

AXBOROT TEXNOLOGIYALARI UNIVERSITETI
HUZURIDAGI ILMIY DARAJALAR BERUVCHI
DSc.13/30.12.2019.T.07.01 RAQAMLI ILMIY KENGASH
TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI

SAIDRASULOV SHERZOD NORBOY O'G'LI

ELEKTRON HUKUMAT MUHITIDA O'ZBEK TILIDAGI
FUQAROLAR MUROJAATLARINI FORMALLASHTIRISH
ALGORITMLARI VA DASTURIY MAJMUASI

05.01.04 – Hisoblash mashinalari, majmualari va kompyuter tarmoqlarining matematik va dasturiy ta'minoti

TEXNIKA FANLARI BO'YICHA FALSAFA DOKTORI (PHD) DISSERTATSIYASI
AVTOREFERATI

Toshkent -2023

**Texnika fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi avtoreferati
mundarijasi**

**Оглавление автореферата диссертации доктора философии (PhD) по
техническим наукам**

**Contents of dissertation abstract of the doctor of philosophy (PhD)
on technical sciences**

Saidrasulov Sherzod Norboy o'g'li
Elektron hukumat muhitida o'zbek tilidagi fuqarolar murojaatlarini
formallashtirish algoritmlari va dasturiy majmuasi 3

**ВОЗВРАТИТЕ КНИГУ НЕ ПОЗЖЕ
обозначенного здесь срока**

ений граждан
..... 21

A
521

TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI
HUZURIDAGI ILMIY DARAJALAR BERUVCHI
DSc.13/30.12.2019.T.07.01 RAQAMLI ILMIY KENGASH
TOSHKENT AXBOROT TEXNOLOGIYALARI UNIVERSITETI

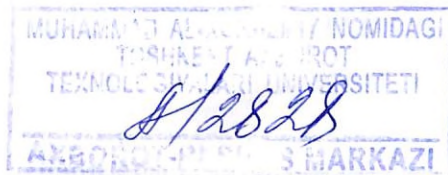
SAIDRASULOV SHERZOD NORBOY O'G'LI

ELEKTRON HUKUMAT MUHITIDA O'ZBEK TILIDAGI FUQAROLAR
MUROJAATLARINI FORMALLASHTIRISH ALGORITMLARI
VA DASTURIY MAJMUASI

05.01.04 – Hisoblash mashinalari, majmualari va kompyuter tarmoqlarining matematik va dasturiy ta'minoti

TEXNIKA FANLARI BO'YICHA FALSAFA DOKTORI (PhD)
DISSERTATSIYASI AVTOREFERATI

Toshkent -2023



Texnika fanlari bo'yicha falsafa doktori (PhD) dissertatsiyasi mavzusi O'zbekiston Respublikasi Vazirlar Mahkamasi huzuridagi Oliy attestatsiya komissiyasida B2022.4.PhD/T.3213 raqam bilan ro'yxatga olingan.

Dissertatsiya Toshkent axborot texnologiyalari universitetida bajarilgan.
Dissertatsiya avtoreferati uch tilda (o'zbek, rus, ingliz (rezyume)) Ilmiy kengash veb-sahifasida (www.tuit.uz) va «Ziyonet» Axborot ta'lim portalida (www.ziyonet.uz) joylashtirilgan.

Ilmiy rahbar: Nishanov Axram Xasanovich
texnika fanlari doktori, professor

Rasmiy opponentlar: Muxamedieva Dilnoza Tulkunovna
texnika fanlari doktori, professor

Eshankulov Hamza Ilxomovich
texnika fanlari falsafa doktori (PhD), dotsent

Yetakchi tashkilot: Berdaq nomidagi Qoraqalpoq davlat universiteti

Dissertatsiya himoyasi Toshkent axborot texnologiyalari universiteti huzuridagi DSc.13/30.12.2019.T.07.01 Ilmiy kengashning 2023-yil «17» mart soat 16:00 dagi majlisida bo'lib o'tadi. (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-70, e-mail: iktuit@tuit.uz).

Dissertatsiya bilan Toshkent axborot texnologiyalari universiteti Axborot-resurs markazida tanishish mumkin (270 raqam bilan ro'yxatga olingan.). (Manzil: 100084, Toshkent shahri, Amir Temur ko'chasi, 108-uy. Tel.: (99871) 238-64-70).

Dissertatsiya avtoreferati 2023-yil «01» mart da tarqatildi.
(2023-yil «10» fevral dagi 7 raqamli reestr bayonnomasi.)



M.M.Musayev
Ilmiy darajalar beruvchi Ilmiy kengash raisi,
texnika fanlar doktori, professor

N.O.Raximov
Ilmiy darajalar beruvchi Ilmiy kengash
ilmiy kotibi, texnika fanlar doktori, dotsent

X.N.Zaynidinov
Ilmiy darajalar beruvchi Ilmiy kengash huzuridagi ilmiy
seminar raisi, texnika fanlar doktori, professor

KIRISH (falsafa doktori (PhD) dissertatsiyasining annotatsiyasi)

Dissertatsiya mavzusining dolzarbligi va zarurati. Dunyoda kechayotgan ijtimoiy-iqtisodiy jarayonlarga hamohang mamlakatimizda ham axborot-kommunikatsiya texnologiyalaridan barcha sohalarda keng foydalanish va uni muntazam ravishda takomillashtirib borish borasida tizimli ishlar olib borilmoqda.

AKT infratuzilmasi, elektron xizmatlar va inson kapitali omillarini inobatga olgan holda «keyingi 20 yildan buyon BMT tomonidan elektron hukumat tizimining rivojlanish indeksi yuritilib, reytingda Buyuk Britaniya, Janubiy Koreya, Daniya, Norvegiya, AQSH, Kanada, Shvetsiya, Avstraliya, Singapur va Niderlandiya kabi davlatlar yetakchilik qilishmoqda»¹.

Elektron hukumat rivojlanganligida asosiy hamda natijaviy shart bu aholining ommaviy «elektron ishtiroki»dir. Ommaviy elektron ishtirok hajmi va sifatini oshirish uchun albatta elektron hukumat doirasida xizmatlarni kengaytirish, aholining barcha qatlamlari uchun turli daraja va doiradagi muammo, ehtiyoj va «murojaat»larini davlat va nodavlat organlariga elektron shaklda yetkaza olishini ta'minlash zarur.

Jahonda elektron hukumat tizimini yanada takomillashtirish borasida yuzaga kelayotgan zamonaviy muammolarni ilmiy hal qilish, jumladan, bu yo'nalishda tobora ortib borayotgan katta hajmdagi turli axborotlarning oqimini maqbul boshqarish, intellektual tahlil qilish va integratsiyalash uchun matematik-dasturiy modellar va algoritmlar yaratish muhim vazifalardan bo'lib hisoblanadi. Shu bilan birga ayniqsa elektron hukumatda xizmat ko'rsatish jarayonini intellektualashtirish va aholining ommaviy elektron ishtirokini oshirish doirasida matematik asoslangan uslubiyat yaratish va amaliyotga joriy etish masalalarning ilmiy-amaliy tadqiqotlariga alohida e'tibor qaratish zarur hisoblanadi.

O'zbekiston Respublikasining «Elektron hukumat to'g'risida»gi (2015) va «Jismoniy va yuridik shaxslarning murojaatlari to'g'risida»gi (2017) Qonunlari, O'zbekiston Respublikasi Prezidentining «Raqamli O'zbekiston-2030» strategiyasini tasdiqlash va uni samarali amalga oshirish chora-tadbirlari to'g'risida» 2020-yil 5-oktyabrdagi PF-6079-son Farmoni va «Sun'iy intellekt texnologiyalarini jadal joriy etish uchun shart-sharoitlar yaratish chora-tadbirlari to'g'risida» 2021-yil 17-fevraldagi PQ-4996-son Qarori, Vazirlar Mahkamasining «Fuqarolarning o'zini o'zi boshqarish organlarida jismoniy va yuridik shaxslarning murojaatlari bilan ishlash tartibi to'g'risida namunaviy nizomni tasdiqlash haqida» 2018-yil 5-yanvardagi 7-son va «Davlat organlarida, davlat muassasalarida va davlat ishtirokidagi tashkilotlarda jismoniy va yuridik shaxslarning murojaatlari bilan ishlash tartibi to'g'risidagi namunaviy nizomni tasdiqlash haqida» 2018-yil 7-maydagi 341-son qarorlari hamda mazkur faoliyatga tegishli boshqa me'yoriy-huquqiy hujjatlarda belgilangan vazifalarni amalga oshirishga ushbu tadqiqot muayyan darajada xizmat qiladi.

Tadqiqotning respublika fan va texnologiyalari rivojlanishining ustuvor yo'nalishlariga mosligi. Mazkur tadqiqot Respublika fan va texnologiyalari

¹<https://publicadministration.un.org> (E-Government Development Index)

rivojlanishining IV. «Axborotlashtirish va axborot-kommunikatsiya texnologiyalarini rivojlantirish» ustuvor yo‘nalishi doirasida bajarilgan.

Muammoning o‘rganilganlik darajasi. Elektron hukumatni doimiy takomillashtirish, iste‘molchilarning axborot va xizmatga bo‘lgan ehtiyojini qanoatlantirish, sifatni oshirish, elektron hukumatda samarali elektron ishtirokni oshirish omillari va modellari, tabiiy tildagi matnli hujjatlarga ishlov berish, intellektual tahlil qilish, tasniflash, matnli hujjatlardan kalit so‘zlarni ajratib olishning matematik va dasturiy ta‘minoti, usullari va texnologiyalarini ishlab chiqish, amaliyotga tatbiq etish va rivojlantirish borasidagi ixtisoslashtirilgan ilmiy muassasalarda ko‘plab olimlar ilmiy-amaliy tadqiqotlar olib borishmoqda.

Jumladan, elektron hukumat tizimi, unga ta‘sir qiluvchi omillar, tizim tatbiqi va tahlillari bo‘yicha B.Wirtz, T.Kvasnisova, Z.V.Arkipova, G.V.Bestolkova, G.O.Barbakov, N.V.Beryoza, S.V.Ponomarev, A.K.Skopintsev, A.V.Sadilova, elektron hukumatda samarali elektron ishtirokni oshirish omillari va modellari bo‘yicha Smith & Dalakiouridou, Fedotova, Teixeira, Alvelos, D.Bennet, Revyakin S.A., Y.Panopoulou, A.Makintosh shug‘ullanganlar. Tabiiy tildagi matnli hujjatlarni qayta ishlash, intellektual tahlil qilish, tasniflash masalalari bo‘yicha Kozlov, Batura, Fabritsio Sebastian, Uilyam B. Cavnar, Vandana Korde, Zakaria Elberichi, Ron Bekkerman, matnli hujjatlardan kalit so‘zlarni ajratib olish bo‘yicha masalalari Stojok, Moskvitina, Grineva, Grinev, Rose, Astraxantsev, Zaxarov, Xoxlova, Lukashevich, Logachev muallifligidagi ishlarda yoritilgan. Shuningdek, respublikamizda M.M.Kamilov, T.F.Bekmuratov, SH.X.Fozilov, A.X.Nishanov, N.S.Mamatov, O.J.Babomuratov, E.S.Babadjanov, L.B.Boboev kabi olim va tadqiqotchilar bu yo‘nalishda ilmiy-tadqiqotlarga o‘z hissalarini qo‘shishgan.

Elektron hukumat muhitida murojaatlar portali orqali yo‘llangan tabiiy o‘zbek tilidagi matnli murojaatlarni avtomatik formallashtirish va predmet sohalar bo‘yicha tasniflash imkonini beruvchi usul va algoritmlarni yaratish muammolari yetarli darajada o‘rganilmagan.

Tadqiqotning dissertatsiya bajarilgan oliy ta‘lim muassasasining ilmiy-tadqiqot ishlari rejalari bilan bog‘liqligi. Dissertatsiya tadqiqoti Muhammad al-Xorazmiy nomidagi Toshkent axborot texnologiyalari universitetida №22/19-F «Milliy elektron hukumat muhiti xizmatlaridan foydalanish samaradorligini oshirishning «iGov – maslahat-muhokama-monitoring tizimi»ni ishlab chiqish va uning tatbiqi» (2019-2020) loyihasi doirasida bajarilgan.

Tadqiqotning maqsadi Elektron hukumat muhitida murojaatlar portali orqali yo‘llangan tabiiy o‘zbek tilidagi matnli murojaatlarni avtomatik formallashtirish va predmet sohalar bo‘yicha tasniflash masalasining algoritmik va dasturiy majmuasini ishlab chiqishdan iborat.

Tadqiqotning vazifalari:

elektron hukumat muhitida jismoniy va yuridik shaxslarning elektron ishtiroki va murojaatlarni tahlil qilish;

tabiiy o‘zbek tilidagi matnli murojaatlarni tasniflash masalasini formal qo‘yish va murojaatlarni qayta ishlashning o‘ziga xos xususiyatlarini ishlab chiqish;

o'zbek tilidagi matnli murojaatlarni avtomatlashgan holda predmet sohalarga ajratish jarayonini loyihalashtirish, qayta ishlashning o'ziga xos infrastrukturasi ishlab chiqish;

o'zbek tilidagi matnli murojaatlarni formallashtirish va formal murojaatlar bazasidan murojaatni qidirish algoritmini ishlab chiqish;

matnli ma'lumotlarni qayta ishlash uchun o'zbek tiliga xos ma'lumotlar va bilimlar bazasi, ekspert va dinamik tanlanmalar, ranjirlash funksiyalari hamda algoritmik va dasturiy majmuasini ishlab chiqish.

Tadqiqotning ob'ekti sifatida elektron hukumat muhitida murojaatlar bilan ishlashga qo'yilgan talablar, murojaatlar portali hamda u orqali yo'llangan tabiiy o'zbek tilidagi matnli murojaatlar bazasi olingan.

Tadqiqotning predmeti elektron hukumat muhitida matnli murojaatlarni formallashtirish va avtomatik predmet sohalarga ajratish usul va algoritmlari hamda ularning dasturiy majmuasini ishlab chiqish bilan bog'liq bo'lgan barcha dasturiy vositalarini tashkil etadi.

Tadqiqotning usullari. Tadqiqot jarayonida matematik modellashtirish, dasturlash texnologiyalari, timsollarni tasniflash, tizimli tahlil va ma'lumotlarning intellektual tahlili, ma'lumotlar va bilimlar bazasini shakllantirish usullaridan foydalanilgan.

Tadqiqotning ilmiy yangiligi quyidagilardan iborat:

tabiiy o'zbek tilidagi matnlarni avtomatik qayta ishlovchi, yagona tizimga integratsiya qiluvchi lingvistik ma'lumotlar bazasi loyihalashtirilgan va yaratilgan;

tabiiy o'zbek tilidagi fuqarolarning matnli murojaatlarini tasniflash uchun kengaytirilgan qatlamlarga ajratish algoritmi, mezonli tasniflash tanlanmalari va ranjirlash funksiyalari ishlab chiqilgan;

matnlarni formallashtirish usuli va formal murojaatlar bazasini kombinatsiyali qidirish algoritmi ishlab chiqilgan;

fuqarolar murojaatlarini tasniflash va formallashtirish algoritmlari asosida elektron hukumat tizimlari uchun dasturiy funksional modullarining tuzilmasi ishlab chiqilgan.

