

DIAXRON KORPUSNING BIZNES-JARAYONI VA MODEL**Xusainova Zilola Yuldashevna***Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti.**Filologiya fanlari falsafa doktori (PhD), dotsent v.b.**E-mail: xusainovazilola@navoiy-uni.uz***Yangibayeva Surayyo Gulumboyevna***Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti.**Kompyuter lingvistikasi mutaxassisligi II bosqich magistranti.**E-mail: surayyoyangibayeva4@gmail.com*

Annotatsiya. Ushbu maqolada o'zbek tilida yaratilayotgan diaxron korpusning nazariy-metodik asoslari, biznes-jarayoni, texnik modeli hamda korpusni shakllantirishning bosqichlari 1991–2021-yillarda chop etilgan o'zbek badiiy asarlari misolida yoritiladi. Tadqiqotda korpus qurish jarayoni nafaqat lingvistik bosqichlar majmuasi, balki BPMN (Business Process Model and Notation) standartiga tayangan holda shakllantirilgan biznes-jarayon sifatida talqin etiladi. Tadqiqot natijasida diaxron korpus uchun BPMN asosida ishlab chiqilgan biznes-jarayon modeli taklif etildi va korpusning texnologik arxitekturasini batafsil yoritildi. Olingan natijalar diaxron korpusning ilmiy-tadqiqot ishlari, tarixiy til qatlamlarini o'rganish va kompyuter lingvistikasi uchun amaliy ahamiyatini ko'rsatadi.

Kalit so'zlar: diaxron korpus, biznes-jarayon, normalizatsiya, metama'lumotlar, indekslash, analitik modul, ER-model.

Abstract. This article describes the theoretical and methodological foundations, business process, technical model, and the stages of constructing a diachronic corpus being developed for the Uzbek language, using Uzbek literary works published between 1991 and 2021 as a case study. The study interprets the corpus-building process not only as a set of linguistic procedures but also as a business process designed in accordance with the BPMN (Business Process Model and Notation) standard. As a result of the research, a BPMN-based business-process model for the diachronic corpus was proposed, and the technological architecture of the corpus was elaborated in detail. The findings demonstrate the practical significance of the diachronic corpus for research activities, the study of historical language layers, and computational linguistics.

Keywords: diachronic corpus, business process, normalization, metadata, indexing, analytical module, ER model.

Аннотация. В данной статье описаны теоретико-методологические основы, бизнес-процесс, техническая модель и этапы формирования диахронного корпуса, создаваемого для узбекского языка, на примере художественных произведений узбекской литературы, изданных в 1991–2021 годах. В исследовании процесс построения корпуса трактуется не только как

совокупность лингвистических процедур, но и как бизнес-процесс, разработанный в соответствии со стандартом BPMN (Business Process Model and Notation). В результате исследования была предложена бизнес-модель процесса для диахронного корпуса на основе BPMN, а технологическая архитектура корпуса детально раскрыта. Полученные результаты демонстрируют практическую значимость диахронного корпуса для научно-исследовательской деятельности, изучения исторических слоёв языка и компьютерной лингвистики.

Ключевые слова: диахронный корпус, бизнес-процесс, нормализация, метаданные, индексирование, аналитический модуль, ER-модель.

Kirish

Tilning tarixiy taraqqiyot jarayonida yuz beradigan leksik, grammatik, uslubiy va semantik o'zgarishlarni aniqlash, ularning dinamik rivojlanish tendensiyalarini ilmiy asosda qayd qilish lingvistikaning dolzarb yo'nalishlaridan biri hisoblanadi. So'nggi yillarda korpus lingvistikasining jadal rivojlanishi natijasida til birliklarini nafaqat sinxron, balki diaxron kesimda ham tizimli tadqiq etish imkoniyati kengaydi. Diahron korpuslar ma'lum tarixiy davrlarni aks ettiruvchi, vaqt bo'yicha tartiblangan matnlar majmuasi bo'lib, tilning vaqt o'tishi bilan qanday o'zgarishini aniqlash, nutqiy birliklarning paydo bo'lishi, kengayishi yoki yo'qolish jarayonlarini statistik hamda avtomatik lingvistik modellar yordamida tahlil qilish imkonini beradi.

