O‘ZBEK TILI KORPUSINI YARATISH VA UNING ISHYURITISHDAGI AHAMIYATI

Nurmatjonova Odinaxon Bexzodbek qizi
Andijon davlat chet tillari instituti, Ingliz tili va adabiyoti fakulteti
Xorijiy til va adabiyoti yo‘nalishi, 1-kurs talabasi
https://doi.org/10.5281/zenodo.24020402

Annotatsiya. Ushbu maqolada o‘zbek tili korpusini yaratishning nazariy va amaliy asoslari,uning zamonaviy tilshunoslikdagi o‘rni hamda davlat va tashkilotlarda ish yuritish jarayonidagi ahamiyati yoritiladi. O‘zbek tilining milliy korpusini shakllantirish, elektron matnlarni yig‘ish, saralash, lingvistik jihatdan belgilash va yagona ma’lumotlar bazasiga joylashtirish masalalari tahlil qilinadi. Shuningdek, korpus texnologiyalaridan rasmiy hujjatlar tayyorlash, imlo va uslubiy xatolarni aniqlash, terminlarni bir xilda qo‘llash, elektron hujjat almashinuvi va tarjima jarayonlarida foydalanish imkoniyatlari ko‘rib chiqiladi. Tadqiqot davomida o‘zbek tilshunosligida yaratilgan ilmiy manbalar va korpus lingvistikasi bo‘yicha mavjud ishlanmalar o‘rganildi. Natijada o‘zbek tili korpusining nafaqat ilmiy tadqiqotlar, balki davlat boshqaruvi, ta’lim, hujjatchilik va raqamli texnologiyalar rivojida ham muhim vosita ekanligi asoslab berildi.

Kalit so‘zlar: o‘zbek tili, milliy korpus, korpus lingvistikasi, kompyuter lingvistikasi, elektron matn, ish yuritish, rasmiy hujjatlar, lingvistik annotatsiya, raqamli texnologiyalar, avtomatik tahlil.

Abstract. This article examines the theoretical and practical foundations of creating an Uzbek language corpus, its role in modern linguistics, and its importance in administrative documentation. The study analyzes the collection, processing, annotation, and systematization of electronic texts. Particular attention is paid to the use of language corpora in preparing official documents, checking spelling, standardizing terminology, and developing electronic document management.

Keywords: Uzbek language, national corpus, corpus linguistics, computational linguistics, office management, official documents, electronic texts, linguistic annotation.

KIRISH

Bugungi globallashuv va axborot texnologiyalari jadal rivojlanayotgan davrda tilning imkoniyatlarini zamonaviy texnologiyalar asosida o‘rganish va rivojlantirish muhim vazifalardan biriga aylanmoqda. Har bir millatning tili uning tarixi, madaniyati, ma’naviyati va milliy o‘zligini ifodalovchi asosiy vositadir. Shu sababli ona tilini asrash, uning boy imkoniyatlarini ilmiy jihatdan o‘rganish va raqamli muhitda keng qo‘llash dolzarb masala hisoblanadi.

O‘zbek tili ham uzoq tarixiy taraqqiyot davomida boy lug‘at tarkibi, grammatik qurilishi, uslubiy imkoniyatlari va ifoda vositalariga ega bo‘lgan tillardan biridir. Bugungi kunda o‘zbek tilidan ta’lim, fan, davlat boshqaruvi, ommaviy axborot vositalari va turli tashkilotlar faoliyatida keng foydalanilmoqda. Biroq tilning raqamli muhitdagi imkoniyatlarini yanada rivojlantirish uchun uning elektron ma’lumotlar bazasini shakllantirish zarur.

Ana shunday ehtiyoj natijasida korpus lingvistikasi yo‘nalishi shakllandi. Korpus lingvistikasi til birliklarini katta hajmdagi elektron matnlar asosida o‘rganadigan, ularning qo‘llanish chastotasi, ma’nosi, grammatik xususiyatlari va o‘zaro munosabatlarini tahlil qiladigan zamonaviy tilshunoslik yo‘nalishidir.

Til korpusi deganda muayyan tilga tegishli bo‘lgan, elektron shaklda saqlanadigan,

tartibga solingan va lingvistik ma’lumotlar bilan boyitilgan matnlar majmuasi tushuniladi. Oddiy elektron matnlar to‘plamidan farqli ravishda korpus maxsus qidiruv, tahlil va saralash imkoniyatlariga ega bo‘ladi.