КОРПОРАТИВ АХБОРОТ-РЕСУРС МАРКАЗЛАРИДА МАЪЛУМОТЛАРНИ ИНТЕЛЛЕКТУАЛ ИЗЛАШ ТИЗИМИНИНГ FSV ТЕХНОЛОГИЯСИ





Баҳодир Мўминов,
Муҳаммад ал-Хоразмий номидаги Тошкент ахборот технлогиялари университетининг Информатика асослари кафедраси мудири, техника фанлари доктори

Кўп жиҳатли корпоратив ахборот-ресурс марказларининг тармоқларида керакли маълумотларни излаш масаласини сифат жиҳатидан янги даражага олиб чиқишни тақозо қилмоқда. Бу масаланинг ечими электрон ресурсларнинг (ЭР) ўзгарувчанлиги ва уларнинг миқдори доимий ўсиб бориш омилларини мураккаблаштириб бориш билан боғлиқдир.

Корпоратив АРМларнинг асосий электрон ресурслар тизимларда излаш модулларининг самарали ишлаши учун бир нечта базавий босқичлар бўлиши лозим *(1-расм)*.

Google, Yahoo, Bing, Yandex, Rambler каби излаш модуллари (ИМ) 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 – излаш модули (ИМ)нинг модификациялари ҳисобланади.

ИМ бу – реалликнинг маълум бир соддалаштириши бўлиб, унинг асосида математик формула ва ушбу формулани ЭРларга татбиқ этишнинг қоидалари асосида ҳосил қилинадиган дастурий модулдир. Бу формула ва қоидалар тизимга, қандай ЭРни излаш сўровига мос бўлган деб ҳисоблаш ҳамда топилган ЭРларнинг тўпламини қандай даражалаш масаласини ҳал этишга имкон беради. Анъанавий ва замонавий усулларнинг негизида учта асосий ёндашув мавжуд.

Биринчи ёндашув – тўпламлар назариясига асосланади. Унинг ҳар хил кўринишлари сифатида қуйидаги турларга ажратиш мумкин: буль модел (мантиқий модел); кенгайтирилган буль модели.

Иккинчи ёндашув – векторли алгебрага асосланади. Ушбу ёндашувни:

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

Учинчи ёндашув – экспертлар гуруҳи, машинали ўқув, интеллектуал таҳлил, норавшан тўпламлар назарияси ва мантиғи, k қийматли мантиқ, эҳтимоллар каби назарияларга асосланиб: эҳтимоллик модели; норавшан моделлар; семантик боғланишни ҳисобловчи моделларга ажратиш мумкин.

Корпоратив ахборот-ресурс марказларида (КАРМ) маълумотларни интеллектуал излаш тизимининг классик моделлари ЭРларни тақдим этувчи калит сўзлар тўпламларидан фойдаланилади, бундай сўзлар одатда терминлар деб юритилади.

Маълумотларни интеллектуал излаш тизимларининг (МИИТ) модификацияланган ва ишлаб чиқилган математик моделлар, усуллар ва алгоритмлар асосида МИИТни 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.

1-расм

ЭЛЕКТРОН РЕСУРС ТИЗИМЛАРДА ИЗЛАШ МОДУЛЛАРИНИНГ БАЗАВИЙ БОСҚИЧЛАРИ

01 излаш сўровига ишлов бериш

02 излаш натижаларининг сўровга мослиги

03 топилган ЭРларни тўғри ва оқилона даражалаш имконияти

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

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

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

3. (V) – ЭРларни тақдим этишнинг усули (V – viewing).

Бу уч таркибий қисмни бирлаштириб, КАМларида МИТлар учун FSV технологияси (FSV платформаси, FSV Framework) деб номлаймиз *(2-расм)*.

FSV технологияси асосини КАРМларида маълумотларни излаш ва қайта ишлаш учун тадқиқотимиз бўйича яратилган моделлар, усуллар ва алгоритмлар ҳамда дастурий модулларнинг функционал тузилмаси, IDEF моделлар асосида сервер иловали мижоз-сервер архитектурасига асосланган ҳолда архитектурасини ишлаб чиқамиз *(3-расм)*.