Diahron korpus mavjud bo'lmagan sharoitda tadqiqotchilar matnlarni qo'lda yig'ish, tasniflash va vaqt bo'yicha tahlil qilishga majbur bo'ladilar. Shuningdek, jarayon katta vaqt va resurs talab qiladi, natijada tadqiqotlarning aniqligi hamda takrorlanish imkoniyati cheklanadi. Mavjud korpuslar lingvistik annotatsiya sifati, morfologik belgilash standartlari, metama'lumotlar tizimi va semantik toifalash kabi jihatlar bo'yicha yagona metodikaga ega emas. Bu esa diaxron tadqiqotlar uchun zarur ilmiy ishonchlilikni pasaytiradi. Shu sababli, diaxron korpus yaratish jarayonining biznes-jarayon modeli va texnologik asoslarini ishlab chiqish, jarayonni avtomatlashtirish va ilmiy standartlashtirish bugungi korpus lingvistikasining dolzarb vazifasidir.

Mazkur tadqiqotning maqsadi – mustaqillik yillari matnlari misolida diaxron korpus yaratishning ilmiy-nazariy asoslarini ishlab chiqish, texnologik biznes jarayonini modellashtirish va korpusning funksional imkoniyatlarini belgilashdan iborat. Tadqiqot doirasida korpusning tuzilishi, metama'lumotlar tizimi, vaqt bo'yicha tabaqalashtirish prinsiplari hamda avtomatlashtirilgan jarayonlar asoslab beriladi.

Diahron korpuslarning yaratilishi nafaqat lingvistik jarayon, balki aniq belgilangan bosqichlardan iborat ***biznes jarayoni*** sifatida ham talqin qilinadi.

Zamonaviy korpus lingvistikasida korpus qurish jarayonlarini standartlashtirish, modellashtirish va optimallashtirish muhim ahamiyat kasb etadi. Shu sababli, diaxron korpusni yaratishda **BPMN (Business Process Model and Notation)** kabi xalqaro standartlardan foydalanish jarayonni izchil, boshqariladigan va takrorlanadigan qilish imkonini beradi. BPMN modeli korpus yaratishning barcha bosqichlarini – matnlarni yig‘ishdan tortib, tozalash, metama’lumotlar bilan boyitish, indekslash, tahlil instrumentlarini qo‘shish va yakuniy foydalanuvchi interfeysini yaratishgacha bo‘lgan jarayonlarni vizual ko‘rinishda aniq aks ettiradi. Biznes-jarayon sifatida qaralganda, korpus qurilishi bir necha asosiy bloklarga bo‘linadi: **kiritish jarayoni (input)** – matnlarni yig‘ish; **qayta ishlash jarayoni (process)** – tozalash, normallashtirish, anotatsiyalash va metama’lumotlarni shakllantirish; **chiqish jarayoni (output)** – tahlil qilish, statistik hisob-kitoblar va foydalanuvchi interfeysi orqali foydalanish. Ushbu bosqichlarning har biri o‘zining alohida vazifalari, qaror nuqtalari va mas’ul rollariga ega bo‘lib, korpus yaratishning umumiy samaradorligiga bevosita ta’sir ko‘rsatadi.

ADABIYOTLAR TAHLILI

Korpus NLP tizimlarining asosidir. Ular AI va mashinali o‘rgatish tizimlarini o‘rgatish uchun ishlatiladi. Ular axborot hayotiy davrlarini modellashtirish va bashorat qilish uchun keng va xilmaxil ma’lumotlar to‘plamini taqdim etadi. Korpus NLP tizimini tabiiy tilni boshqarish va sharhlash uchun tayyorlaydi, bu esa odamlar bilan tabiiy tilda oson muloqot qilish imkonini beradi [Elov B.B., Yuldashev A.U., Yodgorov U.S. 2024]. Ma’lumki diaxron korpuslar bu ma’lum bir tilning yoki lahjaning tarixiy rivojlanishini o‘rganish uchun yaratilgan matnlar to‘plamidir. Ular turli davrlarda chop etilgan matnlarni birlashtirib, tilning leksik, morfologik, sintaktik va semantik o‘zgarishlarini tahlil qilish imkonini beradi [Kilgarriff A. 2005]. Diaxron korpusning biznes modeli esa uning funksional arxitekturasini belgilaydi. Modellashtirish jarayonida korpusning ma’lumotlar bazasi strukturasini, indekslash mexanizmlarini, tahliliy modullarini va foydalanuvchi bilan o‘zaro aloqasini tashkil etuvchi komponentlar aniqlanadi. Bu yondashuv korpusni nafaqat tadqiqot uchun qulay resurs, balki uzoq muddatli rivojlantiriladigan ilmiy platforma sifatida shakllantirish imkonini beradi. BPMN asosidagi yondashuv jarayonning shaffofligi, takomillashtirilishi va kengaytirilishi uchun mustahkam metodologik asos yaratadi. Shu bois, ushbu ilmiy tadqiqotda diaxron korpusning nafaqat lingvistik, balki **jarayon boshqaruvi va texnologik modellashtirish** aspektlari ham yoritiladi.

