // «Хисоблаш ва амалий математика муаммолари». –Т. 2017. №1(7). 92-105 б.
38. Mooers C. From a point of view of mathematical etc. techniques. In R.A.Fairthorneed. Towards information retrieval, 1991.

39. Moreno L., Martinez P. Overlapping factors in search engine optimization and web accessibility. Online Information Review. 2013, Vol. 37 Issue 4. P.564-580. DOI: 10.1108/OIR-04-2012-0063.

40. Muminov B.B. Development of a virtual exchange platform results qualifying master's works with the possibility of sharing the achievements and search. The Strategies of Modern Science Development: Proceedings of the IX International scientific-practical conference. North Charleston, SC, USA, 12-13 October 2016. North Charleston: CreateSpace, 2016. P.17-21.

41. Muminov B.B. Intelligent search for corporate electronic resources. Международной конференции «Перспективы развития информационных технологий ИТРА-2014» Ташкент, 4-5/11/2014г.

42. Muminov B.B. Main classes of membership functions in intellectual search // Journal of Computer Science Engineering and Information Technology Research (JCSEITR) Vol.6. 2016. P.1-10.

43. Muminov B.B. Methods of information retrieval and cash algorithms. Transactions of the international scientific conference "Perspectives for the development of information technologies ITPA – 2014" 4-5 November, 2014. P.194-196.

44. Muminov B.B. Stages of designing data mining // "TATU xabarlari" журналы, 2014. 2-сон. 116-119 б.

45. Muminov B.B. The calculating rating of electronic resources // European Sciences review. Scientific journal № 7-8 2016. P.20-21.

46. Muminov B.B. The mathematic models of searching data in the centers of information resource // IJRET: International Journal of Research in Engineering and Technology. Volume: 05 Issue: 10 | Oct-2016. P.43-46.

47. Muminov B.B. The model of logical semantic searching. Proceedings of the uzbek-japan symposium on ecotechnologies "Innovation for sustainability-harmonizing science technology and economic development with human and natural environment" Tashkent, 14 may 2016. P.72-77.

48. Muminov B.B., Dadenova G.K., Ostonov F.U. Information search. Международная конференция «Перспективы развития информационных технологий ИТРА-2014» Ташкент, 4-5/11/2014.

49. Muminov B.B., Hamroyev E.Z. 3D model retrieval system. Международной конференции «Перспективы развития информационных технологий ИТРА-2014» Ташкент, 4/11/2014. С.324-326.

50. Muminov B.B., Tajiyev J.A. Information processing and signals on dual-core processor. Transactions of the international

scientific conference "Perspectives for the development of information technologies ITPA-2014" 4-5 November, 2014. P.196-199.

51. Muminov B.B., Tuykulov N. A. Changing the achievements of GEW and EW the platform problems which gives availability of searching information. The Strategies of Modern Science Development: Proceedings of the XII International scientific-practical conference. North Charleston, SC, USA, 4-5 April 2017. North Charleston: CreateSpace, 2017. P.15-18.

52. Newquist H. P. Data Mining: The AI Metamorphosis // Database Programming and Design. 1996. № 9. P.115-126.

53. Niemi T., Hirvonen L., Järvelin K. Multidimensional Data Model and Query Language for Informetrics. Journal of the American Society for Information Science & Technology. Aug2003, Vol.54. p939.

54. Oppenheim Ch., Haendler H. Information storage and retrieval. Journal of Academic Librarianship. Jan1994, Vol.19 Issue 6. P.405-407.

55. Ozmutlu H.C., Çavdur F. Application of automatic topic identification on Excite Web search engine data logs. Information Processing & Management. Sep2005, Vol.41 Issue 5. P.1243-1262. DOI: 10.1016/j.ipm.2004.04.018.

56. Pearl J. Probabilistic Reasoning in Expert Systems: Networks of Plausible Inference. San Francisco: Morgan Kaufmann. 1988. P.552.