Tadqiqotning amaliy natijalari quyidagilardan iborat:

elektron hukumat muhitida jismoniy va yuridik shaxslarning elektron ishtiroki va murojaatlarni ta'minlovchi, tabiiy o'zbek tilidagi matnli murojaatlarni tasniflovchi, murojaatlarni formallashtiruvchi va ularni qayta ishlashning o'ziga xos xususiyatlarini aniqlash tamoyillariga mos xizmat ko'rsatuvchi uslubiyatlar ishlab chiqilganligi;

o'zbek tilidagi matnli murojaatlarni avtomatik predmet sohalarga ajratish jarayonini loyihalashtirishga, ma'lumotlarni qayta ishlashning o'ziga xos infrastrukturasi inobatga olgan dasturiy majmuani ishlab chiqilishi;

o'zbek tilidagi matnli murojaatlarni formallashtirish va formal murojaatlar bazasidan murojaatni qidirishning algoritmik ta'minotini ishlab chiqilishi;

matnli ma'lumotlarni qayta ishlash jarayonida o'zbek tilining o'ziga xos ma'lumotlar bazasi, ekspert va dinamik tanlanmalar, ranjirlash funksiyalari hamda algoritmik ta'minotlari asosida fuqarolar murojaatlarini tasniflash va formallashtirish algoritmlari orqali elektron hukumat tizimlari uchun dasturiy funksional modullarining tuzilmasi ishlab chiqilishi bilan izohlanadi.

Tadqiqot natijalarining ishonchliligi. Tadqiqot natijalarining ishonchliligi qo'yilgan muammoning chuqur tahlili, matematik modellashirishning korrektiligi, dasturlash texnologiyalari va algoritmlashtirish nazariyasi hamda tizimli va intellektual tahlil qilish usullarining to'g'ri qo'llanishi, murojaatlar bazasidan murojaatlarni ko'p parametrlil qidirish va tabiiy o'zbek tilida ifoda etilgan ma'lumotlarni formallashtirishda yangicha algoritmik yondashuv qo'llanilishi, shuningdek elektron xizmat ko'rsatish borasidagi nazariy va amaliy tadqiqotlardan olingan natijalar va veb platformalarni ishlab chiqish talablariga muvofiqligi orqali izohlanadi.

Tadqiqot natijalarining ilmiy va amaliy ahamiyati. Tadqiqot natijalarining ilmiy ahamiyati elektron hukumat muhitida tabiiy o'zbek tilidagi matnli murojaatlarni formallashtirish va avtomatik predmet sohalariga ajratish jarayonini loyihalashtirish, qayta ishlashning o'ziga xos infrastrukturasi yaratish, sinflashtirish usul va algoritmlarini ishlab chiqish, tabiiy o'zbek tilidagi matnli murojaatlarni formallashtirish va formal murojaatlar bazasidan murojaatni qidirish algoritmini ishlab chiqish, matnli ma'lumotlarni qayta ishlashda o'zbek tilining o'ziga xos ma'lumotlar bazasi, ekspert va dinamik tanlanmalar, ranjirlash funksiyalari hamda algoritmik ta'minotini ishlab chiqilishi bilan izohlanadi.

Tadqiqot natijalarining amaliy ahamiyati elektron hukumat muhitida jismoniy va yuridik shaxslarning elektron ishtiroki va murojaatlarni ta'minlovchi, tabiiy o'zbek tilidagi matnli murojaatlarni tasniflovchi, murojaatlarni formallashtiruvchi va ularni qayta ishlashning o'ziga xos xususiyatlari aniqlash tamoyillariga mos xizmat ko'rsatuvchi uslubiyatlar, tabiiy o'zbek tilidagi matnli murojaatlarni avtomatik predmet sohalarga ajratish jarayonini loyihalashtirish, ma'lumotlarni qayta ishlashning o'ziga xos infrastrukturasi, o'zbek tilidagi matnli murojaatlarni formallashtirish va formal murojaatlar bazasidan murojaatni qidirish, o'zbek tilining o'ziga xos ma'lumotlar bazasi, ekspert va dinamik tanlanmalar, ranjirlash uslubiyatlari hamda algoritmik ta'minotlari asosida tabiiy o'zbek tilidagi murojaat matnlarni formallashtirish orqali predmet sohasini aniqlash dasturiy majmuasi ishlab chiqilishi bilan izohlanadi.

Tadqiqot natijalarining joriy qilinishi. Elektron hukumat muhitida tabiiy o'zbek tilidagi matnli murojaatlarni avtomatik formallashtirish va predmet sohalari bo'yicha tasniflash masalasi uchun ishlab chiqilgan usul va algoritmlar bo'yicha olingan natijalar asosida:

Tadqiqotchi tomonidan ishlab chiqilgan dasturiy ta'minot virtual qabulxonaning markaziy boshqaruv bo'limiga 2022-yilning aprel oyidan boshlab joriy etilgan bo'lib, 5 oy davomida kelib tushgan 450 mingga yaqin murojaatning 32% ni avtomatik sohalarga ajratish imkonini berdi. Ushbu natijalar murojaatlar turi va sonini qanchalik ortishi bilan bilimlar bazasi kengayishi hamda buning natijasida samaradorlik ko'rsatkichini ko'tarilishi ko'rsatadi (Axborot texnologiyalari va kommunikatsiyalarini rivojlantirish vazirligining 2022-yil 30-avgustdagi 33-8/5583-son ma'lumotnomasi).

Tadqiqotchi tomonidan ishlab chiqilgan dasturiy ta'minot Elektron hukumat muhitida qo'llanilishi, ya'ni tegishli mutaxassislar tomonidan murojaatlarni avtomatik qayta ishlashlarida ularga oqilona qaror qabul qilishlari uchun ko'mak berish imkoniyati yaratilgan. Natijada, bu esa ish unumdorligi o'rtacha 20-30 % gacha

ortishiga erishildi. Dissertatsiya ishi natijasida yaratilgan dasturiy majmuadan O'zbekiston Respublikasi elektron hukumat axborot tizimlarida tabiiy o'zbek tilidagi matnli murojaatlarni avtomatik formallashtirish va predmet sohalar bo'yicha tasniflash masalalarida foydalanilishi ish samaradorligini oshishiga xizmat qiladi («UNICON.UZ» - Fan-texnika va marketing tadqiqotlari markazining 2022-yil 22- avgustdagi 5-3/1465-son dalolatnomasi).

Dasturiy ta'minoti tegishli soha yo'nalishi kesimidagi fuqarolar murojaatlari asosida shakllantirilgan o'quv tanlanmadan foydalangan holda statik termlar to'plamini shakllantirish, murojaatlarni formallashtirish va tasniflash ishlari amalga oshiradi hamda elektron hukumat muhitida qo'llanilishi, ya'ni tegishli mutaxassislar tomonidan murojaatlarni avtomatik qayta ishlashlarida ularga oqilona qaror qabul qilishlari uchun ko'mak berish imkoniyati yaratdi. Natijada, ish unumdorligi o'rtacha 20-30 % gacha ortishiga erishildi ("O'zbekiston pochta" AJning 2022-yil 18-avgustdagi dalolatnomasi).

Kelajakda mazkur dasturiy majmua Elektron hukumat muhitiga joriy etilsa fuqarolarning elektron murojaatlarining tegishi predmet sohalarini aniqlashda dasturiy vosita tomonidan moderatorlarning oqilona qaror qabul qilishlari ish unumdorligining 20-25% gacha ortishiga, ish hajmining esa 10-20% gacha kamayishiga erishiladi («Elektron hukumat loyihalarini boshqarish markazi» davlat markazining 2022-yil 22-avgustdagi dalolatnomasi).

Tadqiqot natijalarining aprobatsiyasi. Mazkur tadqiqot natijalari 2 ta xalqaro va 3 ta respublika ilmiy-amaliy anjumanlarida ma'ruza qilingan va muhokamadan o'tkazilgan.

Tadqiqot natijalarining e'lon qilinganligi. Tadqiqot mavzusi bo'yicha jami 28 ta ilmiy ish chop etilgan, shulardan O'zbekiston Respublikasi Oliy attestatsiya komissiyasining doktorlik dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan ilmiy nashrlarda 12 ta maqola, jumladan 2 tasi xorijiy va 10 tasi respublika jurnallarida chop etilgan hamda 3 ta elektron hisoblash mashinalari uchun yaratilgan dasturiy vositalarni qayd qilish guvohnomolari olingan.

Dissertatsiyaning tuzilishi va hajmi. Dissertatsiya kirish qismi, to'rtta bob, xulosa, foydalanilgan adabiyotlar ro'yxati va ilovalardan iborat. Dissertatsiyaning hajmi 120 betni tashkil etadi.

DISSERTATSIYANING ASOSIY MAZMUNI

Kirish qismida dissertatsiya mavzusining dolzarbligi va zaruriyati asoslangan, tadqiqotning O'zbekiston Respublika fan va texnologiyalari rivojlanishining ustuvor yo'nalishlariga mosligi ko'rsatilgan, maqsad va vazifalari belgilangan va tadqiqot obyektini va predmetini aniqlangan, olingan natijalarning ishonchligi asoslab berilgan, ularning nazariy va amaliy ahamiyati, tadqiqot natijalarini amalda joriy qilish holati, nashr etilgan ishlar va dissertatsiyaning tuzilishi bo'yicha ma'lumotlar keltirilgan.

Dissertatsiyaning «Elektron hukumat muhitida murojaatlarning tashkil etilganlik holati va normativ-huquqiy hujjatlar tahlili» deb nomlangan birinchi bobi to'rtta paragrafdan iboratdir. Ushbu bob jamoatchilik nazorati, fuqarolik jamiyati institutlari rolini kuchaytirish, axborot-kommunikatsiya texnologiyalari orqali davlat xizmatlari ko'rsatishning ochiqligi va sifatini oshirishga qaratilgan. Birinchi

paragrafda O'zbekistonda raqamli iqtisodiyot rivojlanish tahlillari, BMT tomonidan yuritilayotgan «elektron hukumat» tizimini rivojlanganlik tahlillari asosida davlatimizda elektron hukumat rivojlanganlik darajalarini ilg'or davlatlarga nisbatan tahlili, elektron hukumat rivojlanishi reytingi (EGDI) va undagi elektron ishtirok etish indeksleri (EPART) bo'yicha ilg'or davlatlar va MDH davlatlarining tahlillari, elektron hukumatda samarali elektron ishtirokni oshirish omillari va modellari keltirilgan. Xususan keyingi 14 yil davrda EGDI reytingi 109 dan 69 ga (+40), bunga mos elektron ishtirok ko'rsatkichi 0.0909 dan 0.7636 (+0,6727) ga o'sdi. Ikkinchi paragrafda fuqarolar murojaatlari bo'yicha tushunchalar va normativ-huquqiy hujjatlar keltirilgan. Uchinchi paragrafda elektron murojaatlarni tashkil qilish jarayoni va asosiy ko'rsatkichlar, «Murojaatlar» portalining funksional xususiyatlari, murojaatlar bilan ishlashni tashkil qilish yuzasidan xorijiy, jumladan MDH davlatlari tajribasi hamda «Murojaatlar» portalida murojaatlarni tasniflash jarayonini optimallashtirish yondashuvi keltirilgan. To'rtinchi paragrafda masalani qo'yilishi hamda birinchi bob bo'yicha xulosalar keltirilgan

Dissertatsiyaning «**Matnli murojaatlarni avtomatik qayta ishlash masalasi va uning tahlillari**» deb nomlangan ikkinchi bobda tabiiy tildagi matnli hujjatlar va murojaatlarni qayta ishlash borasidagi tadqiqot ishlari va algoritmlar tahlil qilindi hamda mavjud muammolar tasniflandi. Jumladan, matnli intellektual tahlil qilish algoritmlari sinflashtirilgan holda mohiyati ochib berildi. Matnli murojaatlarni qayta ishlashda kalit so'zlar muhim ahamiyat kasb etganligi sababli kalit so'zlarni ajratib olish usullari va algoritmlari alohida paragraf sifatida tadqiq etildi. Jumladan, kalit so'zlarni chiqarishning muammoli sohasi, ishlab chiqilgan usullar va ularni amalga oshiradigan algoritmlar to'g'risidagi nashrlarni tizimlashtirishga va istiqbolli tadqiqot yo'nalishlarini tanlash asoslangan (1-rasm).



1-rasm. Kalit so'zlarni ajratib olish usullarining tasnifi

Tabiiy tillar ko'p va ularning formal shakllanishi turli-tumanligi sababli odatda matnni qayta ishlovchi algoritmlar biror aniq tabiiy tilga nisbatan yaratiladi. Demak, mavjud ko'plab algoritmlar ko'proq ommalashgan tabiiy tillarga mos yaratilgan bo'lib, ularni o'zbek tiliga modifikatsisiz qo'llash natija bermaydi. Shuning uchun tabiiy

tilidagi matnlarni qayta ishlashga qaratilgan tahliliy bobning natijasi sifatida o'zbek tilidagi matnlarni formallashtirish uslubiyatlari tadqiq etilgan.

Dissertatsiyaning «Tabiiy o'zbek tildagi murojaat matnlarni formallashtirish orqali predmet sohasini aniqlash algoritmi» deb nomlangan uchinchi bobida tabiiy o'zbek tilidagi matnli murojaatlarning predmet sohaslarini aniqlash uslubiyati va algoritmik ta'minoti tadqiq qilingan. Ishning dastlabki paragrafida o'zbek tilidagi matnli murojaatlarni tegishli predmet sohalarga ajratish masalasini yechish g'oyasi uchun asosiy vazifalar quyidagicha keltiriladi:

berilgan masala bo'yicha *termlar va qoidalar bazasi* yaratish;

Matnni tahlil qilish algoritmini ishlab chiqish. Mazkur algoritmda berilgan matnga dastlabki ishlov berish va qatlamlarga ajratish amalga oshiriladi. Jumladan, matnni tozalash; matnni gap va so'zlarga ajratish; ismlar, nomlanishlar va raqamlarni ajratish; qisqartmalarni kengaytirish; sintaksis va leksik tahlillar olib boriladi;

Matnni formal qolip asosida qidirish. Bunda kalit so'zlarga tayagan holda joriy murojaat matni mantiqiy matematik ko'rinishdagi formal qolipga tushiriladi. Natijada, bu formal matn oldindan mavjud bo'lgan formal murojaatlar bazasidan qidiriladi. Agar natijaga erishilsa, amaliyot to'xtatiladi, aks holda keying bosqichga o'tiladi;

Tasniflash usuli ishlab chiqish. Joriy qatlamlarga ajratish natijasida yarim formallashtirilgan matn ekspert va dinamik tanlanmalar to'plami bo'yicha predmet sohalari aniqlanadi va u foydalanuvchiga taqdim etiladi;

Natijaga erishilgan murojaat matnning predmet sohasini tasdiqlash va uni to'liq formallashtirish jarayoni. Algoritm natijasi to'g'ri bo'lsa, joriy formal matn formal murojaatlar bazasiga yoziladi.

Keltirilgan vazifalariga asosan murojaatni avtomatik tahlil qilishda dastlab zarur belgilashlar va ma'lumotlar bazalari yaratiladi. Bular quyidagicha:

T – o'zbek tilidagi termlar (iboralar) to'plami: $T = \{t_1, t_2, \dots, t_n\}$, n – mavjud termlar soni. Asosan mustaqil so'z turkumiga mansub so'zlar.

C – so'z turkumlari va ularning xususiyatlari to'plami.

R – raqamlar to'plami: $R = \{r_1, r_2, \dots, r_{rn}\}$, rn – mavjud raqamlar soni.

Q – qisqartmalar to'plami: $Q = \{q_1, q_2, \dots, q_{qn}\}$, qn – qisqartma soni. To'plam elementlari ko'p parametrlil: $q_i = \langle \bar{q}_i^1, \bar{q}_i^2, \bar{q}_i^3 \rangle$. Bu yerda \bar{q}_i^l – qisqartmaning turli shakllari, \bar{q}_i^2 – kengaytmasi, \bar{q}_i^3 – tegishli sohasi

G – yordamchi so'zlar (ko'makchi, so'z bog'lovchilari, yuklamalar) to'plami.

H – affiksalar to'plami $H = \{h_1^{o^1}, h_2^{o^2}, \dots, h_{hn}^{o^{hn}}\}$, bu yerda hn – soni, $h_i^{o^l}$ – affiks va uning old yoki orqa qo'shimcha ekanligi binar $o^l = \langle 0, 1 \rangle$ orqali aniqlanadi (0 bo'lsa, old qo'shimcha, 1 bo'lsa, 'zning orqa qo'shimcha).