Diaxron korpuslar tilning va matnlarning tarixiy rivojlanishini tahlil qilishda muhim vosita hisoblanadi. Ushbu ilmiy ish 1991–2021-yillar oralig‘idagi o‘zbek badiiy asarlarini qamrab olgan diaxron korpus yaratishga qaratilgan bo‘lib, uning shakllanish jarayoni nafaqat lingvistik, balki aniq belgilangan **biznes-jarayon**

modeli asosida tashkil etildi. Korpus yaratishning texnologik bosqichlarini biznes nuqtayi nazardan modellashtirish jarayonni izchil, boshqariladigan va takrorlanadigan qiladi. Zamonaviy ilmiy yondashuvlarda bunday jarayonlar ko‘pincha BPMN kabi rasmiy standartlar orqali ifodalanadi [Dumas M., La Rosa M., Mendling J. and Reijers H. 2013]. Tadqiqotda qo‘llangan BPMN modeli korpus yaratishning barcha bosqichlarini – matnlarni yig‘ish, tozalash, metama’lumotlar yaratish, indekslash va foydalanuvchi interfeysi orqali tahlil qilishni – yagona jarayon sifatida tasvirlash imkonini berdi. Bu yondashuv ilmiy jarayonlarning samaradorligini oshiradi va workflow boshqaruviga oid nazariy asoslar bilan hamohangdir [Reijers H.A. 2003]. Shuningdek, korpus yaratilishida lingvistik jarayonlarning metodik muhiti ham inobatga olindi. Matnlarni yig‘ish, tozalash va anotatsiya qilish jarayonlari korpus lingvistikasining umumiy metodologik standartlariga tayanadi [McEnery T. and Hardie A., 2012]. Har bir matn muallif, asar nomi, yozilgan yili, nashr yili va janr kabi atributlar bilan indekslandi, bu esa korpusni tizimli tahlil qilish imkonini berdi.

ASOSIY QISM

Diaxron korpusning biznes-jarayoni va arxitekturas

Diaxron korpusni yaratish jarayoni nafaqat lingvistik jarayonlar ketma-ketligidan, balki murakkab axborot tizimi sifatida qaraladigan **biznes-jarayonlar tizim**idan iborat. Biznes-jarayonlarni modellashtirish (BPM) raqamli tizimlar uchun funksional oqimlarni aniqlash, ularni standartlashtirish va optimallashtirish imkonini beradi. Shu bois korpus arxitekturas BPMN (Business Process Model and Notation) standarti asosida qayta ishlab chiqildi. BPMN modeli korpus yaratish siklini bosqichma-bosqich, formallashtirilgan va takrorlanuvchi biznes-jarayon sifatida tasvirlashga yordam beradi.

Diaxron korpusning biznes-jarayoni dastlab mustaqillik yillariga oid representativ matnlar tanlash hisoblanib, matnlarni tizimli yig‘ishdan boshlanadi. Tanlov jarayonida quyidagi mezonlar asos qilib olindi:

- Davriy oralig‘i. 1991–2021-yillar.
- Janrlar. Ilmiy-ommabop, rasmiy hujjatlar, publitsistika, matbuot, badiiy matnlar.
- Manba barqarorligi. matnlarning elektron shaklda mavjudligi va ularning autentikligi.