Сервер иловали мижоз-сервер архитектураси фойдаланувчи ва сервер иловаларини бошқариш ва қайта ишлаш учун қулай ва асосий имкониятлари қуйидагича:

1. Нозик мижоз;

2. Мижоз ва сервер орасида минимум маълумот берилади;

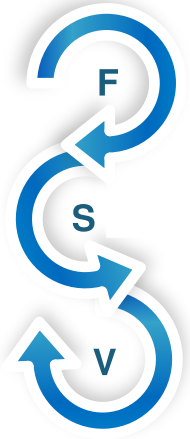
Сервер иловалари бир нечта нусхада, бир нечта компьютерда ишга туширилиши мумкин.

FSV технологияси – корпоратив ахборот муҳитларида МИҚИ моделлари, усуллари ва алгоритмларини интеграция ва модификацияловчи, сервер иловали мижоз-сервер архитектурасига асосланган инструментал дастурий платформа.

FSV технологиясини турли корпоратив тармоқларнинг ахборот тизимларида маълумотларни излаш ва қайта ишлаш учун фойдаланиш мумкин. Бунинг учун ахборот тизимла-

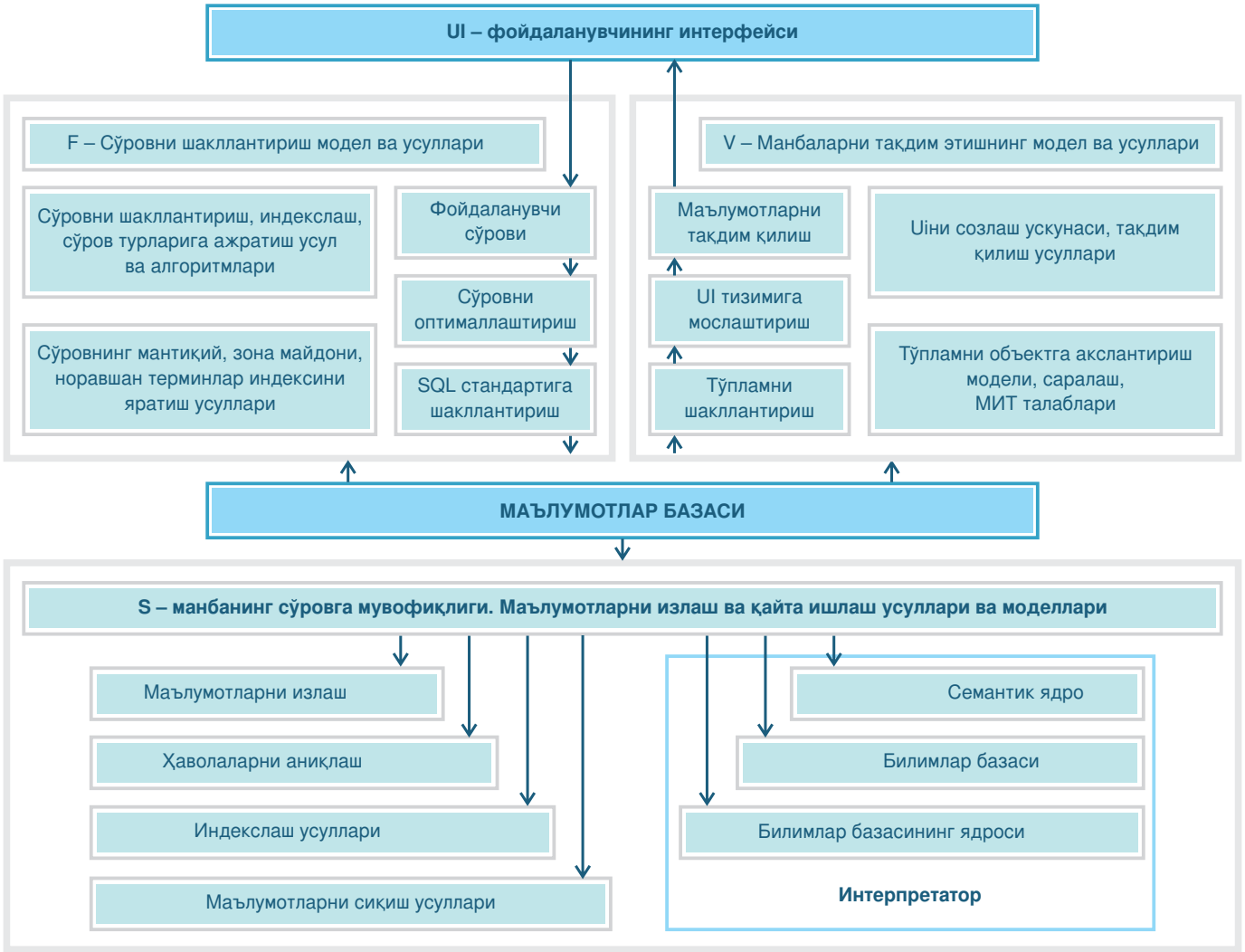
2-расм

FSV ТЕХНОЛОГИЯСИ



3-расм

FSV ТЕХНОЛОГИЯСИНИНГ АРХИТЕКТУРАСИ



4-расм

СЎРОВЛАРНИ ҚАЙТА ИШЛАШ СХЕМАСИ



FSV технологияси – корпоратив ахборот муҳитларида МИҚИ моделлари, усуллари ва алгоритмларини интеграция ва модификацияловчи, сервер иловали мижоз-сервер архитектурасига асосланган инструментал дастурий платформа

ри 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.

4. XML тузилмасини таҳлил қилиш ва МББТга боғлаш мавжуд бўлиши.

5. Дастурлаш тили объектга йўналтирилган.

Ушбу талабларга жавоб берган АТларида FSV технологиясини жорий қилиш ва унинг функционал имкониятларидан МИҚИда фойдаланиш мумкин.

Корпоратив тармоқда сўровларни қайта ишлаш инструментал дастурий модулини яратиш