Ishda T, R, H, G, Q , to'plamlar kalit so'zlar bazasi (KSB) deb yuritiladi.

TS – bir nechta iboralardan turivchi nomlanishlar (tashkilot, manzillar) to'plami: $TS = \{ts_1, ts_2, \dots, ts_{tsn}\}$, tsn – mavjud atamalar soni.

FS – ismlar kitobidan olingan shaxslarning ismlari to'plami: $FS = \{fs_1, fs_2, \dots, fs_{tsn}\}$, fsn – mavjud ismlar soni.

TN – termlar yaqinlik darajasining to‘plami: $TN = \{ \langle \bar{t}_i, \bar{t}_j \rangle, t_{f_{ij}} \}; \forall \bar{t}_i, \bar{t}_j \in T; t_{f_{ij}} \in (0,1]; i, j \leq n; i \neq j$, bu yerda $t_{f_{ij}}$ - berilgan \bar{t}_i va \bar{t}_j termlarning o‘zaro yaqinligi. Agar $t_{f_{ij}} = 1$ bajarilsa, u holda $\langle \bar{t}_i, \bar{t}_j \rangle$ termlar sinonim so‘zlar.

P – predmet sohalari (PS) to‘plami: $P = \{ p_1, p_2, \dots, p_{pn} \}$, np – sohalari soni.

PT – PSga tegishli termlar to‘plami: $PT = \{ pt_k = \langle p_i, t_{j_i} \rangle; i \leq pn; j_i \leq n; pt_k \leq n \times pn; \forall p_i \in P, t_{j_i} \in T$, bunda p_i -predmet soha, t_{j_i} -term indeksi.

PTN – PS bo‘yicha termlarning o‘zaro yaqinlik darajalari to‘plami: $PTN = \{ \langle \overline{pt}_i, \overline{pt}_j \rangle, pt_{f_{ij}} \}; \forall \overline{pt}_i, \overline{pt}_j \in PT; pt_{f_{ij}} \in (0,1]; i, j \leq n \times pn; i \neq j$. Agar $pt_{f_{ij}} = 1$ bo‘lsa, $\langle \overline{pt}_i, \overline{pt}_j \rangle$ soha termlar shartli sinonim so‘zlar hisoblanadi.

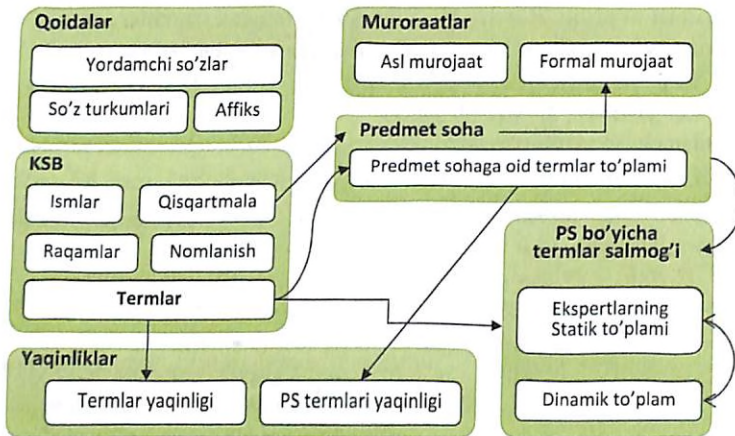
M – asl nusxadagi matnli murojaatlar to‘plami.

E – formallashtirilgan murojaatlar: $E^l = (E^{l,1}, \dots, E^{l,n_l})$. Bunda murojaat bir nechta segmentdan va segmentlar ikki parametrdan iborat: termlar va PS termlari. $E^{l,j} = (e_1^{l,j}, \dots, e_{i_{ij}}^{l,j}); e_{i_{ij}}^{l,j} = \{ (t_{k_1}^{l_{ij}}, \dots, t_{k_n}^{l_{ij}}), (pt_{h_1}^{l_{ij}}, \dots, pt_{h_n}^{l_{ij}}) \}$, $kn < N, nh < M$.

SL – termlarning sohalari kesimida ekspertlar tomonidan berilgan dastlabki statik termlar to‘plami: $SL = \{ sl_k, p_i, t_j, \gamma_{ij} \}; i \leq pn; j \leq n; k \leq n \times pn; \forall p_i \in P, t_j \in T; \gamma_{ij} \in [0, 1]$. Ushbu to‘plam elementlari qiymati oldindan mavjud va ma’lum sondagi murojaatlarni mashinali o‘qitish va ularga ekspertlar aniqlik kiritgan so‘ng olinadi hamda u o‘zgarimas etalon sifatida qabul qilinadi.

DL – sohalari kesimida dinamik termlar to‘plami: $DL = \{ dl_k, p_i, t_j, \delta_{ij} \}; i \leq pn; j \leq n; k \leq n \times pn; \forall p_i \in P, t_j \in T; \delta_{ij} \in [0, 1]$. Dinamik termlar to‘plami statik termlar to‘plamining nusxasi bo‘lib, o‘lcham va tuzilish jihatidan bir xil. Lekin, joriy murojaatlarni mashinali o‘qitishdan keyin DL dagi e_{ij} qiymati doimiy oshib boruvchi hisoblanadi.

Relyatsion modeliga asoslangan yuqoridagi ma’lumotlar bazasi ob’ektlarini o‘zaro o‘xshashlik jihati bo‘yicha bir nechta sinflarga ajratish mumkin (2-rasm).



2-rasm. Ma’lumotlar bazasi ob’ektlarining sinflari

Matni avtomatik tahlil qilish algoritmi qoidalar asosida murojaat matni qatlamlarga ajratish vazifasi bajariladi. Mazkur algoritm KSBga tayangan holda berilgan qoidalar bo'yicha matn bir nechta qatlamlarga ajratadi. Formallashtirish jarayonidagi har bir qatlamni hosil qilishda berilgan matni tahlil qilishning o'ziga xos usullari yoki algoritmi ishlab chiqiladi. Bunda berilgan matnga 9 ta bosqichda ishlov berilib, 10 ta qatlam hosil qilinadi. Qatlamlarga ajratish jarayonida ma'lumot saqlash uchun W^v ikki o'lchamli vektor kiritilib, u har bir bosqichda vertikal va gorizontal o'lchamlarini o'zgartirib boradi. Algoritm qatlamlari quyidagilardan iborat: W^0 -noaniq simvollar; W^1 -nomlanishlar; W^2 -ismlar; W^3 -raqamlar; W^4 -yordamchi so'zlar; W^5 -affikslar; W^6 -qisqartmalar; W^7 -termlar; W^8 -termlarning o'xshashligi; W^9 -PS termlari; W^{10} -PS termlarining o'xshashligi.

Algoritm natijasida olingan oxirgi to'rtta qatlam asosida yarim formal matn hosil qilinadi va unda formal qolip bo'yicha qidirish usuli foydalaniladi. Ya'ni, $W^7 - W^{10}$ qatlamlar asosida formal qolip shakllantiriladi. Buning natijasida oldindan mavjud formallashtirish murojaatlar bazasidan qidirish amaliyoti bajariladi.

Oldindan mavjud E formal murojaatlar to'plamidagi formallashtirilgan murojaat ushbu shaklda tashkil qilingan: segmentlar ketma-ketligida $(e_{t_{ij}}^{i,j})$, segmentlar term va PS termlari ajratilgan kichik segmentlar juftligida (t, pt) , kichik segmentlarga esa termlarning asl nusxalari ketma-ketligi kiradi.

Agar joriy qaralayotgan murojaat matni formal qolipga tushirilgan deb aytiladi, agar, unda term va uning sinonimlari alohida segment hamda PS termlari va ularning yaqinliklari alohida segmentlarga ajratilgan bo'lsa. Shuningdek, hosil qilinadigan segment elementlari $(w_i^t \cup w_i^p$ va $w_i^q \cup w_i^{t0})$, term indekslaridan iborat bo'lib, u quyidagicha ifodalanadi:

$$B = (b^1, b^2, \dots, b^r), b^d = \{(t_{k_1}^d, t_{k_2}^d, \dots, t_{k_d}^d), (p_{h_1}^d, p_{h_2}^d, \dots, h_{h_d}^d)\}.$$

Aksariyat hollarda b^d vektorning ikkinchi segmenti, ya'ni, $h_h^d = \emptyset$ bo'ladi. Chunki, $w_i^t \neq 0$ shartning bajarilishi doimo ushbu term PSga tegishli bo'lmaydi.

Umumiy holda qidirish masalasini ushbu ko'rinishda bayon qilish mumkin: E formallashtirilgan murojaatlar to'plamidan joriy formal qolipga tushirilgan B vektor elementlarini qidirish talab qilinadi. Qidirish algoritmini ishlab chiqishda quyidagicha shart va mezonlar oldindan belgilab olinadi:

qidirishda segmentlarning joylashuv pozitsiyalari hamda ushbu segmentga tegishli elementlar pozitsiyalari muhim emas;

qidirish segment juftligi mosligi bo'yicha olib boriladi;

qidirishda segmentlar va ulardagi elementlari soni belgilangan mezonni qanoatlantirsa, bu natija sifatida olinadi;

E to'plamning ixtiyoriy segmentning ikkinchi jufti mavjud va qidiriluvchi B vektorning segmentidagi ikkinchi jufti bor/yo'q bo'lsa, qidirish mumkin;

E to'plamning ixtiyoriy segmentining ikkinchi jufti mavjud bo'lmasa va qidiriluvchi B vektor segmentidagi ikkinchi juftlik bor bo'lsa, ikkinchi segment bo'yicha qidirilmaydi;

E to'plamdagi ixtiyoriy yozuv segmentlari sonidan qidiriluvchi B vektor segmentlari soni katta bo'lsa, mezon B vektor segmentlari soni bo'yicha olib boriladi,

aks holda, B vektor segmentlari soni E to'plam yozuvining segmentlari soniga nisbati bo'yicha mezon yuritiladi.

Soddaroq qilib aytiladigan bo'lsa, $w_i^7 \neq 0$ bo'lganda $w_i^7 \cup w_i^8$ termlar E to'plamning termlar kichik segmentidan va $w_i^9 \cup w_i^{10}$ termlar E to'plamning PS termi kichik segmentidan parallel qidiriladi. Qidirishda segmentlar joylashuv o'rinlar birinchi darajali muhim emas. Natija mavjud barcha segmentlar soniga nisbatan ularning qoplanish darajasiga bog'liq. Misol uchun matnning 5 ta pozitsiyadagi kalit so'zlarga formal to'plamning 7 ta segmentli qolipga 4 tasi javob bersin. Agar belgilangan mezon 70% bo'lsa, demak, murojaat qolipga tushdi hisoblanadi. Bu esa olib borilayotgan jarayonni natijasi sifatida qaraladi va E formal murojaat to'plamida ko'rsatilgan PS foydalanuvchiga taqdim etiladi.

Keyingi vazifada murojaat matnini tasniflash uchun statik va dinamik tanlanmalar ishlab chiqiladi. Uslubiyat ikki bosqichda amalga oshiriladi, ya'ni: SL va DL tanlanmalar asosida W^7 va W^8 hamda W^9 va W^{10} qatlamlar ustida matematik qayta ishlash.

1. SL va DL vaznlar to'plamidan $w_i^7 \neq 0$ shart asosida W^7 va W^8 vektor elementlari birlashtirilgan holda ($W^7 \cup W^8$) ularning elementlari indeksi bo'yicha sohalari kesimida ajratib olish (filterlash).

W^7 va W^8 vektor elementlarini birlashtirishda W^7 vektori o'zgarishsiz, W^8 vektordagi o'xshash so'zlar vazn koeffitsiyentlariga ko'paytirilgan holdagi yig'indisining o'rta arifmetigi hisoblanadi. Shuningdek, dinamik termlar vaznida sohalarga nisbatan ushbu W^7 va W^8 vektor termlarining umumiy yig'indiga nisbatan ulushi topiladi. Hisoblash uch bosqichda (1) formula yoki bunday murakkab hisoblashlar o'rniga natijasi yaqinroq bo'lgan (2) formula bilan ham amalga oshiriladi:

$$w_i^7 \neq 0 \rightarrow \left\{ \begin{array}{l} F7 = SL(w_i^7) + \frac{DL(w_i^7)}{\sum_{k=1}^m DL(w_k^7)} \\ F8 = \frac{1}{m_{j_l}} \left(\sum_{i=1}^{m_{j_l}} (SL(w_i^{8,j_l}) * t_{f_{ij}}) + \sum_{i=1}^{m_{j_l}} \left(\frac{DL(w_i^{8,j_l})}{\sum_{k=1}^m (DL(w_k^{8,j_l}))} * t_{f_{ij}} \right) \right) \\ FL = \frac{1}{4} \sum_{i=1}^{\bar{n}} (F7_{k,i} + F8_{k,i}), k = \overline{1, p\bar{n}}, FL \in [0,1] \end{array} \right\} (1)$$

$$w_i^7 \neq 0 \rightarrow \left\{ \begin{array}{l} FSL = SL(w_i^7) + \frac{1}{m_{j_l}} \sum_{i=1}^{m_{j_l}} (SL(w_i^{8,j_l}) * t_{f_{ij}}) \\ FDL = \frac{DL(w_i^7)}{\sum_{k=1}^m DL(w_k^7)} + \frac{1}{m_{j_l}} \sum_{i=1}^{m_{j_l}} \left(\frac{DL(w_i^{8,j_l})}{\sum_{k=1}^m (DL(w_k^{8,j_l}))} * t_{f_{ij}} \right) \\ FL_k = \frac{1}{4} \sum_{i=1}^{\bar{n}} (FSL_{k,i} + FDL_{k,i}), \overline{1, p\bar{n}}, FL \in [0,1] \end{array} \right\} (2)$$

2. **PT predmet sohalari termlarini aniqlash.** Bu bosqichda qisqartmalarning predmet sohasini aniqlash kabi natijaga erishiladi. Ma'lumki, $w_i^9 \neq 0$ shart bajarilgan vaqtda ushbu term predmet sohasi aniqlangan bo'ladi. Lekin shuni ham alohida ta'kidlash lozimki, w_i^9 vektorda bir nechta predmet soha termlari bo'lishi mumkinligi uchun ularni $[0, 1]$ intervalga tushirish lozim. Buning uchun dastlab w_i^9 vektordan termlar predmet sohalari bo'yicha mavjudlik soni aniqlanadi. Keyin predmet sohalari bo'yicha eng kattasi topiladi. Demak, ushbu bosqich hisoblashlari (3) dagi ifoda orqali amalga oshiriladi:

$$w_i^9 \neq 0 \rightarrow \left\{ \begin{array}{l} PL_k = \frac{1}{2} \sum_{i=1}^{\bar{n}} \left(PT(w_i^9) + \frac{\sum_{k=1}^{y_{i1}} PTN(w_k^{10, J_i}) * ptf_{ij}}{y_{i1}} \right) \\ a = \max(PL_k), PL_k = \frac{a}{PL_k} \end{array} \right\} (3)$$

Keyingi vazifada statik va dinamik tanlanmalarga tayangan holda termlar va PSga oid termlar hamda ularning yaqinliklarini, ya'ni, $W^7 - W^{10}$ qatlam elementlarini matematik qayta ishlangan natijalarni umumlashtirish, ranjirish va natija foydalanuvchini qanoatlantirsa uni tasdiqlash usuli qaratiladi.