Ushbu tadqiqot doirasida 1991–2021 yillar oralig‘ida nashr qilingan o‘zbek badiiy asarlari asosiy manba sifatida tanlandi. Har bir asar turli formatlarda (PDF, TXT, Word) yig‘ildi va muallif, nashr yili, janr bo‘yicha tasniflandi. Matnlarni yig‘ish jarayonida ularning rasmiy nashr nusxalari va ishonchli manbalar tanlanishi korpusning ilmiy ishonchliligini ta‘minlash uchun muhimdir. Yig‘ilgan matnlar keyingi bosqichda tozalash va normalizatsiyadan o‘tkazildi. Bu bosqichda Kril

yozuvi Lotin alifbosiga o'tkazildi, imlo va orfografik xatolar tuzatildi, ortiqcha bo'sh joylar va noto'g'ri belgilardan xalos bo'linadi. Matnlarni tozalash jarayoni korpusning sifatli tahlil qilinishini kafolatlaydi, shuningdek, avtomatlashtirilgan tahlil algoritmlarining samaradorligini oshiradi. Quyida ushbu bosqichlar ketma-ketlikda yoritiladi.

1. Matnlarni yig'ish, raqamlashtirish va manbalarni aniqlash

Til korpusi matnlari Unicode kodlashtirish tizimidagi simvollardan iborat bo'lib, mashina tomonidan o'qilishi kerak. Korpus matnlari hozirda foydalanilayotgan, "tozalangan" va tabiiy sharoitlarda to'plangan tabiiy tilning elementlari (leksik birliklari)dan iborat bo'lishi kerak [Elov B.B., Yuldashev A.U., Yodgorov U.S. 2024]. Diaxron korpus yaratish jarayonining ham boshlang'ich bosqichi – **manbalarni aniqlash** va **matnlarni to'plash** jarayonlaridir. Ushbu tadqiqot doirasida 1991–2021-yillar oralig'idagi o'zbek badiiy adabiyotiga mansub 120 ga yaqin asar tanlandi. Manbalarni tanlashda quyidagi mezonlar asos qilib olindi:

- adabiy asarning nashr qilingan yili;
- tilning davriy o'zgarishini aks ettirishi;
- mualliflar ko'lami keng bo'lishi;
- matnlarning elektron ko'rinishda mavjud bo'lishi.

Bu bosqich BPMN notatsiyasida **“Collect Textual Resources”** deb nomlanadigan asosiy faoliyat turi sifatida belgilanadi. Matnlar PDF, DOCX va TXT formatlarida yig'ildi. Matnlar turli formatlarda bo'lgani sababli ularni raqamli standartga keltirish uchun quyidagi amallar bajarildi:

- PDF va skanerlangan hujjatlardan OCR (Optical Character Recognition) orqali matnlarni ajratib olish.
- To'g'ri qilingan belgilarni aniqlash va avtomatik ortografik normalizatsiya.
- Dublikatsiyalarni, reklama matnlarini, ortiqcha belgilarni tozalash.
- Matn bloklari yaxlitligini tiklash (sarlavha, paragraf, izohlar).
- Natijada korpus uchun yagona toza va standartlashtirilgan matnlar bazasi yaratildi.

2. Matnlarni tozalash va normalizatsiya qilish

Matnlar turli formatlarda kelishi ularni yagona standartga moslashtirish ehtiyojini tug'diradi. Shu bois korpus yaratish jarayonining navbatdagi bosqichi – **tozalash** va **normalizatsiya** jarayonlaridir.

Ushbu bosqich quyidagi jarayonlarni o'z ichiga oladi:

1. **Kril** → **Lotin transliteratsiyasi**. Mustaqillik yillarida turli yozuvlarda chop etilgan asarlarda transliteratsiya muammosi mavjud bo'lgani sababli

avtomatlashtirilgan transliteratsiya moduli qoʻllandi. Transliterator imloviy meʼyorlarni inobatga olgan holda yaratilgan.

2. **Orfografik xatolarni avtomatik tuzatish.** Matnlarda uchraydigan eng koʻp uchraydigan xatoliklar (qoʻsh harflar, oʻrin almashish, ortiqcha belgilalar) algoritmik tarzda aniqlanib, korpusda yagona shaklga keltirildi.

3. **Belgilarni standartlashtirish.** Qoʻshtirnoqlar, tire, nuqta, vergulning turlicha kod variantlari bir xil koʻrinishga oʻtkazildi.