Сўровларни қайта ишлаш FSV технологиясининг муҳим қисми ҳисобланади ва F – сўровларни шакллантириш деб айтилади. Бу жараённи қуйидаги схема билан тасвирлаймиз.

4-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 $0:fqq®$.

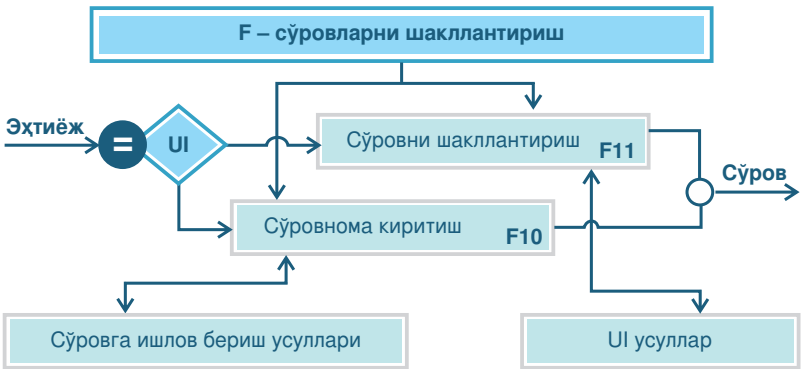
2-босқич. q* нормаллашган сўровни яратиш $*1:fqq®$.

Бу иккала босқич алоҳида инструментал дастурий модул бўлиб, уларни мос равишда f ва f1 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 биланоқ бу усуллар сўровни яратиш учун таъсир кўрсатади.

5-расм

f ОПЕРАТОРИНИНГ IDEF0 МОДЕЛИ



f операторининг IDEF0 моделини қуйидагича ишлаб чиқамиз *(5-расм)*.

f оператори UI га асосланган ақлий ҳужум, синов, тестлаш, масалани ечиш, муаммо, савол-жавоб каби турли фойдаланувчининг ахборот эҳтиёжини аниқлашга мўлжалланган UI усуллар мажмуасидан ҳам фойдаланилади. 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 айтамиз.

f операторининг UI усуллар асосида яратиладиган сўров – сунъий (машина) сўрови деб айтамиз.

f1 оператори эса рекурент кетма-кетликда бўлиб, қўшимча ва ёрдамчи терминлардан тозалаш, сўровнинг турини аниқлаш, зона майдон, мантиқий ва норавшан термларнинг индексларини яратиш орқали амалга оширилади. f1 операторининг IDEF0 моделини келтирамиз.

f1 операторининг алгоритми қуйидагича амалга оширилади:

1 – алгоритм. Рекурент кетма-кетликда сўровни қайта ишлаш.

1-қадам. Табиий тилда ёзилган сўровда қўшимча терминлардан тозалаш.

2-қадам. Сўровни ёрдамчи терминлардан тозалаш.