1. **Natijalarni umumlashtirish.** 7, 11 va 12 qadamlarda paydo bo'lgan NL (qisqartmalar), FL , PL bir o'lchamli vektorlarning elementlari soni bir xil bo'lib, bu o'lcham predmet sohalari soni pn ga teng. Natijaviy vektorlar mos vertikal indekslari bo'yicha berilgan mezon asosida umumlashtiriladi:

$$R_k = m_1 NL_k + m_2 FL_k + m_3 PL_k, m_1 + m_2 + m_3 = 100\%, k = \overline{1, np}.$$

bu yerda m_1, m_2, m_3 ekspertlar tomonidan berilgan vazniy koeffitsiyentlar.

2. **Ranjirish va natijani taqdim qilish.** Bunda belgilangan mezondan katta (odatda 70%) va eng yuqori ko'rsatkichga ega bo'lgan m ta (odatda 3 ta) element R_k vektor elementlari orasidan ajratib olinadi.

$$F(R) \rightarrow \max_{count m} (R |_{> 70})$$

Bu esa o'z o'rnida berilgan matn tegishli PS bo'lib hisoblanadi.

3. **Tasdiqlash.** Bu bosqichda DL dinamik termlar to'plami yangilanadi va E – formallashtirilgan murojaatlar to'plamiga yangi element qo'shiladi. Sharti $F(R)$ funktsiya qaytargan qiymatlar (predmet sohalari) foydalanuvchi (ekspert) tomonidan to'g'riligi, ya'ni haqiqatdan ham ushbu murojaat matni sohalari aniq bo'lganda tasdiqlanishi mumkin. Buni amalga oshirish quyidagicha bo'ladi:

a) Foydalanuvchi $F(R)$ dagi sohalarni tanlashi: $EL = F(R) |_{user\ select}$;

b) DL to'plamini yangilash. Faqat EL dagi PS bo'yicha W^7 va W^8 vektordagi elementlarga mos DL termlar ko'rsatkichi bittaga orttiriladi:

$$DL \leftarrow 1 + DL |_{W^7 \cup W^8}^{EL}$$

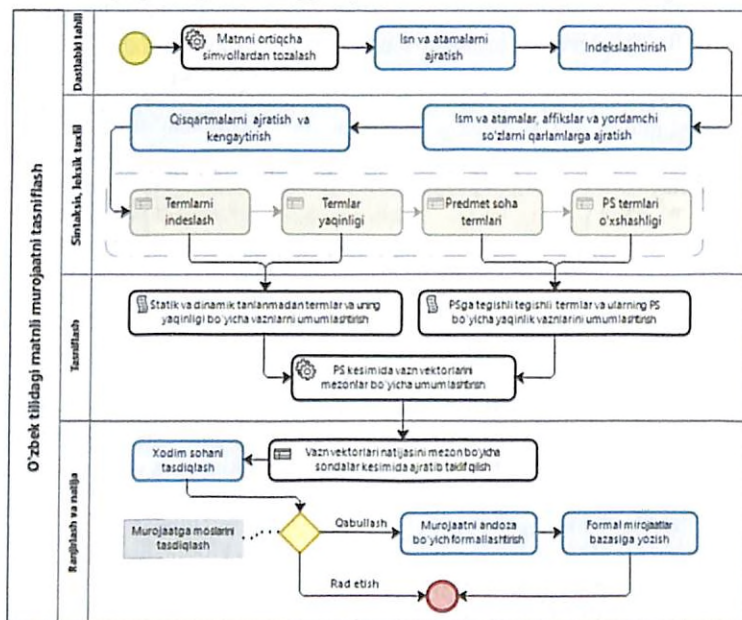
c) E to'plamiga W^{7-10} vektorning nol bo'lmagan elementlari formallashtirilgan holda qo'shiladi. Formallashtirish shunday amalga oshiriladiki, $w_i^7 \neq 0$ bo'lgan indeksdagi termlar bir segmentga birlashtirilib, segmentlar ketma-ketligi tuziladi. Keyin, E to'plamiga segmentlar ketma-ketligi bilan birga EL -PSning indeksi yoziladi. Buni quyidagicha ifodalash mumkin:

$$E \leftarrow (\{w_i^7 \cup w_i^8\} \{w_i^9 \cup w_i^{10}\}) * (\{...\} \{...\}) * \dots * () + EL;$$

d) Murojaatning asl nusxasini M to'plamga qayd qilinadi: $M = M + S$;

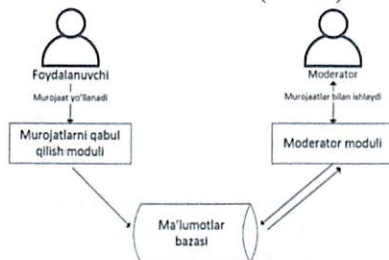
e) Agar natijaga erishilsa, w_i^9 vektorining mos indekslariga pt_k parametrlari (p_i -predmet soha, t_j -termning indeksi) qayd etiladi:

Ishlab chiqolgan algoritmning jarayon sxemasi 3-rasmda tasvirlangan.



3-rasm. O'zbek tilidagi matnlarni tasniflashning jarayon sxemasi

Dissertatsiyaning «Tabiiy o'zbek tilidagi murojaat matnlarni formallashtirish orqali predmet sohasini aniqlash dasturiy vositasi» deb nomlangan to'rtinchi bobida tadqiqot davomida erishilgan natijalar bo'yicha dasturiy vosita ishlab chiqilgan. Jumladan, EHMda murojaatlarni qayta ishlash samaradorligini oshirishni ta'minlovchi tizim yaratish va joriy etish ucun texnika va dasturiy talablar aniqlangan. Dasturiy vosita predmet sohasi murojaatchi tomon, ijrochi tashkilotlar, markaziy apparat va murojaatlarga ishlov berish markazi nuqtai nazaridan loyihalashtirildi. Dasturiy majmua asosiy ikki moduldan, ya'ni: murojaatlarni qabul qilish va murojaatlar bilan ishlash moduli iborat (4-rasm).



4-rasm. Murojaatlar bilan ishlash tuzilmasi

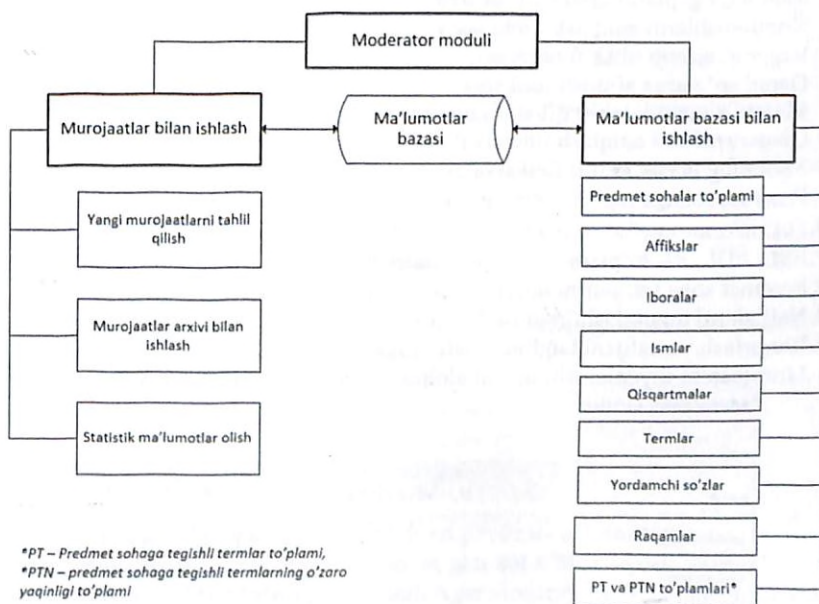
Dasturiy majmuaning asosiy amallarni bajaradigan qismi moderatr moduli hisoblanadi. Moderator moduli ishlari asosiy ikki guruhga ajratilgan bo'lib, uning funkcionalligi 5-rasmda ko'rsatilgan.

Dasturiy majmua veb-ilova ko'rinishida ishlab chiqilgan. Bunda asosiy interfeysi Php va HTML tillarida yozilganligi dasturni istalgan platformada foydalanish imkoniyatini beradi. Dasturning asosiy funksiyalari Python dasturlash tilida yozilganligi asosan foydalanuvchiga mavhum ravishda serverda saqlanadi.

Dasturiy majmua veb-ilova shaklida ishlab chiqilgan. Dastur interfeysi veb-ilova ko'rinishida bo'lib foydalanuvchi undan foydalanish uchun veb brauzerlar qo'llanadi. Bu yesa dasturiy majmuadan istalgan platforma (shaxsiy kompyuter, mobil telefonlar, planshet) orqali foydalanish imkonini beradi. Dasturning asosiy matn bilan moduli Python dasturlash tilida yozilgan. Ma'lumotlar bazasi esa MySql ma'lumotlar bazasini boshqarish tizimida ishlab chiqilgan.

Avtorizatsiya moduli – ushbu modul dasturiy majmua asosiy moduli bo'lib foydalanuvchi (moderator) tizimga kirishini nazorat qiladi. Moderator tizimdan avtorizatsiyadan o'tgan xolda tizimga kiradi.

Moderator moduli – bu moderatorning asosiy moduli bo'lib u asosiy ishini ushbu modulda amalga oshiradi. Moderator moduli bir necha modullarda iborat bo'lib ular murojaatlar bilan ishlash va termlar bilan ishlash moduli.

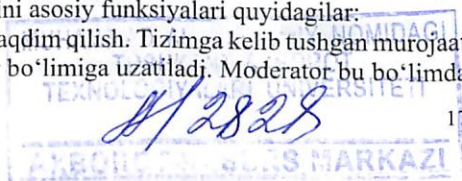


*PT – Predmet sohaga tegishli termlar to'plami,
 *PTN – predmet sohaga tegishli termlarning o'zaro yaqinligi to'plami

5-rasm. Moderator modulining funkcionalligi

Murojaatlar bilan ishlash modulini asosiy funksiyalari quyidagilar:

Yangi kelib tushgan murojaatlarni taqdim qilish. Tizimga kelib tushgan murojaat ma'lumot bazasiga yoziladi va moderator bo'limiga uzatiladi. Moderator bu bo'limda



avtomatik tahlil qilingan murojaatning taklif etilgan predmet sohalaridan birini tanlashi yoki rad qilishi lozim.

Ma'lumotlar bazasidagi murojaatlar bilan ishlash. Bu bo'lim ma'lumotlar bazasiga yozilgan barcha murojaatlar amallarni bajarishga mo'ljallangan.

Statistik ma'lumotlarni olish. Kelib tushgan murojaatlar haqida statistik ma'lumotlar shakllantiriladi.

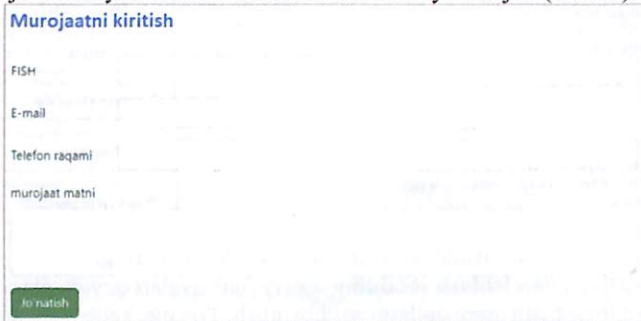
Termlar bilan ishlash moduli dasturiy majmuaning ma'lumotlar bazasidagi termlar bilan ishlash, ya'ni ko'rish, tahrirlash, o'chirish va o'zgartirish amallarini bajarish uchun mo'ljallangan. Dasturiy majmua ishlab chiqilishi paytida ma'lumot bazasi termlar bilan to'ldirilgan, ammo yangi termlar paydo bo'lgan holatda uni ma'lumotlar bazasiga kiritish ushbu modul orqali amalga oshiriladi.

Ma'lumotlar bazasi dasturiy majmuaning asosi bo'lib, unda dastur ishlashi uchun zarur bo'lgan barcha ma'lumotlar (termlar, predmet sohalar, baholar va hokazo) va tizimga kelib tushgan murojaatlar saqlanadi.

Dasturiy majmuaning asosiy moduli murojaatlarni tahlil qilish moduli Python dasturlash tilida yozilgan bo'lib, u bosh funksiya (main) va qo'shimcha funksiyalardan iborat. Qo'shimcha funksiyalar bular ishning uchinchi bobida ko'rsatilgan algoritmnii amalga oshirish uchun mo'ljallangan funksiyalar:

1. Matnni grammatik tozalash funksiyasi;
2. Matndagi gaplarni ajratish funksiyasi;
3. Nomlanishlarni aniqlash funksiyasi;
4. Ismlarni ajratib olish funksiyasi;
5. Gapni so'zlarga ajratish funksiyasi;
6. Matnni sintaktik tahlil qilish funksiyalari;
7. Qisqartmalarni aniqlash funksiyalari;
8. Matnning leksik tahlili funksiyalari;
9. Predmet sohaga tegishli matnni leksik tahlili funksiyalari;
10. Formal murojaatlar to'plamidan qidirish funksiyasi;
11. FSL, FDL, FL to'plamlarini shakllantirish funksiyalari;
12. Predmet soha termlarini aniqlash funksiyasi;
13. Natijalarni umumlashtirish funksiyasi;
14. Ranjirlash va natijani taqdim qilish funksiyasi.

Murojaatchi foydalanishi uchun alohida interfeys mavjud (6-rasm).



Murojaatni kiritish

FISH

E-mail

Telefon raqami

murojaat matni

Jo'natish

6-rasm. Murojaatlarni qabul qilish interfeysi

Moderator uchun alohida interfeysi ishlab chiqilgan (7-rasm).

Moderator moduli Bosh sahifa murojaatlar Termilar Ma'lumotlar bazasi

Yangi murojaatlar ro'yhati

T/r	FISH	E-mail	Telefon	murojaat matni	
1	Azamov Murod	murod@mail.ru	+998919566740	Assalom aleykum. Men TATUGa o'qishga kimgandim. Shu sababli kontrakt to'lovi uchun ta'lim krediti olmoqchiman. Shunga yordam bersangiz. -	
2	Vohidova Komila	komila@gmail.com	+998915986654	Men Jahon tillari universitetida 1 kursda o'qiyman. O'qishimni boshqa OTMga ko'chirmoqchiman. Shunga yordam bersangiz.	
3	Murodov Azamat	murodov@gmail.com	+998995456655	Assalom aleykum. Yangi uy Toshkent shahrida yashayman. Sovuq suv hisoblagichni davlat standartida qaytadan o'tkazmoqchiman.	
3	Saparov Sherzod	sherzod@gmail.com	+998994458789	Assalom aleykum. Men Samarqand shahrida yashayman. Saraton kasallik bilan kasallangandim. Emlanishga uchun iqtisodiy tarafdin yordam berishingizni so'rayman.	

7-rasm. Moderator interfeysi. Yangi murojaatlarni qabul qilish moduli

Yangi murojaat kelib tushgandan keyin dastur avtomat uning sohasini aniqlaydi va natijani ma'lumotlar bazasiga yozib qo'yadi (8-rasm).

Murojaatni tegishli sohaga biriktirish

T/r: 1

FISH: Azamov Murod

E-mail: murod@mail.ru

Telefon raqami: +998919566740

murojaat matni

Assalom aleykum. Men TATUGa o'qishga kimgandim. Shu sababli kontrakt to'lovi uchun ta'lim krediti olmoqchiman. Shunga yordam bersangiz.

Taklif qilinayotgan sohalar:

8-rasm. Murojaat tegishli bo'lgan predmet sohani tanlash

Soha tanlanganda ushbu murojaat uchun shu soha biriktiriladi va ma'lumotlar bazasiga formal murojaat to'plamiga yoziladi. Agar moderator sohalardan birini to'g'ri deb qabul qilmasdan natijani bekor qilsa, u holda unga barcha sohalar ro'yxati taqdim qilinadi va shu sohalardan birini tanlash taklif qilinadi. Soha tanlangandan so'ng mazkur murojaatga aynan shu soha biriktiriladi va formal murojaatlar to'plamiga yoziladi.