57. Pertesis D., Doulkeridis Ch. Efficient skyline query processing in SpatialHadoop. information Systems. Dec2015, Vol.54. P.325-335. DOI: 10.1016/j.is.2014.10.003

58. Rakhmatullaev M.A., Rakhmatullaev I.M. Algorithms for constructing semantic relationships between metadata of digital archives // Материалы научно-технической конференции «Современное состояние перспективы применения информационных технологий в управлении». 7-8 сентября. 2015. 65-70 с.

59. Rakhmatullaev M.A., Muminov B.B. System of intelligence retrieval for corporate academical information network // IJRET: International Journal of Research in Engineering and Technology. Volume: 05 Issue: 09 | Sep-2016, P.85-87.

60. Silberschatz A., Henry F., Sudarshan S. Database system concepts / Library of Congress Cataloging-in-Publication Data 6th ed. Published by McGraw-Hill, Inc., 1221 Avenue of the Americas, New York, NY 10020. 2011. P.1376.

61. Sparck-Jones K., Walker S., Robertson S.E. A probabilistic model of information retrieval: Development and comparative experiments. Inf. Process. Manag. 36(6), 2000. P.779-808.

62. Sturdy D. Search Engine Optimisation and Automatic Classification. *Legal Information Management*. Spring2010, Vol.10 Issue 1. P.24-28. DOI: 10.1017/S1472669610000071.
63. Taniar D. The use of Hints in SQL-Nested query optimization By: Taniar D., Khaw H., Tjioe H., Pardede E. // *Information Sciences*. Jun2007, Vol.177 Issue 12. P.2493-2521. DOI: 10.1016/j.ins.2006.12.015.
64. Thenmalar S., Geetha T.V. Enhanced ontology-based indexing and searching. *Aslib Journal of Information Management*. 2014, Vol.66. Issue 6. P.678-696. DOI:10.1108/AJIM-08-2014-0098.
65. Vechtomoova O. Introduction to Information Retrieval. *Computational Linguistics*. Jun2009, Vol.35 Issue 2. P.307-309.
66. Wang G., Chan Ch. Efficient processing of enumerative set-based queries. *Information Systems*. Jan2016, Vol.55. P.54-72. DOI: 10.1016/j.is.2015.08.005.
67. Wessel K. Variations on language modeling for information retrieval. –Enschede: *Neslia Paniculata*. 2004. P.273.
68. Witten, I., Nevill, C. G. Indexing and compressing full-text databases for CD-ROM. *Journal of Information Science*. 1991, Vol.17 Issue 5, P.265-272.
69. Yan E. Topic-based Pagerank: toward a topic-level scientific evaluation. *Scientometrics*. Aug2014, Vol.100 Issue 2, P.407-437. DOI:10.1007/s11192-014-1308-5
70. Zadeh L.A. The concept of a linguistic variable and its application to approximate reasoning. *American Elsevier Publishing Company, New York*, 1973. P.167.
71. Алиев Р.А., Алиев Р.Р. Теория интеллектуальных систем и ее применение. –Баку, “Изд-во Чашыюглы”, 2001. – 700 с.
72. Алоев Р.Д., Мухамедиева Д.Т. Норавахан ахборотни қайта ишлаш асосида талабалар ўқув жараёнини муқобиллаштириш ва ўзлаштиришини башорат қилиш моделлари. –Т.: 2010. –282 б.
73. Бекмурадов Т.Ф. Систематизация задач интеллектуальных систем поддержки принятия решений // *Проблемы информатики и энергетики*. –Т.: 2003. № 4. –С.24-35.
74. Бекмурадов Т.Ф., Дадабаева Р.А., Мухамедиева Д.Т. Принятие решений в нечеткой среде // *Научный журнал «Проблемы информатики»*. –Новосибирск. Институт вычислительной математики и математической геофизики СО РАН. 2010. Вып.1. –С.52-61.
75. Борисов А.Н. Модели принятия решений на основе логико-лингвистической переменной. –Рига: Зинатне, 1982. 256 с.