3. Metamaʼlumotlarni yaratish va atributlash

Har bir matnni tizimli tahlil qilish uchun metamaʼlumotlar yaratish muhim bosqich hisoblanadi. Tadqiqot doirasida har bir asarga quyidagi atributlar kiritildi: muallif, sarlavha, nashr yili, janr, til va matn uzunligi. Ushbu maʼlumotlar standart formatda bazaga kiritildi, bu esa keyinchalik korpusni indekslash va avtomatik soʻrovlar orqali tez tahlil qilish imkonini beradi. Metamaʼlumotlar yaratish jarayoni diaxron korpusning ilmiy qadriyatini oshiradi, chunki ular orqali har bir matn konteksti va xususiyatlari aniq belgilanadi, shuningdek, turli davrlar va adiblar boʻyicha solishtirish tahlillarini oʻtkazish imkonini beradi [Biber D. 1993].

Mustaqillik davri nashrlariga asoslangan diaxron korpusning muhim tarkibiy qismi – **metamaʼlumotlar (metadata)** boʻlib, ular matnlarning kontekstual maʼnosini aniqlashga yordam beradi va ilmiy tahlilning qamrovini kengaytiradi. Ushbu bosqich quyidagi atributlarni aniqlashni oʻz ichiga oladi:

- muallif (ismi, familiyasi);
- asar nomi;
- yozilgan yili
- nashr qilingan yili;
- janr;
- matnning umumiy hajmi (soʻzlar soni, gaplar soni).

4. Korpusni indekslash va strukturasini shakllantirish

Korpus strukturasini yaratish bosqichi matnlarni indekslash va lingvistik birliklarni belgilashni oʻz ichiga oladi. Har bir matn gaplar boʻyicha alohida satrlarga ajratildi, soʻngra n-gramlar, morfologik va sintaktik birliklar aniqlanib, bazaga kiritildi. Bu jarayon orqali korpus ichidagi lingvistik birliklar tizimli tarzda saqlanadi va keyinchalik avtomatlashtirilgan tahlil, qidiruv va vizualizatsiya imkoniyatlari yaratadi [Gries S. 2009]. Shuningdek, korpus strukturasini shakllantirish jarayoni diaxron tahlil uchun asosiy platforma sifatida xizmat qiladi, chunki u tilning vaqt oʻtishi bilan oʻzgarishini kuzatish, leksik va sintaktik tendensiyalarni aniqlash imkonini beradi.

Korpusning tuzilishi uning ilmiy qiymatini aniqlovchi asosiy koʻrsatkichlardan biridir. Ushbu bosqichda korpus quyidagi lingvistik birliklarga ajratildi:

4.1. Gaplar bo'yicha ajratish (Sentence Segmentation)

Har bir asar maxsus algoritmlar yordamida gaplarga ajratildi. O'zbek tilida tinish belgilari ba'zan noaniq vaziyatlarda kelishi (masalan, qisqartmalar, dialoglar, ohang belgilarida) sababli qo'shimcha qoidalar ishlab chiqildi.

4.2. N-gramlar generatsiyasi

Korpusning diaxron tahlilida n-gramlar (unigram, bigram, trigram) muhim statistik birlik hisoblanadi, chunki ular tilning leksik dinamikasini ko'rsatadi. Ushbu bosqich "*N-gram Extraction*" deb nomlangan ma'lumotlarni qayta ishlash bosqichi sifatida modellashtirildi.

Lingvistik belgilash

Zarur joylarda qo'shimcha grammatik yoki leksik teglash amalga oshirildi. Bu bosqich keyinchalik sintaktik tahlil va semantik jarayonlar uchun mustahkam asos yaratadi.

5. Diaxron qatlamlarga ajratish

Matnlar quyidagi davriylik asosida davrlashtirildi:

1991–2000-yillar;