3-қадам. Сўровда норавшан терм мавжудлигини текшириш.

4-қадам. Агар сўровда норавшан терм бўлса, сўров норавшан сўров ҳисоблансин ва 8-қадамга.

5-қадам. Сўровда мантиқий амал мавжудлигини текшириш.

6-қадам. Агар сўровда мантиқий амал бўлса, сўров мантиқий сўров ҳисоблансин ва 9-қадамга

7-қадам. Сўров сўзли сўров ҳисоблансин ва 10-қадамга

8-қадам. Норавшан термларни индекслаш.

9-қадам. Мантиқий амалларни индекслаш.

10-қадам. Зона майдонларини аниқлаш ва индекслаш.

11-қадам. Сўров семантикасини аниқлаш.

12-қадам. Нормаллаштирилган сўров.

f ва f1 операторлар ёрдамида фойдаланувчининг ахборотга бўлган эҳтиёжини нормаллаштирилган сўровга олиб келинади. Шуни ҳам таъкидлаш лозимки, UI усуллари ёрдамида тайёрланган сўровлар жуда тез оптималлашган сўровга ўтади.

F – сўровларни шакллантириш жараёнида бошланғич маълумотлар, яъни қўшимча терминлар тўплами, ёрдамчи сўзлар тўплами, норавшан термлар тўплами, мантиқий 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 ва сўров терминининг тахминий семантик алоқаси шакллантирилади. Буларнинг асосида FSV технологиясига асосланган МИТларда тайёр олдиндан маълум сўров ва терминлар учун электрон ресурсларнинг семантик алоқаси яратилади. 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 асосан фикрларни излаш усулларига таянади ва буларни 3 гуруҳга ажратиш мумкин:

1. Психологик фикрлашни фаоллаштириш усуллари (ақлий ҳужум, тескари ақлий ҳужум).

2. Фикрни тизимли излаш усуллари (саволлар рўйхати, тест).

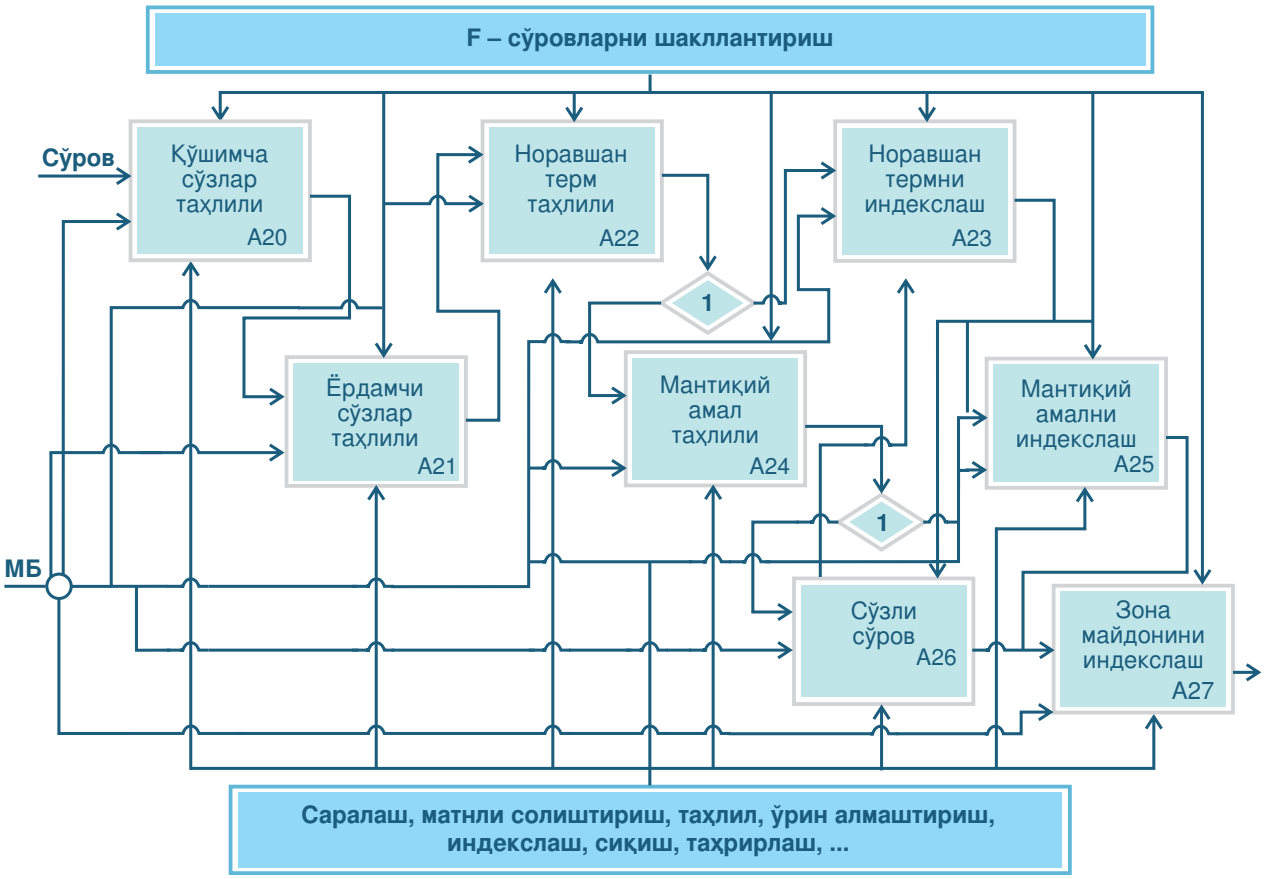
3. Фикрни излашга қаратилган усуллари (венн диаграмма, 7 бошқарув элементи).