XULOSA

Dissertatsiyada elektron hukumat muhitida fuqarolarning tabiiy o'zbek tilida ifodalangan matnli murojaatlarini formallashtirish algoritmlari va dasturiy moduli ishlab chiqildi.

Tadqiqotni amalga oshirishdan olingan asosiy natijalar sifatida quyidagilarni qayd etish mumkin:

1. Elektron hukumat muhitida jismoniy va yuridik shaxslarning elektron ishtiroki va murojaatlarni tahlili, elektron hukumat tizimi rivojlanganligida asosiy hamda natijaviy sharti bo'lgan aholining ommaviy «elektron ishtiroki» hajmi va sifatini oshirish, elektron hukumat doirasida xizmatlarni kengaytirish, aholining barcha qatlamlarining turli daraja va doiradagi muammo, ehtiyoj va «murojaatlar»ni davlat va nodavlat organlariga elektron shaklda yetkaza olishini ta'minlash masalasining dolzarbligini ko'rsatadi.

2. Berilgan matnni qoidalar asosida dastlabki qayta ishlash va qatlamlarga ajratishni amalga oshiradigan matnni tahlil qilish algoritmi ishlab chiqildi. Murojaatni kalit so'zlar bazasi orqali qaysi sohaga tegishli ekanligini aniqlash jarayoning boshlang'ich matematik ko'rinishi murojaat matnini avtomatik formallashtirish o'zbek tilidagi matnlarni avtomatik tanib olish masalasini to'laqonli yechishga imkon beradi.

3. O'zbek tilidagi matnli murojaatlarni avtomatik predmet sohalarga ajratish jarayonini loyihalash murojaatni avtomatik tahlil qilish masalasida jarayonni tashkil qilish uchun zarur belgilashlar va ma'lumotlar bazalari shakllantirildi. Tahlil qilish qoidalari, ekspertlarning xulosalari kabi ma'lumotlar bazalar berilgan matnni anglash uchun xizmat qiladi.

4. O'zbek tilidagi matnli murojaatlarni kalit so'zlarga tayangan holda formal qolipga tushirish uchun mantiqiy matematik ko'rinishdagi formallashtirish algoritmi ishlab chiqildi. Formal murojaatlar bazasidan murojaatni qidirish algoritmi qatlamlashtirish algoritmi natijasida olingan oxirgi to'rtta qatlam asosida shakllantirilgan formal qolipga tushirilgan murojaatlarni mavjud formallashtirish murojaatlar bazasidan qidirish imkonini beradi.

5. Termlarning sohalar kesimida ekspertlar tomonidan berilgan dastlabki statik termlar to'plami va sohalar kesimida dinamik termlar to'plami ishlab chiqildi. Matnli murojaatlarni tasniflashda statik va dinamik tanlanmalar asosida matematik uslubiyat ishlab chiqildi. Qatlamlashtirish algoritmi natijasida olingan oxirgi to'rtta qatlam elementlarini matematik qayta ishlash statik va dinamik tanlanmalarga tayanadi. Olingan natijalar umumlashtirilgandan so'ng ranjirlash uslubiyatlari eng yuqori ko'rsatkichga ega bo'lgan m ta predmet sohani ajratib olish, ushbu predmet sohalar ro'yxatini foydalanuvchiga tanlash uchun taqdim qilishga hamda ish unumdorligining 20-25% gacha ortishiga, ish hajmining esa 10-20% gacha kamayishiga imkon beradi.

НАУЧНЫЙ СОВЕТ DSc.13/30.12.2019.Т.07.01
ПО ПРИСУЖДЕНИЮ УЧЕНЫХ СТЕПЕНЕЙ ПРИ ТАШКЕНТСКОМ
УНИВЕРСИТЕТЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

ТАШКЕНТСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙСАИДРАСУЛОВ ШЕРЗОД НОРБОЙ УГЛИ

АЛГОРИТМЫ И ПРОГРАММНЫЙ КОМПЛЕКС
ФОРМАЛИЗАЦИИ ОБРАЩЕНИЙ ГРАЖДАН НА УЗБЕКСКОМ ЯЗЫКЕ
В СРЕДЕ ЭЛЕКТРОННОГО ПРАВИТЕЛЬСТВА

**05.01.04 – Математическое и программное обеспечение вычислительных машин,
комплексов и компьютерных сетей**

АВТОРЕФЕРАТ
ДИССЕРТАЦИИ ДОКТОРА ФИЛОСОФИИ (PhD) ПО ТЕХНИЧЕСКИМ НАУКАМ

Ташкент-2023

Тема диссертации доктора философии (PhD) по техническим наукам зарегистрирована в Высшей аттестационной комиссии при Кабинете Министров Республики Узбекистан за номером B2022.4.PhD/T.3213

Диссертация выполнена в Ташкентском университете информационных технологий.

Автореферат диссертации на трёх языках (узбекский, русский, английский (резюме)) размещен на веб-странице научного совета (www.tuit.uz) и на Информационно-образовательном портале «Ziyouet» (www.ziyouet.uz).

Научный руководитель:

Нишанов Ахрам Хасанович
доктор технических наук, профессор

Официальные оппоненты:

Мухамедиева Дилноза Тулкуновна
доктор технических наук, профессор

Эшанкулов Хамза Илхомович
доктор философии технических наук (PhD), доцент

Ведущая организация: Каракалпакский государственный университет имени Бердаха

Защита диссертации состоится «17» мая 2023 г. в 16:00 часов на заседании Научного совета DSc.13/30.12.2019.T.07.01 при Ташкентском университете информационных технологий. (Адрес: 100084, г. Ташкент, улица Амира Темура, дом 108. Тел.: (99871) 238-64-70, e-mail: iktuit@tuit.uz).

С диссертацией можно ознакомиться в Информационно-ресурсном центре Ташкентского университета информационных технологий (регистрационный номер №270). (Адрес: 100084, г.Ташкент, улица Амира Темура, дом 108. Тел.: (99871) 238-64-70).

Автореферат диссертации разослан «01» мая 2023 года.
(протокол рассылки №7 от «10» апреля 2023 года).



М.М.Мусаев
Председатель Научного совета по присуждению ученых степеней, доктор технических наук, профессор

Н.О.Рахимов
Ученый секретарь Научного совета по присуждению ученых степеней, доктор технических наук, доцент

Х.Н.Зайнидинов
Председатель Научного семинара при научном совете по присуждению ученых степеней, доктор технических наук, профессор

ВВЕДЕНИЕ (Аннотация диссертации доктора философии (PhD))

Актуальность и востребованность темы диссертации. В соответствии с проводимыми в мире социально-экономическими процессами в нашей стране налажена системная работа по широкому использованию и регулярному совершенствованию информационно-коммуникационных технологий во всех сферах.

С учетом инфраструктуры ИКТ, электронных услуг и факторов человеческого капитала последние 15 лет ООН ведет индекс развития системы электронного правительства, а в рейтинге лидируют такие страны, как Великобритания, Южная Корея, Дания, Норвегия, США, Канада, Швеция, Австралия, Сингапур и Нидерланды¹.

Основным и продуктивным условием развития электронного правительства является массовое «электронное участие» населения. Для повышения объема и качества массового электронного участия необходимо, конечно, расширить услуги в рамках электронного правительства, обеспечить, чтобы все слои населения получили доступ к проблемам, потребностям и «обращениям» в электронном виде в государственные и негосударственные органы.

Важнейшими задачами являются научное решение современных проблем, возникающих при дальнейшем совершенствовании системы электронного правительства в мире, в том числе создание математико-программных моделей и алгоритмов для оптимального управления потоками растущего большого объема информации, интеллектуального анализа и интеграции. Вместе с тем особое внимание необходимо уделить научно-практическим исследованиям вопросов создания и внедрения в практику математически обоснованной методики интеллектуализации процесса оказания услуг в электронном правительстве и повышения массового электронного участия населения.

Данное диссертационное исследование в определенной степени служит выполнению задач, предусмотренных Законами Республики Узбекистан «Об электронном правительстве» (2015) и «Об обращениях физических и юридических лиц» (2017), Указом Президента Республики Узбекистан от 5 октября 2020 года № УП -6079 «Об утверждении Стратегии «Цифровой Узбекистан-2030» и мерах по ее эффективной реализации», Постановлением Президента от 17 февраля 2021 года ПП № 4996 «О мерах по созданию условий для ускоренного внедрения технологий искусственного интеллекта», Постановлениями Кабинета Министров «Об утверждении типового положения о порядке работы с обращениями физических и юридических лиц в органах самоуправления граждан» от 5 января 2018 года №7, «Об утверждении типового положения о порядке работы с обращениями физических и юридических лиц в государственных органах, государственных учреждениях и организациях с государственным участием» от 7 мая 2018 года №341, а также другими нормативно-правовыми документами, принятыми в данной сфере.

¹<https://publicadministration.un.org> (E-Government Development Index)

Соответствие исследования приоритетным направлениям развития науки и технологий республики. Данное исследование выполнено в соответствии с приоритетным направлением развития науки и технологий IV. «Развитие информатизации и информационно-коммуникационных технологий».

Степень изученности проблемы. Многие учёные научно-исследовательских институтов проводили научно-практические исследования по совершенствованию электронного правительства, удовлетворения потребностей потребителей в информации и услугах, повышения качества, повышения эффективности факторов и моделей электронного участия в электронном правительстве, обработке интеллектуального анализа классификации текстовых документов на естественном языке, математическое и программное обеспечение выделения ключевых слов из текстовых документов, а также по разработке, внедрению и развитию методов и технологий.

Заслуживают внимания научные работы иностранных ученых Б.Вирц, Т.Кваснисова, З.В.Архипова, Г.В.Бестолкова, Г.О.Барбаков, Н.В.Берёза, С.В.Пономарев, А.К.Скопинцева, А.В.Шадилова по совершенствованию системы электронного правительства, изучению влияющих факторов на нее, внедрению и анализу системы. Smith & Dalakiouridou, Fedotova, Teixeira, Alvelos, D.Bennet, Revyakin S.A., Y.Panopoulou, A.Makintosh занимались факторами и моделями повышения эффективного электронного участия в электронном правительстве. По вопросам обработки, интеллектуального анализа, классификации текстовых документов на естественном языке работали Kozlov, Batura, Fabritsio Sebastian, Uilyam B. Cavnar, Vandana Korde, Zakaria Elberichi, Ron Bekkerman, вопросы выделения ключевых слов из текстовых документов изучены в исследованиях Стоянка, Москвитина, Гринева, Гринева, Розе, Астраханцев, Захарова, Хохлова, Лукашевича, Логачева. Также такие ученые, как М.М.Камилов, Т.Ф.Бекмуратов, Ш.Х.Фозилов, А.Х.Нишанов, Н.С.Маматов, О.Ж.Бабомурадов, Е.С.Бабаджанов, Л.Бобоев и другие вносят свой вклад в развитие этого направления в Узбекистане.

Проблемы создания методов и алгоритмов, позволяющих автоматически формализовать текстовые обращения, направленные через портал обращений на естественном узбекском языке и классифицировать их по предметным областям в среде электронного правительства. недостаточно изучены, до настоящего времени в должной мере не решены.

Связь диссертационного исследования с планами научно-исследовательских работ высшего образовательного учреждения, где выполнена диссертация. Диссертационное исследование выполнено в рамках научно-исследовательских проектов Ташкентского университета информационных технологий имени Мухаммада-ал Хоразмий №22/19-F «Разработка и внедрение iGov - консультативно-дискуссионно-мониторинговой системы для повышения эффективности использования услуг национальной среды электронного правительства» (2019-2021 гг.)

Цель исследования заключается в разработке алгоритмов и программного обеспечения для автоматической формализации и классификации текстовых

обращений через портал обращений на естественном узбекском языке в среде электронного правительства.

Задачи исследования:

анализ электронного участия и обращений физических и юридических лиц в среде электронного правительства;

разработка формальной постановки задачи классификации текстовых обращений на естественном узбекском языке и их особая обработка;

проектирование автоматизированного (интеллектуального) процесса распределения на предметные сферы текстовых обращений на узбекском языке, разработка соответствующей инфраструктуры обработки;

формализация текстовых обращений на узбекском языке и разработка алгоритма поиска обращения из базы формальных обращений;

разработка базы данных, экспертных и динамических подборов, функций ранжирования, комплекса алгоритмического и программного обеспечения для обработки текстовых данных, характерных для узбекского языка;

Объектом исследования являются требования для работы с обращениями в электронном правительстве, портал обращений, а также база поступивших обращений на естественном узбекском языке.

Предметом исследования является формализация и автоматическое распределение текстовых обращений на предметные сферы в среде электронного правительства, а также методы, алгоритмы и все программные средства, связанные с разработкой программного комплекса.

Методы исследований. В процессе исследования использованы методы математического моделирования, технологии программирования, классификации образов, системного анализа и интеллектуального анализа данных, формирования базы данных и знаний.

Научная новизна исследования заключается в следующем:

спроектирована и создана лингвистическая база данных, которая автоматически обрабатывает тексты на естественном узбекском языке и интегрирует их в единую систему;

разработаны алгоритмы разделения на расширенные слои для классификации текстовых обращений граждан на естественном узбекском языке, критериальный классификационные выборки и функции ранжирования;

разработан метод формализации текстов и алгоритм комбинированного поиска базы формальных обращений;

разработан программный комплекс на основе алгоритмов классификации и формализации обращений граждан для систем электронного правительства;

разработана структура программно-функциональных модулей для систем электронного правительства на основе алгоритмов классификации и формализации обращений граждан.

Практические результаты исследования заключаются в:

разработке методов в среде электронного правительства, обеспечивающих электронное участие физических, юридических лиц и обращений, классифицирующих текстовые обращения на естественном узбекском языке,

формализующих обращения и обслуживающих их в соответствии с принципами определения особенностей обработки;

проектировании процесса автоматического распределения текстовых обращений на узбекском языке на предметные сферы, разработке программного комплекса с учетом особенностей инфраструктуры обработки данных;

разработке алгоритмического обеспечения поиска обращения из базы формальных обращений и формализации текстовых обращений на узбекском языке;

разработке программного средства для определения предметной сферы обращений на естественном узбекском языке, в процессе обработки текстовых данных на основе баз данных особенностей узбекского языка, экспертных и динамических подборов, методов ранжирования и алгоритмического обеспечения путем формализации текстов.

Достоверность результатов исследования подтверждается глубоким анализом поставленной проблемы, корректностью математического моделирования, применением технологий программирования, теории алгоритмизации, методов системного и интеллектуального анализа, применением технологий интеллектуализации при многопрофильном поиске потребностей в базе услуг и формализации данных, выраженных на естественном узбекском языке, а также результатами теоретических и практических исследований по электронному обслуживанию и соответствии требованиям разработанной веб-платформы.

Научная и практическая значимость результатов исследования. Научная значимость результатов состоит в проектировании процесса формализации и автоматического распределения текстовых обращений на естественном узбекском языке в среде электронного правительства, создании соответствующей инфраструктуры обработки, разработки методов и алгоритмов классификации, формализации текстовых обращений на естественном узбекском языке и разработке алгоритма поиска обращения из базы формальных обращений, специфической базы данных узбекского языка при обработке текстовых данных, разработке экспертных и динамических подборов, методов ранжирования и алгоритмического обеспечения.

Практическая значимость результатов исследования определяется разработкой программного комплекса определения предметной сферы путем формализации текстов и заключается в обеспечении электронного участия и обращений физических и юридических лиц в среде электронного правительства, методов, классифицирующих текстовые обращения на естественном узбекском языке, формализующих обращения и обслуживающих их в соответствии с принципами определения особенностей их обработки, проектирования процесса автоматического распределения текстовых обращений на предметные области, соответствующих инфраструктуре обработки данных, формализации текстовых обращений на естественном узбекском языке и поиск обращения из базы формальных обращений, специфической базы данных узбекского языка, экспертной и динамической выборки на основе методов ранжирования и алгоритмического обеспечения обращения на естественном узбекском языке.