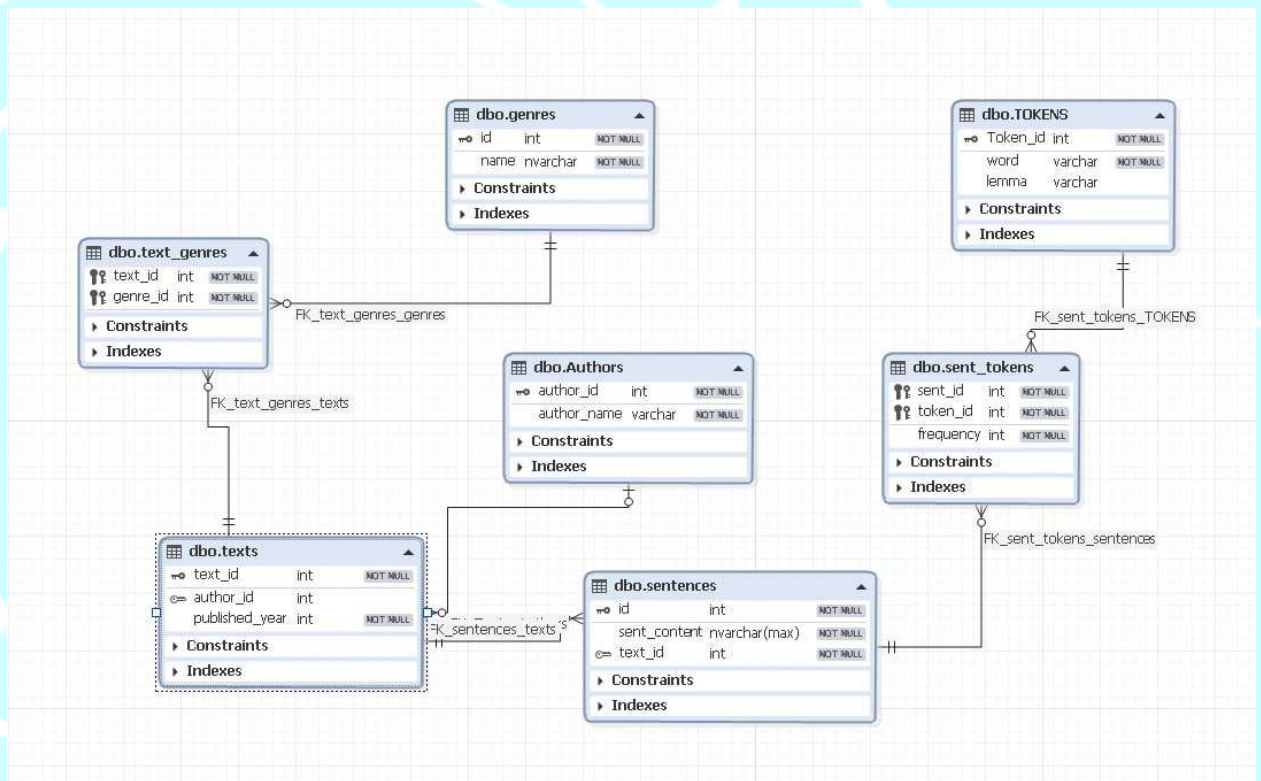
2001–2010-yillar;

2011–2020-yillar.

Bu ajratish tilshunoslik tadqiqotlarida vaqt bo'yicha leksik va grammatik o'zgarishlarni solishtirish imkonini beradi.

6. Ma'lumotlar bazasini shakllantirish

Ushbu jarayonlarning ma'lumotlar bazasi darajasidagi amalga oshirilishi diaxron korpusning ER-modeli orqali ifodalanadi. ER-model korpusning asosiy tarkibiy qismlari – matnlar, metama'lumotlar, gaplar, n-gramlar va statistik birliklarning bir-biri bilan bog'lanish mexanizmini aks ettiradi. Ushbu model korpus ichida ma'lumotlarning saqlanishi, izlash, filtrlash va tahlil qilish jarayonlarining optimallashtirilgan arxitekturasini ta'minlaydi. Quyida diaxron korpus uchun ishlab chiqilgan ma'lumotlar bazasi ER-modeli keltiriladi (1-rasm) [Xusainova Z., Yangibayeva S. 2025].