76. Галахов И. Проектирование корпоративной информационно-аналитической системы. Открытые системы. № 4, 2003. –С.27-32
77. Голицына О.Л. Основы алгоритмизации и программирования // О.Л.Голицына, И.И.Попов. –М.: Форум, 2008. –432 с.
78. Грубер М. Понимание SQL. –М.: Мир, 1993. –481 с.
79. Дадажанов Т.К., Нусратов Т.С., Рахматуллаев М.А. Нечеткие алгоритмические модели соответствий в планирование решений. Сборник Республиканской научно-технической конференции «Современное состояние и перспективы применения информационных технологий в управлении» Джизак, 5-6/9/2016. 383-386 б.
80. Дональд К. Искусство программирования, том 3. Сортировка и поиск. The Art of Computer Programming, vol.3. Sorting and Searching. 2-е изд. –М.: Вильямс, 2007. –824 с.
81. Кабулов В.К. Алгоритмизация в социально-экономических системах. –Т.: Фан. 1989. –264 с.
82. Каримов У.Ф., Мўминов Б.Б., Каримов Ў.У. Корпоратив тармоқда очик ресурслар электрон кутубхонасини яратиш усуллари. Электрон кутубхона тармоқларида илмий-таълимий ахборотлар яратиш ва улардан фойдаланиш технологиялари. –Т.: “E_LINE PRESS”, 2016. –48-58 б.
83. Каримов У.Ф., Каримов Ў.У. Миллий корпоратив электрон кутубхона шакллантириш истиқболлари: муаммолар ва ечимлар // Электрон кутубхона тармоқларида илмий-таълимий ахборотлар яратиш ва улардан фойдаланиш технологиялари. –Т.: ООО “E-LINE PRESS”, 2017. 41-51 б.
84. Кулинкович Т.О. Основы научного цитирования. –Минск: БГУ, 2010. –58 с.
85. Мўминов Б.Б. e-AR тизимида библиографик ёзувларни яратиш ва тасвирлаш. ТАТУ, “Ахборот ва телекоммуникация технологиялари муаммолари” мавзуйдаги республика илмий-техник анжуман материаллари. Тошкент, 2016. 32-35 б.
86. Мўминов Б.Б. IS-ICT ПЛАТФОРМАСИ. Иқтисодийнинг реал тармоқларини инновацион ривожланишида ахборот-коммуникация технологияларининг аҳамияти // Республика илмий-техник анжумани тўпламлари. 1-қисм. –Т.: ТАТУ. 323-325 б.
87. Мўминов Б.Б. Ахборот муҳитларини классификациялаш // “ТАТУ хабарлари” журнали, 2016. №1. 54-61 б.
88. Мўминов Б.Б. Ахборот-ресурс марказларининг корпоратив тармоғида излаш модулининг элементлари. Сборник Республиканской научно-технической конференции «Современное состоя-

ние и перспективы применения информационных технологий в управлении» Джизак, 5-6 сентября 2016. -242-247 б.

89. Мўминов Б.Б. Корпоратив тармоқда сўровларни қайта ишлаш инструментал дастурий воситанинг функционал тузилмаси. Иктисодиётнинг реал тармоқларини инновацион ривожланишида ахборот-коммуникация технологияларининг аҳамияти / Республика илмий-техник анжуманининг тўпламлари. 2-қисм. –Т.:2017. 71-74 б.

90. Мўминов Б.Б. Маълумотларни излаш тизими. –Т.: Фан ва технология, 2016. 210 б.

91. Мўминов Б.Б. Маълумотларни излаш усуллари. –Т.: Фан ва технология, 2016. 276 б.

92. Мўминов Б.Б. Маълумотларни интеллектуал излаш ва қайта ишлашнинг тарихи ва ривожланиши // “ТАТУ хабарлари” журнали. 2015, №3. 48-52 б.

93. Мўминов Б.Б. Норавадан сўровларни қайта ишлаш. ТАТУ, “Radiotexnika, telekommunikatsiya va axborot texnologiyalari: muammolari va kelajak rivoji” Xalqaro ilmiy-texnik konferensiya maqolalar to‘plami. –Т.: 2015. 27-31 б.

94. Мўминов Б.Б. Параметрли тегишлилик функциясини лойиҳалаш. Материалы научно-технической конференции «Современное состояние перспективы применения информационных технологий в управлении». Ташкент, 7-8/09/2015. С.172-175.