Внедрение результатов исследования. На основе научных результатов, полученных по разработанному методу и алгоритмам для автоматической формализации и классификации по предметным областям текстовых обращений на естественном узбекском языке в среде электронного правительства:

Программное обеспечение, разработанное исследователем, было применено в центральном управлении при виртуальной приёмной с апреля 2022 года, что позволило автоматическому распределению 32% из почти 450000 обращений, поступивших за эти 5 месяцев. Эти результаты показывают, при увеличении типов и количества обращений, расширяется база знаний и, как следствие, внедрение повышает эффективность (Справка министерства по развитию информационных технологий и коммуникаций № 33-8/5583 от 30 августа 2022 г.).

Программное обеспечение, разработанное исследователем, может быть использовано в среде электронного правительства в автоматической обработке обращений соответствующими специалистами, чтобы помочь им принять разумное решение. В результате достигается увеличение производительности труда в среднем на 20-30%. Использование программного комплекса диссертационной работы в информационных системах электронного правительства Республики Узбекистан при решении вопросов автоматической формализации текстовых обращений на естественном узбекском языке и классификации по предметным областям способствует повышению эффективности работы (Акт центра научно-технических и маркетинговых исследований «UNICON.UZ» № 5-3/1465 от 22 августа 2022 г.).

Программное обеспечение, разработанное исследователем, может быть использовано в среде электронного правительства в автоматической обработке обращений соответствующими специалистами, чтобы помочь им принять разумное решение. В результате достигается увеличение производительности труда в среднем на 20-30% (АКТ АО «Узбекистон почтаси» от 18 августа 2022 г.).

В дальнейшем, при внедрении данного программного комплекса в среду электронного правительства за счёт автоматического определения программным инструментом предметных областей электронных обращений граждан будет достигнуто повышение производительности труда модераторов до 20-25%, а снижения объема работы до 10-20% (АКТ государственного учреждения «Центр управления проектами электронного правительства» от 22 августа 2022 г.).

Апробация результатов исследования. Результаты исследования докладывались и обсуждались на 2 международных и 3 национальных научных конференциях.

Публикация результатов исследования. Всего по теме исследования опубликовано 28 научных работ, из них 12 статей опубликованы в научных изданиях, рекомендованных ВАК Республики Узбекистан, в том числе 2 в зарубежных и 10 в отечественных журналах и 3 свидетельства на программное обеспечение для ЭВМ.

Структура и объём диссертации. Диссертация состоит из введения, четырех глав, заключения, списка использованной литературы и приложений. Объем диссертации составляет 120 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Вводная часть основывается на актуальности и необходимости диссертации, соответствии исследования приоритетам науки и техники Республики Узбекистан, поставленным целям и задачам, объекту и предмету исследования, достоверности результатов, их теоретическое и практическое значение, внедрение результатов исследований, сведения о состоянии, опубликованных работах и структуре диссертации.

Первая глава диссертации под названием «Состояние организации обращений и анализ нормативных правовых актов в среде электронного правительства» состоит из четырех параграфов. Данная глава направлена на усиление роли общественного контроля, институтов гражданского общества, повышение открытости и качества предоставления государственных услуг посредством информационно-коммуникационных технологий. Первый параграф содержит анализ развития цифровой экономики в Узбекистане, анализ уровня развития электронного правительства в нашем государстве по отношению к развитым государствам на основе анализа развития системы «Электронное правительство», проводимого ООН, анализ передовых государств и государств СНГ по рейтингу развития электронного правительства (EGDI) и индексу электронного участия (epart), приведены факторы и модели повышения эффективности электронного участия в электронном правительстве.

В частности, в течение следующих 14 лет рейтинг EGDI вырос со 109 до 69 (+40), а соответствующий показатель электронного участия вырос с 0,0909 до 0,7636 (+0,6727). Во втором параграфе приведены понятия и нормативные правовые акты по обращениям граждан. В третьем параграфе изложены основные показатели и процесс организации электронных обращений, функциональные особенности портала «Мурожаатлар», опыт зарубежных стран, в том числе СНГ, по организации работы с обращениями, а также подходы к оптимизации процесса классификации обращений на портале «Мурожаатлар». Четвертый параграф содержит постановку вопроса, а также выводы по первой главе.

Во второй главе диссертации под названием «Задачи автоматической обработки текстовых обращений и их анализ» были проанализированы исследовательские работы и алгоритмы обработки текстовых документов и обращений на узбекском языке, а также классифицированы существующие проблемы. В частности, раскрыта сущность алгоритмов интеллектуального анализа текста с классификацией. Поскольку ключевые слова становятся все более важными при обработке текстовых ссылок, методы и алгоритмы извлечения ключевых слов были исследованы в отдельном параграфе. В частности, обосновано на основе систематизации публикаций и выбора перспективных направлений исследований в проблемной области вывода ключевых слов, разработанных методах и алгоритмах их реализации (рис.1).



Рис.1. Методы вывода ключевых слов

Естественных языков много и из-за разнообразия их формальных формирований алгоритмы обработки текста обычно создаются по отношению к определенному естественному языку. Следовательно, многие существующие алгоритмы созданы в соответствии с более популярными естественными языками и их применение к узбекскому языку без модификаций не работает. Поэтому как результат аналитической главы были исследованы методики формализации текстов на естественном узбекском языке, направленной на обработку текстов на узбекском языке.

В третьей главе диссертации под названием «Алгоритм определения предметной области посредством формализации текстов обращений на естественном узбекском языке» исследована методика и алгоритмическое решение определения предметных областей текстовых обращений на естественном узбекском языке. В предыдущем параграфе работы основные задачи идеи решения вопроса о распределении текстовых обращений на узбекском языке на соответствующие предметные области изложены следующим образом:

создание базы терминов и правил по заданному вопросу;

Разработан алгоритм анализа текста. В данном алгоритме осуществляется первичная обработка и наложение заданного текста. Включая очистку текста; разделение текста на предложения и слова; разделение имен, названий и чисел; развёртывание сокращений; синтаксический и лексический анализ;

Поиск текста по формальному шаблону. При этом текст, опираясь на ключевые слова, в логическом математическом представлении сводятся к формальному шаблону. В результате текущий формальный текст ищется в уже существующей базе формальных обращений. Если результат достигнут, операция прекращается, иначе цикл продолжается;

Разработка метода классификации. В результате разделения на текущие слои полужормализованный текст по набору экспертных и динамических выборок определяется предметная область и представляется пользователю;

Результат процесса утверждения формализации текста предметной области обращения. Когда результат алгоритма верен, текущий формальный текст записывается в базу данных формальных обращений.

Исходя из поставленных задач автоматического анализа обращения изначально создаются необходимые обозначения и базы данных. Это следующие:

T – набор терминов (выражений) на узбекском языке: $t = \{t_1, t_2, \dots, t_n\}$, n – количество доступных терминов. Слова, относящиеся в основном к самостоятельной категории слов.

C – совокупность категорий слов и их характеристик.

R – набор цифр: $R = \{r_1, r_2, \dots, r_m\}$, m – количество доступных цифр.

Q – набор аббревиатур: $Q = \{q_1, q_2, \dots, q_n\}$, q_n – количество аббревиатур.

Элементы набора многопараметрические: $q_i = \langle \bar{q}_i^1, \bar{q}_i^2, \bar{q}_i^3 \rangle$. Вот \bar{q}_i^1 – различные формы аббревиатуры, \bar{q}_i^2 – расширение, \bar{q}_i^3 – соответствующая область

G – набор вспомогательных слов (вспомогательные, связующие слова, предлоги).

H – набор аффиксов $H = \{h_1^{o^1}, h_2^{o^2}, \dots, h_{hn}^{o^{hn}}\}$, hn – число. Здесь $h_i^{o^l}$ – это аффикс, и то, является ли он передним или задним дополнением, определяется двоичным $o^l = \langle 0, 1 \rangle$ (если 0, то аффикс является предлогом слова, а если 1, то аффикс является задним дополнением слова).

TS – совокупность производных существительных (организация, адрес) из нескольких фраз: $TS = \{ts_1, ts_2, \dots, ts_{tsn}\}$, tsn – количество доступных определений.

FS – совокупность лиц, взято из книги имен: $FS = \{fs_1, fs_2, \dots, fs_{tsn}\}$, fsn – количество существующих имён.

В работе **T, R, H, G, Q**, называются базой ключевых слов (БКС).

TN – совокупность степени близости терминов:

$TN = \{\langle \bar{t}_i, \bar{t}_j \rangle, tf_{ij}\}; \forall \bar{t}_i, \bar{t}_j \in T; tf_{ij} \in (0, 1]; i, j \leq n; i \neq j$, здесь tf_{ij} – данные \bar{t}_i и \bar{t}_j близкие по смыслу. Если выполняется условие $tf_{ij} = 1$, в этом случае $\langle \bar{t}_i, \bar{t}_j \rangle$ термины являются синонимами.

P – совокупность предметных областей (ПО): $P = \{p_1, p_2, \dots, p_{pn}\}$, pn – количество областей.

PT – совокупность терминов, относящихся к PS: $PT = \{pt_k = \langle p_i, t_{j_i} \rangle; i \leq pn; j_i \leq n; pt_k \leq n \times pn; \forall p_i \in P, t_{j_i} \in T$. Здесь p_i – предмет области, t_{j_i} – индекс термина.

PTN – совокупность степени близости среди терминов PS: $PTN = \{\langle \overline{pt}_i, \overline{pt}_j \rangle, ptf_{ij}\}; \forall \overline{pt}_i, \overline{pt}_j \in PT; ptf_{ij} \in (0, 1]; i, j \leq n \times pn; i \neq j$. Если $ptf_{ij} = 1$ бо'lsa, $\langle \overline{pt}_i, \overline{pt}_j \rangle$ термины области условно называются синонимами.

M – совокупность оригинальных текстовых обращений.

E – формализованные обращения: $E^l = (E^{l_1}, \dots, E^{l_{nh}})$. В этом случае ссылка состоит из нескольких сегментов, а сегменты состоят из двух параметров: терминов и терминов PS.

$E^{i,j} = (e_1^{i,j}, \dots, e_{i_j}^{i,j}); e_{i_j}^{i,j} = \{(t_{k_1}^{i,j}, \dots, t_{kn}^{i,j}), (pt_{h_1}^{i,j}, \dots, pt_{nh}^{i,j})\}, kn < N, nh < M$.

SL – совокупность экспертных первичных статичных терминов в разрезе терминов областей: $SL = \{sl_k, p_i, t_j, \gamma_{ij}\}; i \leq pn; j \leq n; k \leq n \times pn; \forall p_i \in P, t_j \in T; \gamma_{ij} \in [0, 1]$. Значение элементов заранее известны этого набора, которое принимается после машинного чтения и уточнения экспертами

определённого количества обращений, а также принимается в качестве неизменяемого ориентира.

DL – совокупность динамических терминов в разрезе областей: $DL = \{dl_k, p_i, t_j, \delta_{ij}\}$; $i \leq pn$; $j \leq n$; $k \leq n \times pn$; $\forall p_i \in P, t_j \in T$; $\delta_{ij} \in [0, 1]$. Совокупность динамических терминов является копией набора статических терминов, идентичен по размеру и структуре. Но после машинного чтения текущих обращений значение e_{ij} в **DL** постоянно увеличивается.

Вышеуказанные объекты БД можно разделить на несколько классов по их взаимному сходству на основе реляционной модели (рис.2).



Рис.2. Классы объектов базы данных

На основе правил алгоритма автоматического анализа текста разбивает его на слои. Этот алгоритм в соответствии с БКС разбивает на несколько слоев. При формировании каждого слоя в процессе форматирования текста вырабатывается свой метод или алгоритм анализа данного текста. При этом исходный текст обрабатывается в 9 этапов, создавая 10 слоев. Во время наложения вводится двумерный вектор W^v , который на каждом этапе изменяет свое вертикальное и горизонтальное разрешение. Слои алгоритма можно представить следующим образом: W^0 -неопределенные символы; W^1 -имена; W^2 -имена; W^3 -цифры; W^4 -вспомогательные слова; W^5 -аффиксы; W^6 -сокращения; W^7 -термины; W^8 -сходство терминов; W^9 -сходство терминов PS; W^{10} -сходство терминов PS.

На основе последних четырех слоев, полученных в результате алгоритма, выполняется метод поиска текста по формальному шаблону. То есть формальная форма формируется на основе слоев $W^7 - W^{10}$. В результате этого выполняется поисковая практика по уже существующей базе формализованных обращений.

Формализованное обращение к уже существующему сборнику Е-формальных обращений будет организовано в этой форме: в последовательности сегментов $(e_{ij}^{i,j})$, в паре малых сегментов с разделением терминов сегментов и терминов PS (t, pt) , а в малых сегментах - последовательность оригиналов терминов.

В случае, если текущее рассматриваемое текстовое обращение считается формальным при условии, что термины и его синонимы разделены на отдельный сегмент, их сходства на отдельные сегменты терминов PS. Также элементы производимого сегмента $(w_i^7 \cup w_i^8 \text{ и } w_i^9 \cup w_i^{10})$, состоящие из индексов терм, выражают его следующим образом:

$$B = (b^1, b^2, \dots, b^r), \quad b^d = \{(t_{k_1}^d, t_{k_2}^d, \dots, t_{k_d}^d), (p_{h_1}^d, p_{h_2}^d, \dots, h_{h_d}^d)\}.$$

В большинстве случаев второй сегмент вектора b^d будет $h_n^d = \emptyset$. Поскольку выполнение условия $w_i^7 \neq 0$ не всегда означает, что этот термин относится к *PS*.

Обобщённо вопрос поиска можно описать в этом виде: требуется поиск векторных элементов B текущей формальной строки из формализованного набора обращений E . При разработке алгоритма поиска предварительно определяются следующие условия и критерии:

при поиске не важны позиции расположения сегментов и позиции элементов соответствующие этому сегменту;

поиск ведётся по соответствию пары сегментов;

если при поиске количество сегментов и их элементов соответствует заданным критериям, то принимается как результат;

Множество E имеет вторую пару добровольного сегмента и может быть найден при наличии/отсутствии второй пары в сегменте искомого вектора B ;

Если отсутствует вторая пара добровольного сегмента набора E и есть вторая пара в векторном сегменте B , поиск по второму сегменту не производится;

Если количество сегментов любой строки превышает количество сегментов искомого B вектора, то критерий определяется по количеству векторных сегментов B , в противном случае ведётся критерий по соотношению числа векторных сегментов B к числу сегментов записи E .

Проще говоря, если $w_i^7 \neq 0$, то $w_i^7 \cup w_i^8$ термы E ищутся из подмножества малых термов E множества и $w_i^9 \cup w_i^{10}$ термы множества E ищут параллельно из подмножества малых термов *PS*. При поиске сегментов местоположение позиций не имеет первостепенного значения. Результат зависит от степени их покрытия относительно количества всех имеющихся сегментов. Например, пусть ключевые слова в 5 позициях текста отвечают 4 из 7 сегментных шаблонов формального набора. Если указанный критерий равен 70%, это означает обращение в формальном шаблоне. Это рассматривается как результат текущего процесса, и предоставляется пользователю, указанный *PS* в наборе формальных обращений E .

В следующей задаче разрабатываются статические и динамические варианты классификации текстовых обращений. Методика выполняется в два этапа, а именно: математическая обработка слоев W^7 и W^8 , также W^9 и W^{10} на основе выборки *SL* и *DL*.

1. Извлечение (фильтрация) из весовых множеств *SL* и *DL* с условием $w_i^7 \neq 0$ с присоединением векторных элементов W^7 и W^8 в сечении полей по индексу элементов ($W^7 \cup W^8$).