Бундай усуллардан жуда 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 усулларидан ҳам

UI усуллари фойдаланувчига ҳеч қандай қийинчилик туғдирмасдан сўровни сунъий тилда яратади. Буни эса машина махсус тилда тузилган сўровлар каби қайта ишлайди. UI усуллари асосан фикрларни излаш усулларига таянади

6-расм

***f1* ОПЕРАТОРИНИНГ IDEF0 МОДЕЛИ**



фойдаланиш мумкин. FSV технологиясида UI усулларнинг 3 тасини, яъни ақлий ҳужум (5 босқич), саволлар рўйхати, 7 элемент бошқаруви учун инструментал дастурий модуллар яратилган.

FSV технологияси архитектураси 3 та асосий қисмдан иборат (3-расмга қаранг). Бу қисмлардан бирининг функционал имкониятлари S – ЭРнинг сўровга мувофиқлигини ҳисоблаш жараёнлари ва V – ЭРларни тақдим этишнинг усули қисмларнинг функционал имкониятларини ишлаб чиқамиз. Бу иккита қисм бир-биридан алоҳида инструментал дастурий модул сифатида қаралади.

IDEF0 модели FSV технологиясининг ядроси бўлиб унга кирувчи параметрлар «сўров», «сўров тури», «МБ» ва чиқиш параметрлари «ЭРлар тўплами»дир

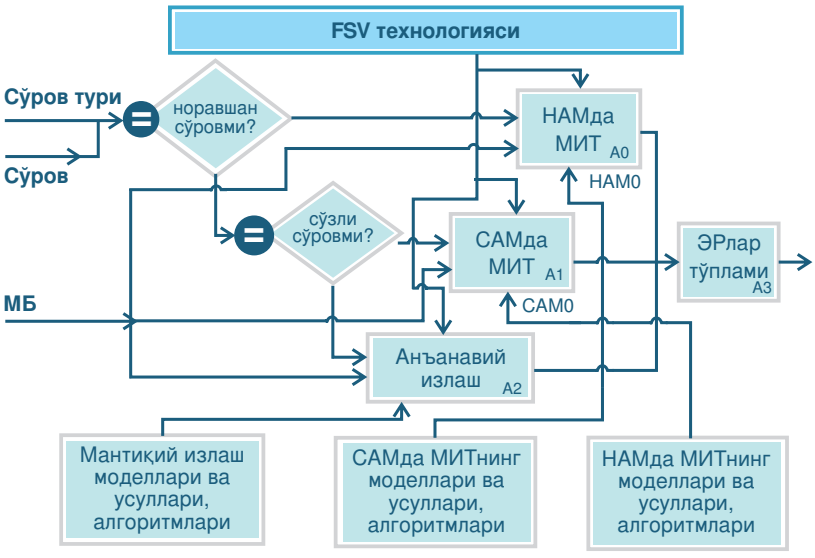
S – ЭРнинг сўровга мувофиқлигини ҳисоблаш жараёнлари ҳаволалар асосида электрон ресурснинг долзарблигини ҳисоблаш, КАРМда маълумотларни излаш тизимининг дастурий модулининг IDEF0 моделлари ҳамда F-сўровларни шакллантириш қисмининг сўровлар ва терминлар тўплами, V-ЭРларни тақдим этишнинг ҳисобчи-назоратчи функцияларини статистик маълумотларга асосланади *(7-расм)*.