95. Мўминов Б.Б. Реализация этапов поиска данных в нечетких информационных средах. ТАТУ, “Ахборот ва телекоммуникация технологиялари муаммолари” Илмий-техник конференциянинг маърузалар тўплами, 1-қисм. Тошкент, 2015. 116-118 б.

96. Муминов Б.Б. Формирование запросов для интеллектуального поиска в нечеткой информационной среде // International Scientific Journal SCIENCE and World, №3(31), 2016, Vol.1. P.79-85.

97. Мўминов Б.Б. Электрон ахборот-ресурсларида мантикий сўровларни қайта излаш. Материалы Республиканская научная конференция с участием зарубежных ученых "Математическая физика и родственные проблемы современного анализа". Бухара, 26-27/11/2015, С.367-369.

98. Мўминов Б.Б. Электрон ахборот таълим ресурслари маълумотларни излаш. // «Узлуксиз таълим сифат ва самарадорлигини оширишнинг назарий-услугий муаммолари» мавзуидаги илмий конференция. Самарқанд. 2015. 25-27б.

99. Муминов Б.Б., Абидова Ш.Б. Запрос и векторная модель. Актуальные научные исследования в современном мире. X

международная научная конференция, iScience, Переяслав-хмельницкий, Украина, 23-24 февраля 2016. С.103-104.

100. Муминов Б.Б., Гапунова А.А. Обработка нечетких запросов. Актуальные научные исследования в современном мире., X международная научная конференция, iScience, Переяслав-хмельницкий, Украина, 23-24 февраля 2016. С.105-109.

101. Муминов Б.Б., Рахматуллаев М.А., Корпоратив тармоқда фан ва таълимга оид маълумотларни интеллектуал излаш модели ва воситаси // Электрон кутубхона тармоқларида илмий-таълимий ахборотлар яратиш ва улардан фойдаланиш технологиялари. –Т.: ООО “E-LINE PRESS”, 2017. -28-35 б.

102. Муминов Б.Б., Тўраев Б.З., Ҳамроев Э.З. Вариантли сўровларни қайта ишлаш. // Материалы Республиканская научная конференция с участием зарубежных ученых "Математическая физика и родственные проблемы современного анализа". Бухара, 2015. -С.363-364.

103. Муминов Б.Б., Ҳамроев Э.З. Ахборотларни излаш усуллари. ТАТУ, “Ахборот ва телекоммуникация технологиялари муаммолари”. Илмий-техник конференциянинг маърузалар тўплами. I-қисм. Тошкент, 2015. -199-201 б.

104. Муминов Б.Б., Ҳамроев Э.З. Матнли файлларни индекслаш ва маълумотларни излаш. // Материалы Республиканская научная конференция с участием зарубежных ученых "Математическая физика и родственные проблемы современного анализа". Бухара, 2015, -С.365-367.

105. Муминов Б.Б., Ширинов З. Маълумотларни излашда долзарблик ва вазни ҳисоблаш усуллари // Бухоро давлат университети илмий ахбороти. Бухоро, 2016. №4. -35-42 б.

106. Мухамедиева Д.Т. Норавадан ахборотни қайта ишлаш асосида суст шаклланган жараёнларни башоратлаш ва муқобиллаштириш моделлари. –Т.: Fan va texnologiya, 2012. -376 б.

107. Острейковский В.А. Информатика. –М.: Высш. шк., 2005.

108. Оутей М., Конте П. Эффективная работа: SQL Server 2000. – СПб.: Питер; Киев: Издательская группа BHV, 2002. -992 с.

109. Пряжинская В.Г. Компьютерное моделирование в управлении водными ресурсами / В.Г.Пряжинская, Д.М.Ярошевский, Л.К.Левит-Гуревич. –М.: 2002. 323 с.

110. Рахматуллаев М.А. Мировые информационные ресурсы. –Т.: Fan va texnologiya, 2017. 196 с.