1-rasm. Diaxron korpusning ma'lumotlar bazasi ER-modeli

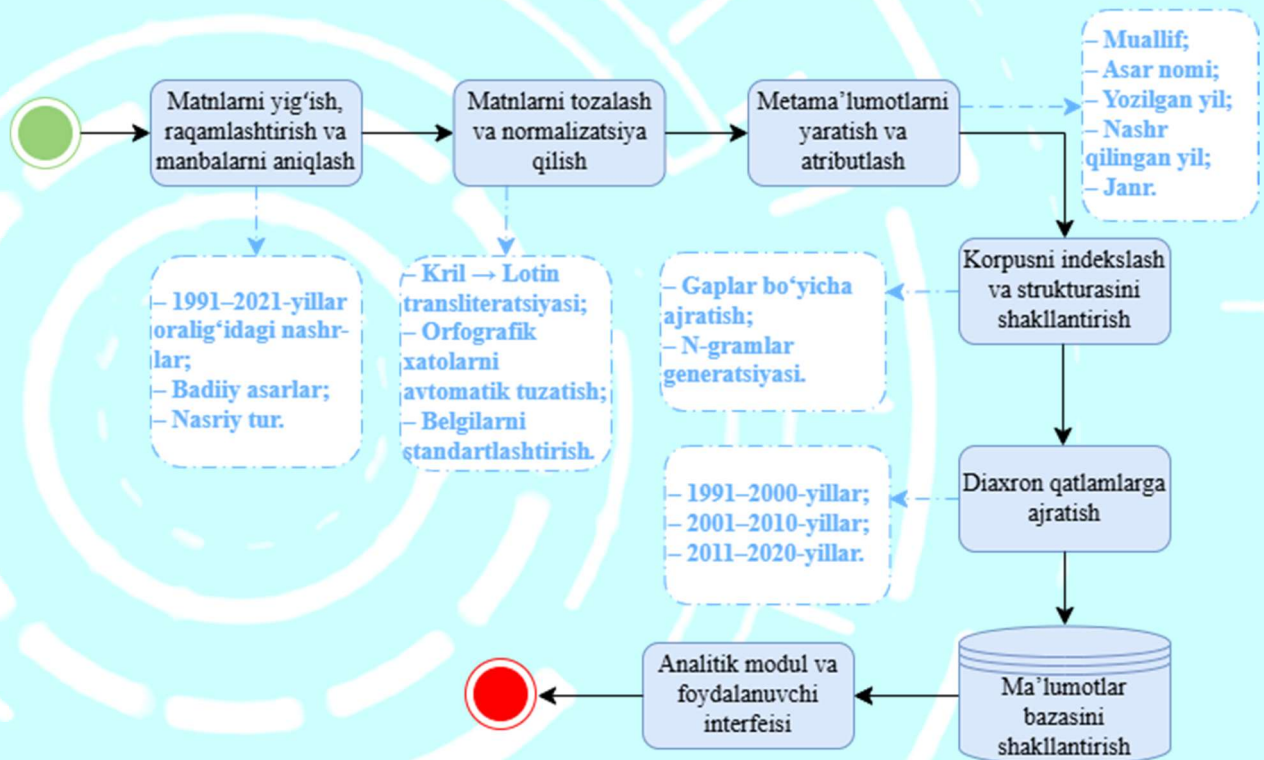
Mazkur strukturaviy model korpusning keyingi bosqichlari – analitik modulni shakllantirish, foydalanuvchi so'rovlarini qayta ishlash va statistik tahlillarni bajarish jarayonlari uchun asos vazifasini bajaradi.

7. Analitik modul va foydalanuvchi interfeisi

Korpusdan foydalanishning keyingi bosqichi tahlil va foydalanuvchi interfeisi bilan bog'liq. Analitik modul orqali matnlar statistik tahlildan o'tkazildi, eng ko'p ishlatilgan so'zlar, iboralar va n-gramlar aniqlab olindi. Natijalar grafik va jadval ko'rinishida vizualizatsiya qilindi, bu esa korpusni ilmiy maqolalarda va pedagogik maqsadlarda qo'llash imkoniyatini oshiradi. Foydalanuvchi interfeysi orqali izlash va filtrlar yordamida matnlar bo'yicha tezkor so'rovlar o'tkazish mumkin, shuningdek, turli davrlar va adiblar bo'yicha solishtiruvlar amalga oshiriladi. Shu tarzda, diaxron korpus ilmiy ishlar uchun qulay va samarali tahlil vositasi sifatida xizmat qiladi.

Diachron korpusning ilmiy qiymati uni qanday tahlil qilish imkoniyatlari bilan belgilanadi. Shu sababli korpus uchun **analitik modul** ishlab chiqildi. Modul quyidagi imkoniyatlarni beradi: matn va metama'lumotlar bo'yicha kengaytirilgan qidiruv; yillar bo'yicha til birliklari dinamikasini tahlil qilish; so'z chastotasi, n-gramlar, gap uzunligi va boshqa statistik ko'rsatkichlarni hisoblash; grafik vizualizatsiyalarni yaratish; filtrlar asosida matnlarni solishtirish.