Қуйида келтирилган IDEF0 модели FSV технологиясининг ядроси бўлиб унга кирувчи параметрлар «сўров», «сўров тури», «МБ» ва чиқиш параметрлари «ЭРлар тўплами»дир.

Кирувчи параметрларда «сўров» бу терминлар тўплами, «сўров тури» сўровнинг тури, яъни норавшан, сўзли, мантиқий сўров бўлиши мумкин, «Маълумотлар базаси» ЭРларнинг тўплами, семантик ядро, норавшан билимлар базаси ва ядроси, сўровлар ва терминлар тўплами, статик маълумот ва ҳисоботлар, мантиқий семантик алоқалар, йиғма динамик ва статик маълумотлардан иборат. «сўров» ва «сўров тури» жуфтлик ҳисобланади ва доим биргаликда бўлиши лозим.

7-расм

S – ЭРНИНГ СЎРОВГА МУВОФИҚЛИГИНИ ҲИСОБЛАШ ЖАРАЁНЛАРИНИНГ IDEF0 МОДЕЛИ



IDEF0 моделида «НАМда МИТ»да норавшан сўров асосида КАРМда маълумотларни излаш тизимининг моделлари ва усуллари, алгоритмларига асосланади. «САМда МИТ»да сўзли сўров (зона майдонли сўров) асосида КАРМда маълумотларни излаш тизимининг моделлари ва усуллари, алгоритмларига асосланади. «Анъанавий излаш»да мантиқий, махсус лингвистик восита асосида яратилган сўровлар мантиқий излаш моделларининг индекслаш усуллари, 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 деб ҳисобланган ЭРларнинг сараланмаган тўплами ҳосил бўлади. Бу тўплам ЭР номи, ЭРнинг уникал коди, долзарблик қиймати, мурожаатлар сони, ЭРнинг бошқа ЭРлар билан алоқаси, асосий калит сўзлари (термин, сўров кўринишида), муаллифи, яратилган санаси, охирги мурожаат санаси, ёрлиғи, тўлиқ матннинг ҳажми, тили, файл типи каби маълумотлар тузилмасига эга. Бу FSV технологиясининг V – ЭРларни тақдим этиш қисми учун зарур.

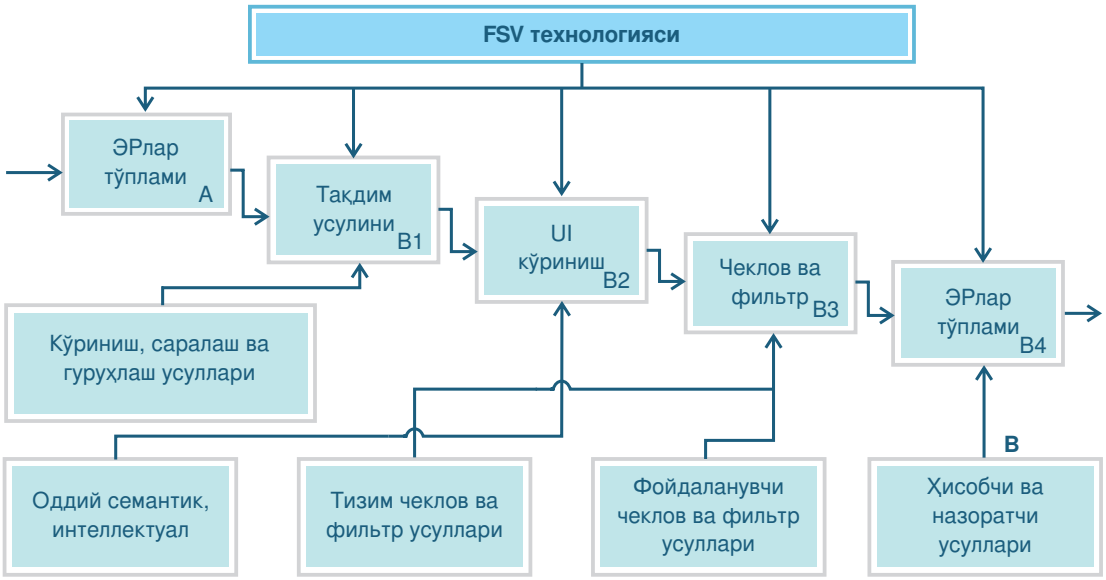
FSV технологиясининг V – ЭРларни тақдим этиш қисми асосан фойдаланувчининг хоҳишига ва тизимнинг имкониятларидан келиб чиққан ҳолда амалга оширилади. Бу қисм ЭРлар тўпламини қабул қилиш, тақдим усулини танлаш, UI кўринишни аниқлаш, чеклов ва фильтрларни таъминлаш ва шуларга асосланган ЭРлар тўпламини тақдим қилиш, фойдаланувчининг ЭРлар тўпламидаги мурожаатларини ҳисоблаш, назорат қилиш жараёнларини ўз ичига олади *(8-расм)*.