При сложении элементов W^7 и W^8 , вектор W^7 не меняется, вектор W^8 является средним арифметическим суммы аналогичных слов умноженной на весовые коэффициенты. Также вычисляется относительная доля динамических терминов в разрезе областей относительно общей суммы W^7 и W^8 . Расчет производится в три этапа (1) по формуле или вместо таких сложных вычислений также можно применить выражение (2) для близкого результата:

$$w_i^7 \neq 0 \rightarrow \left\{ \begin{array}{l} F7 = SL(w_i^7) + \frac{DL(w_i^7)}{\sum_{k=1}^{pn} DL(w_k^7)} \\ F8 = \frac{1}{m_{ji}} \left(\sum_{l=1}^{m_{jl}} (SL(w_i^{8,jl}) \cdot tf_{lj}) + \sum_{l=1}^{m_{jl}} \left(\frac{DL(w_i^{8,jl})}{\sum_{k=1}^{pn} (DL(w_k^{8,jl}))} \cdot tf_{lj} \right) \right) \\ FL = \frac{1}{4} \sum_{i=1}^n (F7_{k,i} + \bar{F}8_{k,i}), k = \overline{1, pn}, FL \in [0,1] \end{array} \right\} (1)$$

$$w_i^7 \neq 0 \rightarrow \left\{ \begin{array}{l} FSL = SL(w_i^7) + \frac{1}{m_{ji}} \sum_{l=1}^{m_{jl}} (SL(w_i^{8,jl}) \cdot tf_{lj}) \\ FDL = \frac{DL(w_i^7)}{\sum_{k=1}^{pn} DL(w_k^7)} + \frac{1}{m_{ji}} \sum_{l=1}^{m_{jl}} \left(\frac{DL(w_i^{8,jl})}{\sum_{k=1}^{pn} (DL(w_k^{8,jl}))} \cdot tf_{lj} \right) \\ FL_k = \frac{1}{4} \sum_{i=1}^n (FSL_{k,i} + FDL_{k,i}), \overline{1, pn}, FL \in [0,1] \end{array} \right\} (2)$$

2. Определение терминов предметных областей PT. На этом этапе определяется предметная область сокращений. Известно, что в момент выполнения условия $w_i^9 \neq 0$ это поле термина является определённым. Но также следует особо отметить, что, поскольку в векторе w_i^9 может быть несколько терминов предметной области, необходимо уменьшить интервал до $[0, 1]$. Для этого сначала из вектора w_i^9 определяется количество присутствий в предметных областях терминов. Затем субъект находит наибольшее по полям. Следовательно, вычисления этого шага выполняются через ifida (3):

$$w_i^9 \neq 0 \rightarrow \left\{ \begin{array}{l} PL_k = \frac{1}{2} \sum_{i=1}^n \left(\frac{PT(w_i^9) + \frac{\sum_{k=1}^{y_{ji}} PTN(w_k^{10,jl}) \cdot pt_{fij}}{y_{ji}}}{a} \right) \\ a = \max(PL_k), PL_k = \frac{a}{PL_k} \end{array} \right\} (3)$$

В следующей задаче рассматриваются термины и термины, на основе статической и динамической выборки, относящиеся к PS, а также их сходства, то есть суммирование, ранжирование, проверка математически обработанных результатов элементов слоя $W^7 - W^{10}$ пока результат не удовлетворит пользователя.

1.Подведение итогов. Количество элементов одномерных векторов NL, FL, PL которые появляются на шагах 7, 11 и 12 выше, одинаково, и это измерение равно количеству предметных областей pn . Результирующие векторы суммируются на основе критерия, заданного соответствующими вертикальными индексами:

$R_k = m_1 NL_k + m_2 FL_k + m_3 PL_k, m_1 + m_2 + m_3 = 100\%, k = \overline{1, pn}$,
здесь m_1, m_2, m_3 весовые коэффициенты, обозначенные экспертами.

2.Ранжировка и представление результата. Здесь из векторных элементов R_k выделяется элемент m (обычно 3) с наибольшим значением и больше указанного критерия (обычно 70%).

$$F(R) \rightarrow \max_{count\ m} (R \Big|_{> 70})$$

А также этот текст относится к PS.

3.Подтверждение. На этом этапе обновляется набор динамических терминов DL и добавляется новый элемент в E множество формализованных обращений. Подтверждение возможно при выполнении условия значения возвращаемые функцией F (R) (предметные области) проверены пользователем (экспертом) на правильность, то есть соответствие текстового обращения к предметной области. Это выполняется в следующем виде:

а) Пользователь выбирает предметную область из $F(R)$:

$$EL = F(R) \Big|_{user\ select}$$

б) Обновление множества DL. Только показатель степени DL, соответствующий элементам в векторе W^7 и W^8 на PS в EL, увеличивается на единицу: $DL \leftarrow 1 + DL \Big|_{W^7 \cup W^8}^{EL}$;

в) К множеству E формально добавляются ненулевые элементы вектора W^{7-10} . Формализация выполняется таким образом, что термы индекса $w_i^7 \neq 0$ объединяются в один сегмент, образуя последовательность сегментов. Затем индекс el-PS записывается в E вместе с последовательностью сегментов. Это можно выразить как:

$$E \leftarrow ((w_i^7 \cup w_i^8) [w_i^9 \cup w_i^{10}]) \cdot ((\dots) [\dots]) \cdot \dots \cdot () + EL;$$

д) Копия оригинального обращения записывается во множестве M: $M = M + S$;

е) Если результат достигнут, в соответствующие индексы вектора w_i^9 записываются параметры pt_k (p_i -предметная область, t_{ji} -индекс терма):

Технологическая схема разработанного алгоритма изображена на рисунке 3.

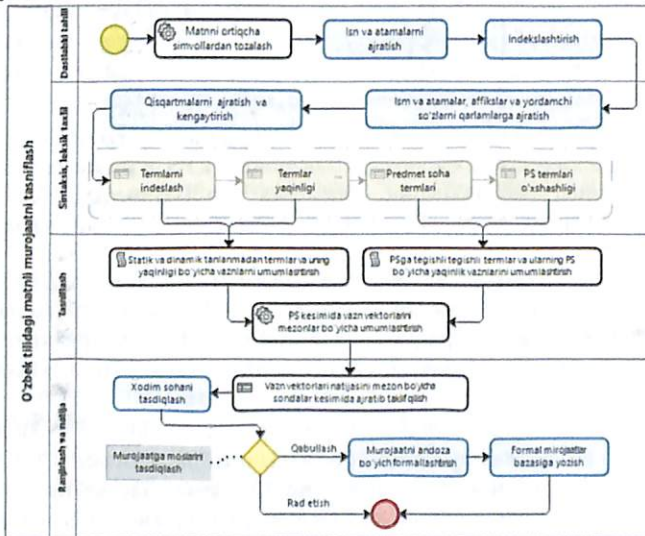


Рис.3. Схема процесса классификации текстов на узбекском языке

В четвертой главе диссертации, под названием «Программное средство определения предметной области посредством формализации текстов обращений на естественном узбекском языке» как результат исследования был разработан программный продукт. В частности, определены технические и программные требования к созданию и внедрению системы, обеспечивающие повышение эффективности обработки обращений с помощью ЭВМ. Программное средство определения предметной области спроектировано с точки зрения заявителя, исполнительных организаций, центрального аппарата и центра обработки обращений. Программный комплекс состоит из двух основных модулей: модуль приема обращений и модуль работы с обращениями (рис.4).

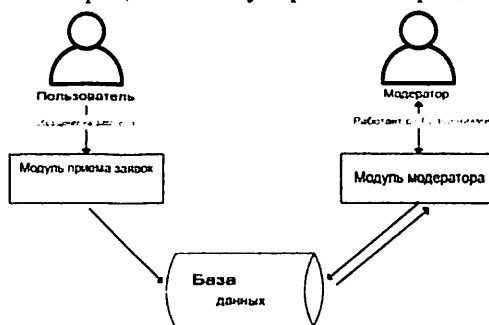


Рис.4. Схема работы с обращениями

Основная часть выполняемых действий программного комплекса выполняет модуль модератора. Работа модуля модератора разделена на две основные группы, функциональность которых показана на рисунке 5.

Программный комплекс разработан в виде веб-приложения. При этом тот факт, что основной интерфейс написан на языках PHP и HTML, позволяет использовать программу на любой платформе. Основные функции программы написаны на языке программирования Python и в основном скрыты от пользователя и хранятся на сервере.

Программный комплекс разработан в виде веб-приложения. Интерфейс программы представлен в виде веб-приложения и пользователю достаточно иметь веб-браузер для использования. Это позволяет пользователю с помощью любой платформы (персональный компьютер, мобильные телефоны, планшет) иметь возможность пользования. Модуль программы с основным текстом написан на языке программирования Python. База данных же разработаны в системе управления базами данных MySQL.

Модуль авторизации – этот модуль является основным модулем программного комплекса, который контролирует вход пользователя (модератора) в систему. Модератор входит в систему, пройдя авторизацию.

Модуль модератора-это основной модуль модератора, который выполняет свою основную работу в этом модуле. Модуль модератора состоит из нескольких модулей, которые являются модулем для работы с обращениями и модулем для работы с терминами.

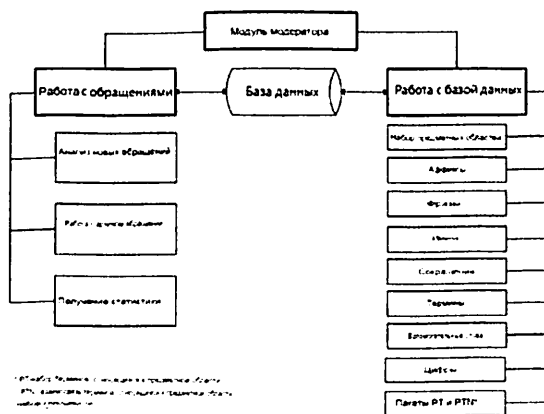


Рис.5. Функциональность модуля модератора

Модуль работы с обращениями. Основными функциями модуля работы с обращениями являются:

Предоставлять новые поступившие обращения. Поступившее в систему обращение записывается в базу данных и передается в отдел модераторов. Модератор должен выбрать или отклонить одну из предложенных автоматически проанализированных предметных областей обращения.

Работа с базой данных обращений. В этом разделе выполняются предназначенные возможные действия над обращениями.

Получение статистических данных. Формируется статистика поступивших обращений.

Модуль работы с терминами предназначен для работы с терминами в базе данных программного комплекса, то есть для выполнения операций просмотра, редактирования, удаления и изменения. Во время разработки программного комплекса база данных заполняется терминами, но в случае появления новых терминов ее ввод в базу данных осуществляется через этот модуль.

База данных основа программного комплекса, в которой хранится вся информация, необходимая для работы программы (термины, предметные области, оценки и т. д.) и ссылки, поступающие в систему.

Базовый модуль программного комплекса модуль анализа ссылок написан на языке программирования Python и состоит из базовой функции (main) и дополнительных функций. Дополнительные функции это функции, предназначенные для реализации алгоритма, указанного в третьей главе работы:

1. Функция грамматической очистки текста;
2. Функция разделения предложений в тексте;
3. Функция распознавания имен;
4. Функция извлечения имен;
5. Функция разбиения предложения на слова;
6. Функции синтаксического анализа текста;
7. Функции определения сокращений;

8. Функции лексического анализа текста;
 9. Функции лексического анализа текста, относящиеся к предметной области;
 10. Функция поиска из базы формальных обращений;
 11. Функции формирования множеств FSL, FDL, FL;
 12. Функция определения терминов предметной области;
 13. Функция суммирования результатов;
 14. Функция ранжирования и представления результата.
- Для заявителя существует отдельный интерфейс (рис.6).

Рис.6. Интерфейс приёма обращений

Разработан отдельный интерфейс для модератора (рис.7).

Moderator moduli Bosh sahifa Murojaatlar Terminlar Hisobotlar bazasi

Yangi murojaatlar ro'yxati

T/r	FISH	E-mail	Telefon	murojaat matni
1	Azamov Murod	murod@mail.ru	+998919566740	Assalom aleykum. Men TATUga o'qishga kirdim. Shu sababli kontrakt to'lovi uchun ta'lim krediti olmoqchiman. Shunga yordam bersangiz.
2	Vafedova Kamola	kamola@gmail.com	+99891508654	Men sahon filiali universitetida 1 kursda o'qiyman. O'qishimni boshqa OTMga ko'chirmoqchiman. Shunga yordam bersangiz.
3	Murodov Azamat	murodov@gmail.com	+99899545655	Assalom aleykum. Yangi uy toshkent shahrida yashayman. So'ziz kirib keladigan davlat standartida qaytadan o'rnatmoqchiman.
4	Saparov Sherzod	sherzod@gmail.com	+99899445789	Assalom aleykum. Men Samarqand shahrida yashayman. Sarafin kasallik bilan kasallangandim. Endiranga uchin o'qibodiy tashkilot yordam bersangizmi va hayman.

Рис.7. Интерфейс модератора. Модуль принятия новых обращений

После поступления нового обращения программа автоматически определяет его предметную область и записывает результат в базу данных (рис.8).

Murojaatni tegishli sohaga biriktirish

T/r: 1

FISH: Azamov Murod

E-mail: murod@mail.ru

Telefon raqami: +998919566740

murojaat matni

Assalom aleykum. Men TATUga o'qishga kirdim. Shu sababli kontrakt to'lovi uchun ta'lim krediti olmoqchiman. Shunga yordam bersangiz.

Taklif qilinayotgan sohalar:

Рис.8. Выбор соответствующей области обращения

Когда область выбрана, это поле прикрепляется к этому обращению и записывается в базу данных формальных обращений. Если модератор отменяет результат, не считая предложенные области правильными, то ему

предоставляется список всех областей и предлагается выбрать одно из них. После выбора области к данному обращению прикрепляется соответствующая область и записывается в базу формальных обращений.

ЗАКЛЮЧЕНИЕ

В диссертации разработаны алгоритмы и программный комплекс формализации обращений граждан на естественном узбекском языке в среде электронного правительства.

В качестве основных результатов, полученных в результате проведения исследования, можно отметить:

1. Электронное участие физических и юридических лиц в среде электронного правительства и анализ обращений, повышение объемов и качества массового «электронного участия» населения, являющегося основным и результативным условием развития системы электронного правительства, расширение услуг в рамках электронного правительства, предоставление всем слоям населения в электронном виде государственных и негосударственных органов информации о проблемах, потребностях и «обращениях» различного уровня и объема показывает актуальность вопроса.

2. Разработан алгоритм анализа текста, осуществляющий первичную обработку и раслоение заданного текста. Автоматическое форматирование текста обращения позволяет полноценно решить задачу автоматического распознавания текстов на узбекском языке.

3. Проектирование процесса автоматического распределения обращений на узбекском языке на предметные области в вопросе автоматического анализа обращений сформированы необходимые обозначения и базы данных для организации процесса. Такие базы данных, как правило, служат для анализа, заключения экспертов, понимания данного текста.

4. Был разработан алгоритм формализации в логико-математическом виде, позволяющий формализовать обращения на узбекском языке в формальный шаблон, основанного на ключевых словах. Алгоритм поиска обращений из базы формальных обращений позволяет искать в существующей базе формализованных обращений формализованные обращения, сформированные на основе последних четырех слоев, полученных в результате алгоритма раслоения.

5. Был разработан исходный набор статических и динамических терминов, заданных экспертами на пересечении областей терминов. При классификации текста обращений была разработана математическая методика, основанная на статическом и динамическом отборе. Математическая обработка последних четырех элементов слоя, полученных в результате алгоритма наслоения, основана на статической и динамической выборке. После суммирования полученных результатов с помощью ранжирования выделяется m предметных областей с наивысшими показателями, что позволяет сделать список предметных областей доступным пользователю для выбора.