111. Рахматуллаев М.А., Муминов Б.Б. Модели принятия решений в нечетких информационно-библиотечных средах “TATU xabarları” журналі, 2014. №3. 111-117 б.
112. Рахматуллаев М.А. Повышения детерминированности информационных сред для интеллектуализации информационного поиска // Eighth World Conference on Intelligent Systems for Industrial Automation WCIS-2014, Tashkent, Uzbekistan. 25/11/2014. 329-332 p.
113. Рахматуллаев М.А. Ахборот-кутубхона муассасаларининг автоматлаштирилган тизими (ARMAT). –Т.: ELine Press, 2014. 230б.
114. Рахматуллаев М.А. Корпоратив ахборот-ресурс марказларининг автоматлаштирилган тизими: электрон катологи маълумотлар базасини шакллантириш (KARMAT-U). –Т.: 2012. -113 б.
115. Роб П., К.Коронел. Системы баз данных: проектирование, реализация и управление / Пер.с англ. –СПб.:BHV, 2004. -1040 с.
116. Рыбин В.В. Основы теории нечетких множеств и нечеткой логики. –М.: Изд-во МАИ, 2007. -96 с.
117. Рыжов А.П. Модели поиска информации в нечеткой среде. –М.: Издательство Центра прикладных исследований при механико-математическом факультете МГУ. 2004. -96 с.
118. Стрелковский А.С. Элементы невыпуклой оптимизации. –Новосибирск: Наука, 2003. -356 с.
119. Таха Х.А. Введение в исследование операций. 7-изд.: Пер.с англ. –М.: Вильямс, 2007. -912 с.
120. Трофимова В.В. Информатика. –М.: Юрайт. 2011. -911с.
121. Ухоботов В.И. Избранные главы теории нечетких множеств. Челябинск: Изд-во ЧГУ, 2011. -245 с.
122. Фозилов Ш., Маматов Н.С. Маълумотларга дастлабки ишлов бериш масалалари // Материалы научно-технической конференции «Современное состояние перспективы применения информационных технологий в управлении». 7-8/09/2015. –С. 288-306.
123. Шокин Ю.И. Проблемы поиска информации. –Новосибирск: Наука, 2010. –220 с.
124. Юсупбеков Н.Р. Вычислительный интеллект и его составляющие / Международный научно-технический журнал “Химическая технология. Контроль и управление”. –Т. 2013. №3. –С. -26-31.
125. Юсупбеков Н.Р. Интеллектуальные системы управления и принятия решений. –Т.: «Ўзбекистон миллий энциклопедияси», 2014. -490 с.

ШАРТЛИ БЕЛГИЛАР ВА АТАМАЛАР РЎЙХАТИ

MVC	- Model View Control
DOM	- Document Object Model
FSV	- Formation-Searching-Viewing
HTML	- HyperText Markup Language
IDEF	- Integrated Computer-Aided Manufacturing
ORM	- Object-Relational Mapping
SEO	- Search engine optimization
SMART	- Salton's Magic Automatic Retriever of Text
SQL	- Structured query language
UI	- User interface (фойдаланувчи интерфейси)
XML	- Integration Definition for Function Modeling
ААТ	- Автоматлаштирилган ахборот тизими
АКТ	- Ахборот-кутубхона тизими
АР	- Ахборот ресурси
АРМ	- Ахборот-ресурс маркази
АТ	- Ахборот тизими
ББ	- Билимлар базаси
ДАМ	- Детерминан ахборот муҳити
ИМ	- Излаш модули
КАКТ	- Корпоратив ахборот-кутубхона тизими
КАМ	- Корпоратив ахборот муҳити
ЛЎ	- Лингвистик ўзгарувчи
МБ	- Маълумотлар базаси
МББТ	- Маълумотлар базаларини бошқариш тизими
МЖ	- Маълумотлар жамланмаси
МИҚИ	- Маълумотларни излаш ва қайта ишлаш
МИТ	- Маълумотларни излаш тизими
НАМ	- Норавадан ахборот муҳити
НСАМ	- Норавадан ва стохастик ахборот муҳити
РМБ	- Реляцион маълумотлар базаси
САМ	- Стохастик ахборот муҳити
ЭР	- Электрон ресурс
ЭҲМ	- Электрон ҳисоблаш машинаси