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» 3 та гуруҳга ажратилади. Булар:

- кўриниш усуллари (жадвал, ёрлиғи асосида, рўйхат, тартибсиз);
- саралаш усуллари (ЭР коди, долзарблиги, мурожаатлар сони, санаси, охирги мурожаат санаси, матн тили, матн ҳажми);
- гуруҳлаш усуллари (мурожаатлар сони, санаси, охирги мурожаат санаси, матн тили, матн ҳажми).

8-расм

V- ЭЛЕКТРОН РЕСУРСЛАРНИ ТАҚДИМ ЭТИШНИНГ IDEF0 МОДЕЛИ



Моделда «UI кўриниши» ҳам 3 та кўринишда оддий, яъни фақат ЭР номи ва бир нечта параметрлари билан, семантик, яъни ЭРнинг бошқа ЭРлар билан алоқаси, мурожаатлар асосида, интеллектуал, яъни ББсининг ядроси асосидаги семантик алоқалари, ҳаволаларининг рейтинги асосида амалга оширилади.

Моделда «Чеклов ва фильтр» бу асосан тизим ва фойдаланувчининг талаби асосида аниқланади ва 2 гуруҳга ажратилади:

- тизимнинг чекловлари ва фильтрлари. Бунда тизимнинг махсус талаблари сўровга мос маълумотларнинг тўпламини тақдим қилиш, мурожаатлар кам ЭРларни тўпламга киритиш, янги ЭРларни олдинроққа жойлаштириш ва бошқалар қўйилади.

- фойдаланувчининг чеклов ва фильтрлари. Бунда фойдаланувчи йил бўйича чеклов, нашр тури бўйича, соҳа бўйича, файл типи, матн тили, файл ҳажми, махсус калит сўзлар бўйича чеклов ва фильтрларни ўрнатиши мумкин.

Моделда асосан созлашлар фойдаланувчининг шахсий кабинетидан амалга оширилади. Агар созлаш амалга оширилмаса одатий (жорий) тизимнинг созланмалари билан тақдим қилади.

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, сўровларнинг тахминий семантик алоқаси, Билимлар базасини ривожлантиришда фойдаланади.

Маълумотларни излаш тизимининг фойдаланувчилар сони қанча кўп бўлса, интеллектуал излаш ва қайта ишлаш аниқлиги шунча юқори бўлади.

Бахадыр Муминов, заведующий кафедрой по основам информатики при Ташкентском Университете информационных технологий, доктор технических наук

В данной статье, описывается разработанная технология FSV для интеллектуального поиска данных в Информационно-ресурсных центрах (ИРЦ). Разработаны необходимые требования для информационных систем ИРЦ по внедрению в технологию FSV. IDEF0 – модели созданы для каждого этапа технологий FSV. При разработке методов UI для запросов основных этапов интеллектуального формирования пользовательского интерфейса использованы методы поиска, которые делятся на 3 группы. Третий этап IDEF0 – моделей разделен на 3 группы и осуществляется на основе пользователей, присутствующих в модели «фильтр и установить ограничения». Технологии FSV – это интерактивные технологии, которые должны отличаться семантическим ядром, пакетом запросов, запросами приблизительных семантических связей, развитием баз знаний. ■