**SCIENTIFIC COUNCIL AWARDING SCIENTIFIC DEGREES
DSc.13/30.12.2019.T.07.01 AT TASHKENT UNIVERSITY OF
INFORMATION TECHNOLOGIES**

TASHKENT UNIVERSITY OF INFORMATION TECHNOLOGIES

SAIDRASULOV SHERZOD NORBOY O'G'LI

**ALGORITHMS AND SOFTWARE PACKAGE FOR FORMALIZING
CITIZENS' APPEALS IN UZBEK LANGUAGE IN THE
E-GOVERNMENT ENVIRONMENT**

05.01.04 – Mathematical and software of computers, complexes and computer networks

**DISSERTATION ABSTRACT OF THE DOCTOR OF PHILOSOPHY (PhD)
ON TECHNICAL SCIENCES**

Tashkent – 2023

The theme of doctor of philosophy (PhD) on technical sciences was registered at the Supreme Attestation Commission at the Cabinet of Ministers of the Republic of Uzbekistan under number B2022.4.PhD/T.3213

The dissertation has been prepared at Tashkent University of Information Technologies.

The abstract of the dissertation is posted in three languages (Uzbek, Russian, English (resume)) on the website (www.tuit.uz) and on the website of «Ziyonet» Information and educational portal (www.ziyonet.uz.).

Scientific adviser: Nishanov Axram Xasanovich
Doctor of Technical Sciences, Professor

Official opponents: Muxamedieva Dilnoza Tulkunovna
Doctor of Technical Sciences, Professor

Eshankulov Hamza Ilxomovich
PhD in technical sciences, Docent

Leading organization: Karakalpak State University named after Berdakh

The defense will take place «17» March 2023 at 16:00 at the meeting of Scientific council No. DSc.13/30.12.2019.T.07.01 at Tashkent University of Information Technologies (Address: 100084, Tashkent city, Amir Temur street, 108. Ph.: (+99871) 238-64-70, e-mail: iktuit@tuit.uz).

The dissertation can be reviewed at the Information Resourse Centre of Tashkent University of Information Technologies (is registered under No. 270). (Address: 100084, Tashkent city, Amir Temur street, 108. Ph.: (+99871) 238-64-70).

Abstract of dissertation sent out on «01» March 2023 y.
(Dispatching protocol No. 7 on «10» February 2023 y.).



M.M.Musayev

Chairman of the scientific council awarding scientific degrees, Doctor of Technical Sciences, Professor

N.O.Rakhimov

Scientific secretary of scientific council awarding scientific degrees, Doctor of Technical Sciences, dotsent

H.N. Zaynidinov

Chairman of the academic seminar under the scientific council
Awarding scientific degrees, Doctor of Technical Sciences, Professor

INTRODUCTION (abstract of PhD dissertation)

The aim of the research work is to develop algorithms and software for automatic formalization and classification of text appeals through the portal of appeals in the natural uzbek language in the e-government environment.

The object of research work is the requirements for working with appeals in e-government, the portal of appeals, as well as the database of received appeals in the natural uzbek language.

The scientific novelty of the research work is consist on follows:

designed and created a linguistic database that automatically processes texts in natural Uzbek and integrates them into a single system;

developed an algorithm for dividing into extended layers for classifying text appeals of citizens in the natural uzbek language, criteria-based classification selections and ranking functions;

a method of formalization of texts and an algorithm for the combined search of the base of formal appeals have been developed;

the structure of software-functional modules for e-government systems was developed based on algorithms for classifying and formalizing citizens' appeals.

Introduction of research results. Based on the scientific results obtained using the developed method and algorithms for automatic formalization and classification by subject areas of text appeals in the natural uzbek language in the e-government environment:

The software developed by the researcher has been applied in the central office at the virtual reception since April 2022, which allowed the automatic distribution of 32% of the almost 450,000 calls received during these 5 months. These results show that with an increase in the types and number of requests, the knowledge base expands and, as a result, the implementation increases efficiency (Reference of the Ministry for the Development of Information Technologies and Communications No. 33-8 / 5583 of August 30, 2022).

The software developed by the researcher can be used in the e-government environment to automatically process applications by relevant professionals to help them make an intelligent decision. As a result, an increase in labor productivity by an average of 20-30% is achieved. The use of the dissertation work software package in the information systems of the electronic government of the Republic of Uzbekistan in solving the issues of automatic formalization of text appeals in the natural Uzbek language and classification by subject areas helps to increase the efficiency of work (Act of the center for scientific, technical and marketing research "UNICON.UZ" No. 5-3 /1465 dated August 22, 2022).

The software developed by the researcher can be used in the e-government environment to automatically process applications by relevant professionals to help them make an intelligent decision. As a result, an increase in labor productivity by an average of 20-30% is achieved (ACT of JSC "Uzbekiston pochta" dated August 18, 2022).

In the future, when this software package is introduced into the e-government environment, due to the automatic determination by the software tool of the subject

areas of citizens' electronic appeals, an increase in the labor productivity of moderators up to 20-25% will be achieved, and a decrease in the amount of work up to 10-20% (ACT of the state institution "Center e-government project management" dated August 22, 2022).

Publication of research results. A total of 29 scientific papers were published on the research topic, of which 12 articles were published in scientific publications recommended by the Higher Attestation Commission of the Republic of Uzbekistan, including 2 in foreign and 10 in national journals and 3 certificates for computer software.

The structure and volume of the thesis. The dissertation consists of an introduction, four chapters, a conclusion, a list of references and appendices. The volume of the dissertation is 120 pages.

E'LON QILINGAN ISHLAR RO'YXATI
СПИСОК ОПУБЛИКОВАННЫХ РАБОТ
LIST OF PUBLISHED WORKS

I бўлим (I часть; I part)

1. Саидрасулов Ш.Н. Ўзбекистонда рақамли иктисодиёт ва электрон ҳукуматни ривожланишини рейтингли баҳолаш услубияти таҳлили // Ахборот коммуникациялар: Тармоқлар, Технологиялар, Ечимлар илмий-техник журнал (Infocommunication: networks-technologies-solutions), 4 (60)-сон, 2021 йил, –Б. 49-56 (05.00.00; №2).
2. Саидрасулов Ш.Н. Электрон ҳукумат тизимида жисмоний ва юридик шахсларнинг матли мурожаатларини автоматик қайта ишлаш масаласи // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 4(18)-сон, 2021 йил, –Б. 121-124 (05.00.00; №10).
3. Акбаралиев Б.Б., Рўзиев О.Б., Хўжаев О.К., Саидрасулов Ш.Н., Махмудов А.З., Махмудова Ш.А. Moodle тизими базаси элементлари кесимида DATA Mining усулларида фойдаланган ҳолда ўқувчи фаолиятини баҳолаш // ТАТУ хабарлари журнали, 2-сон, 2015 йил. –Б. 11-16 (05.00.00; №31).
4. Хўжаев О.К., Алламов О.Т., Саидрасулов Ш.Н. Маълумотларни интеллектуал таҳлили дастурий воситаларининг қўлланиш соҳалари // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 1(3)-сон, 2018 йил, –Б. 98-101 (05.00.00; №10).
5. Алламов О.Т., Саидрасулов Ш.Н. Транспорт воситалари ҳаракатини тартибга солишда катта масшабли графга тақсимланган ҳолда параллел ишлов бериш // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 1(3)-сон, 2018 йил, –Б. 7-10 (05.00.00; №10).
6. Нишанов А.Х., Бабаджанов Э.С., Саидрасулов Ш.Н. «Электрон ҳукумат» ривожланганлигини баҳолаш услубиятларининг таҳлили (2001-2018 йиллар) // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 1(7)-сон, 2019 йил, –Б. 106-113 (05.00.00; №10).
7. Хўжаев О.К., Саидрасулов Ш.Н., Эгамбердиев О.Н. Синфларга ажратиш масаласини ечишда самарали уулни танлаш алгоритми // Ахборот коммуникациялар: Тармоқлар, Технологиялар, Ечимлар илмий-техник журнал, 1 (49)-сон, 2019 йил, –Б. 39-43 (05.00.00; №2).
8. Нишанов А.Х., Бабаджанов Э.С., Саидрасулов Ш.Н., Кенжаев Х.Б. «Электрон ҳукумат» муҳитида самарали электрон иштирокни ошириш омиллари // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 3(9)-сон, 2019 йил, –Б. 97-104 (05.00.00; №10).
9. Алламов О.Т., Саидрасулов Ш.Н., Қутлибоев Т.Қ. Мобиль қурилмалар ёрдамида объектларнинг геоjoyлашув маълумотларини олиш // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 3(9)-сон, 2019 йил, –Б. 71-75 (05.00.00; №10).
10. Нишанов А.Х., Бабаджанов Э.С., Саидрасулов Ш.Н., Кенжаев Х.Б. Электрон мурожаатларни қайта ишлаш ва электрон хизматлар самарадорлигини ошириш масалалари // Муҳаммад ал-Хоразмий авлодлари илмий-амалий ва ахборот-таҳлилий журнали, 1(15)-сон, 2021 йил, –Б. 109-113 (05.00.00; №10).

11. Nishanov A.X., Babadjanov E.S., Saidrasulov Sh.N. Analysis of methodology of rating evaluation of digital economy and e-government development in Uzbekistan // International Journal of Early Childhood Special Education (INT-JECSE) DOI: 10.9756/INT-JECSE/V14I2.230 ISSN:1308-5581 Vol 14, Issue 02, 2022, pp. 2447-2452, (№1, Web of Science).

12. Nishanov A.X., Babadjanov E.S., Saidrasulov Sh.N. Logical-algorithmic processing of text information in the electronic government system // Journal of Contemporary Issues in Business and Government Vol. 28, No. 03, 2022, pp. 256-260, (№1, Web of Science).

II бўлим (II часть; II part)

13. Nishanov A.X., Babadjanov E.S., Saidrasulov Sh.N., Mamasaidov U.E., Toliev X.I. Mathematical Statement of Dynamic Factors Affecting the Development of Electron Government// International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 1, Number 12 (2020), pp. 5240-75246, © International Research Publication House. <http://www.irphouse.com>

14. Nishanov A.X., Akbaraliev B.B., Saidrasulov Sh.N., Achilov U.S.H., Rayimov T.G., Sobirov B.T. Criteria Study in Solving Data Science Classification Problems // Turkish Journal of Computer and Mathematics Education, Vol.12 No.9 (2021), 1461-1467

15. Комилов М.М., Саидрасулов Ш.Н. Информационный подход для поддержки системы принятия решений // «Информационно-коммуникационные технологии в высшем образовании» Республиканская научно-методическая конференция. Ташкент, 2010. –Б. 96-98.

16. Хўжаев О.К., Саидрасулов Ш.Н. ХОР масаласини нейрон тўрлари ёрдамида МАТЛАБ мухитида ечиш // «Ахборот технологиялари ва телекоммуникация муаммолари» Республика илмий-техник анжумани. Ташкент 21-22 апрель 2011 йил. –Б. 220-224.

17. Нишанов А.Х., Хўжаев О.К., Саидрасулов Ш.Н. Хопфилд нейрон тўри асосида тимсолларни таниб олиш масаласи дастурий таъминоти // «Ахборот технологиялари ва телекоммуникация муаммолари» Республика илмий-техник анжумани. Ташкент, 21-22 апрель, 2011 йил. –Б. 224-227.

18. Нишанов А.Х., Хўжаев О.К., Саидрасулов Ш.Н. Объектга йўналтирилган дастурлаш тилларини ўргатишдаги муаммолар // «Алоқа ва ахборотлаштириш соҳаси учун кадрлар тайёрлаш сифатини ошириш муаммолари» Тошкент ахборот технологиялари университети ва филиаллари профессор ўқитувчиларининг илмий услубий конференцияси. Тошкент, 5-6 январь, 2012 й. –Б. 96-99.

19. Нишанов А.Х., Хўжаев О.К., Саидрасулов Ш.Н. Таълимда эркин ва очик кодли дастурий таъминотлар // «Ахборот технологиялари ва телекоммуникация муаммолари» Ёш олимлар, тадқиқотчилар, магистрант ва талабаларнинг Республика илмий-техник конференцияси. Тошкент, 15-16 март, 2012 й. –Б. 121-124.

20. Нишанов А.Х., Алламов О.Т., Саидрасулов Ш.Н. Коммуникация тармоқларида маршрутизация алгоритмларини қўллаш // «Иновацион ғоялар, технологиялар ва лойиҳаларни ишлаб чиқаришга тадбик этиш муаммолари»

Республика илмий-техник конференцияси илмий ишлар тўплами. Тошкент, 11-12 май, 2012 й. –Б. 257-259.

21. Саидрасулов Ш.Н., Мардиев У.Р. «Операцион тизимлар» фанидан тажриба ишлари доирасидаги топшириқларни бажаришда «Виртуал машина» дастуридан фойдаланиш // «Ахборот ва телекоммуникация технологиялари муаммолари» Республика илмий-техник конференциясининг маърузалар тўплами. Тошкент, 12-13-март, 2015 й. –Б. 31-33.

22. Саидрасулов Ш.Н., Имомов Ш.Ш. Параллеллаштириш жараёнида OPENMP технологиясининг хусусиятлари // «Радиотехника, телекоммуникация ва ахборот технологиялари: муаммолари ва келажак ривож» **Халқаро илмий-техник конференция** мақолалар тўплами. Тошкент, 21-22-май, 2015 й. –Б. 111-113 б.

23. Холикназаров Р.Х., Саидрасулов Ш.Н. Давлат бошқаруви ва тадбиркорликда ахборот-коммуникация технологияларининг ўрни // «Радиотехника, телекоммуникация ва ахборот технологиялари: муаммолари ва келажак ривож» **Халқаро илмий-техник конференция** мақолалар тўплами. Тошкент, 21-22-май, 2015 й. –Б. 575-578.

24. Саидрасулов Ш.Н., Рўзиев О.Б. Анализ мобильных приложений по учёту успеваемости и посещаемости студентов // «Ахборот ва телекоммуникация технологиялари муаммолари» Республика илмий-техник конференциясининг маърузалар тўплами, Тошкент, 10-11 март, 2016 й. –Б. 492-495.

25. Саидрасулов Ш.Н., Рўзиев О.Б., Хамдамбеков У.Б. Эффективность использования системы онлайн обучения детей с ограничениями возможностями // «Современные образовательные технологии в мировом учебно-воспитательном пространстве» VI **Международной научно-практической конференции**. г. Новосибирск, 20-30-июня, 2016 г. –Б. 40-43.

26. Рўзиев О.Б., Хўжаев О.К., Алламов О.Т., Саидрасулов Ш.Н. Программа для решения задачи маршрутизации в динамическом многопараметрном графе // O'zbekiston Respublikasi (с) AVTOR.UZ Elektron depozitariyasi. Intellektual mulk ob'yekti saqlashga kiritilganligi to'g'risida guvohnoma №1680.

27. Нишанов А.Х., Бабаджанов Э.С., Саидрасулов Ш.Н. IGOV 2.0 - Электрон хизматлар самарадорлигини ошириш бўйича маслаҳат-муҳокамамониторинг тизими // O'zbekiston Respublikasi Adliya vazirligi huzuridagi Intellektual mulk agentligi. Ma'lumotlar bazasining rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma №BGU 00486. 05.12.2021.

28. Нишанов А.Х., Бабаджанов Э.С., Саидрасулов Ш.Н. Табиий ўзбек тилдаги мурожаат матнларини формаллаштириш орқали предмет соҳасини аниқлаш дастурий таъминоти // O'zbekiston Respublikasi Adliya vazirligi huzuridagi Intellektual mulk agentligi. Elektron hisoblash mashinalari uchun yaratilgan dasturning rasmiy ro'yxatdan o'tkazilganligi to'g'risidagi guvohnoma №DGU 18169. 22.08.2022.

Avtoreferat «Muhammad al-Xorazmiy avlodlari» O‘zbekiston ilmiy jurnali tahririyatida tahrirdan o‘tkazildi hamda o‘zbek, rus va ingliz tillaridagi matnlarini mosligi tekshirildi (21.02.2023).