Қисқачча мазмуни.....	3
Tasks and results of the research.....	6
Задачи и результаты исследования.....	8
Кириш.....	10
I-боб. Корпоратив ахборот муҳитларида маълумотларни излаш ва қайта ишлаш усуллари, воситаларининг ҳозирги ҳолати.....	19
1.1. Маълумот излашнинг ривожланиш босқичлари ва фундаментал моделлари, усуллари.....	19
1.2. Корпоратив ахборот-кутубхона тизимлари ва сўровларни қайта ишлаш усуллари.....	30
1.3. Фан ва таълимга оид ахборот муҳитларини классификациялаш.....	35
1.4. Корпоратив ахборот муҳитларида маълумотларни излаш tizimini яратиш муаммолари.....	41
Биринчи боб бўйича хулосалар.....	45
II-боб. Стохастик ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ва дастурий модулини яратиш.....	46
2.1. Стохастик ахборот муҳитларида маълумотларни излаш tiziminинг элементлари ва моделлаштириш омиллари.....	46
2.2. Маълумотларни излаш тизимларида электрон ресурс ва термин частоталарининг вазнини ҳисоблаш усуллари.....	51
2.3. Тўлиқ функционалли маълумотларни излаш тизимларида долзарбликни ҳисоблаш усулларига эвристик ёндашув.....	59
2.4. Маълумотларни излаш тизимларида мантиқий семантик излаш модели.....	63
2.5. Стохастик ахборот муҳитларида маълумотларни излаш tiziminинг дастурий модулини яратиш.....	68
Иккинчи боб бўйича хулосалар.....	74
III. Норавшан ахборот муҳитларида маълумотларни излаш ва қайта ишлаш моделлари, усуллари ва дастурий модулини яратиш.....	76
3.1. Норавшан сўровларни қайта ишлаш усули.....	76
3.2. Параметрли тегишлилик функциясини лойиҳалаш усули..	80
3.3. Маълумотларни интеллектуал излашда тегишлилик функцияларининг синфларини танлаш.....	83

3.4. Фан ва таълимга оид маълумотларни интеллектуал излашда семантик ядро модели ва алгоритмларини ишлаб чиқиш.....	95
3.5. Маълумотларни интеллектуал излаш тизимида норавшан излаш модели ва қоидаларни ишлаб чиқиш.....	99
3.6. Норавшан ахборот муҳитларида маълумотларни излаш тизимининг дастурий модулини яратиш.....	106
Учинчи боб бўйича хулосалар.....	113
IV. Корпоратив тармоқда маълумотларни излаш ва қайта ишлаш дастурий таъминотни яратиш ҳамда жорий қилиш.....	115
4.1. Корпоратив тармоқда маълумотларни излаш ва қайта ишлашнинг FSV технологияси.....	115
4.2. Корпоратив тармоқда сўровларни қайта ишлаш инструментал дастурий модулини яратиш.....	118
4.3. Корпоратив тармоқда маълумотларни излаш ва тақдим қилиш инструментал дастурий модулини яратиш.....	122
4.4. FSV технологиясини жорий қилишда маълумотлар базасини ишлаб чиқиш.....	126
4.5. FSV технологиясини дастурий комплексларга жорий қилиш ва унинг самарадорлиги.....	130
Тўртинчи боб бўйича хулосалар.....	136
Хулоса.....	137
Фойдаланилган адабиётлар рўйхати.....	139
Шартли белгилар ва атамалар рўйхати.....	149

БАХОДИР БОЛТАЕВИЧ МЎМИНОВ

**НОРАВШАН ВА СТОХАСТИК
АХБОРОТ МУҲИТЛАРИДА МАЪЛУМОТЛАРНИ
ИЗЛАШ ВА ҚАЙТА ИШЛАШ МОДЕЛЛАРИ,
УСУЛЛАРИ**

Монография

Мухаррир М.А.Хакимов

Босишга рухсат этилди 24.12.2018й. Бичими 60X84 ¹/₁₆.
Босма табағи 9,5. Шартли босма табағи 12,25. Адади 100 нусха.
Буюртма № 232. Баҳоси келишилган нарҳда.
«Университет» нашриёти. Тошкент, Талабалар шаҳарчаси, ЎзМУ
маъмурий биноси.
Ўзбекистон Миллий университети босмахонасида босилди.
Тошкент, Талабалар шаҳарчаси, ЎзМУ.