Yuqoridagi bosqichlarga asoslangan holda mustaqillik davri nashrlariga asoslangan diachron korpusning quyidagicha biznes-jarayoni shakllantirildi (2-rasm):



1-rasm. Diaxron korpusning biznes-jarayoni sxemasi

XULOSA

Ushbu maqolada Mustaqillik yillariga oid o'zbek matnlari asosida diaxron korpusni yaratish jarayoni, uning biznes-jarayoni modeli va texnologik asoslari tahlil qilindi. BPMN notatsiyasiga tayangan holda korpus yaratishning to'liq avtomatlashtirilgan jarayoni ishlab chiqildi, bu esa korpus lingvistikasida zamonaviy texnologik yondashuvni tatbiq etish imkonini berdi. Korpusni shakllantirishning barcha bosqichlari – matnlarni yig'ish, raqamlashtirish, tozalash, normalizatsiya qilish, metama'lumotlar bilan boyitish va ma'lumotlar bazasiga joylashtirish yaxlit tizim sifatida modellashtirildi.

Mazkur jarayonda korpusning sifatini, uzluksiz yangilanishini va foydalanish qulayligini ta'minlash uchun biznes-jarayonlarning strukturasi va ularning integratsiyasi muhim ekani aniqlanadi. Korpus biznes-jarayoni samarali tashkil etilganida, ilmiy tadqiqotlar uchun ishonchli ma'lumotlar bazasi yaratiladi, bu esa lingvistik tahlil, tilshunoslik va kompyuter lingvistikasi sohalarida yangi imkoniyatlar ochadi. Tadqiqotda ishlab chiqilgan algoritmlar, texnologik arxitektura va ontologik tizimlarning jarayon bilan uyg'unligi korpusning funksionalligi va avtomatlashtirilgan ishlashiga katta hissa qo'shadi.

Shuningdek, ushbu tadqiqot diaxron korpusni yanada kengaytirish, yangi til va davrlarni qamrab olish, shuningdek, avtomatlashtirilgan jarayonlarni rivojlantirish imkoniyatlarini ko'rsatadi. Kelajakda diaxron korpusni boshqa lingvistik va ilmiy sohalarida qo'llash, uning interaktiv va foydalanuvchi uchun qulay platforma sifatida rivojlanishini ta'minlash mumkin. Natijada, diaxron

korpusning biznes-jarayoni – bu nafaqat texnik va lingvistik masala, balki ilmiy, texnologik va boshqaruv nuqtai nazaridan muhim strategik tizim sifatida namoyon bo‘ladi. U ilmiy tadqiqotlarni samarali qo‘llab-quvvatlash, ma’lumotlarni tizimli ravishda boshqarish va kelajakdagi lingvistik modellarni yaratish uchun asos bo‘lib xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR RO‘YXATI

1. Elov B.B., Yuldashev A.U., Yodgorov U.S. *Til korpusi turlari va umumiy xususiyatlari*. Xorazm ma’mun akademiyasi axborotnomasi. 2024. 118 –124.
2. Kilgarriff A. *Language is never, ever, ever random*. Corpus Linguistics 2005. 263-275.
3. Dumas M., La Rosa M., Mendling J. and Reijers H. *Fundamentals of Business Process Management*. Springer. 2013. 45-48.
4. Reijers H.A. *Workflow Management: Modeling Concepts, Architecture and Implementation*. MIT Press. 2003.
5. McEnery T. and Hardie A. *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press. 2012. 193-194.
6. Elov B.B., Yuldashev A.U., Yodgorov U.S. *Lingvistik korpuslar tipologiyasi va ularga qo‘yilgan talablar*. Xorazm ma’mun akademiyasi axborotnomasi. 2024. 128–129.
7. Biber D. *Representativeness in Corpus Design*. Literary and Linguistic Computing, 1993. 8(4), 243–257.
8. Gries S. *Quantitative Corpus Linguistics with R*. Routledge. 2009.
9. Xusainova Z., Yangibayeva S. *Diaxron korpusning ER-modelini yaratish*. O‘zbekiston: til va madaniyat. Kompyuter lingvistikasi. 2025.