

SKETCH ENGINE PLATFORMASIDA KORPUSLAR YARATISHNING AMALIY TADQIQI

Yusupova Ra'no Davronovna

Buxoro Davlat Universiteti, xorijiy tillar fakulteti,
ingliz adabiyotshunosligi yo'nalish talabasi.

Annotatsiya:

Maqolada Sketch Engine platformasida korpuslar yaratishning amaliy tadqiqi haqida va maqsadi, vazifalarini kengroq qilib berilgan.

Kalit so'zlar: Tilshunoslik, korpus, elektron, matnlar, lingvistik, ob'ekt, reprezentativli, tarixiy, zamonaviy.

Аннотация: В статье описано практическое исследование создания кейсов на платформе Sketch Engine, а также его цели и задачи.

Ключевые слова: Лингвистика, корпус, электронные, тексты, лингвистический, предметный, репрезентативный, исторический, современный.

Abstract: The article describes the practical research of creating cases on the Sketch Engine platform, as well as its goals and tasks.

Key words: Linguistics, corpus, electronic, texts, linguistic, object, representative, historical, contemporary.

Tilshunoslikda korpus (ko'plikda corpora) yoki matn korpusi katta va tizimlangan matnlar to'plamidan (hozirgi kunda, odatda, elektron saqlanadi va qayta ishlanadi) iborat til manbayidir. Korpus tilshunosligida ular muayyan til doirasidas tahlillarni amalga oshirish va gipotezani tekshirish, tildagi hodisalarni kuzatish yoki nazariy lingvistik qoidalarni tekshirish uchun foydalaniladi.

Tilshunoslikning muhim vazifalaridan biri lingvistik tadqiqotlar uchun faktik materiallar manbalarini elektron shaklda to'plash va saqlashdir. Ayni paytda ushbu

muammoni hal qilish uchun qulay tarzda saqlanadigan turli xil funksiyalarga ega maxsus texnologiyalar nafaqat katta hajmdagi matnlarni elektron shaklda saqlashga imkon beradi, balki ularni qidirish, qayta ishlash imkonini ham yaratadi.

Elektron shaklda yoki korpusda matnlar to'plamini yaratish vazifasi zamonaviy tilshunoslik uchun shunchalik muhimki, elektron matnlar amaliy tilshunoslikning maxsus bo'limi - korpus lingvistikasining tadqiqot ob'ektiga aylanadi.

Korpuslarni qulay yaratib olish, undan foydalanish uchun korpus menejeri tizimlari yaratilgan. Bunda biz tayyor korpus menejerlarida korpus yarata olamiz. Bunda bizga faqat korpus uchun kerak bo'ladigan matnlar to'plami kerak bo'ladi. Bundan tashqari korpus menejerida tayyor korpuslardan foydalanish ham mumkin. Biz bugun mavjud korpusdan foydalanish qismini ko'rib chiqamiz. Korpus menejerining qulayliklaridan tezda bizga natija taqdim qiladi, korpus menejeri dizayni ham yaxshi ishlab chiqilgan.

Korpuslarni yaratishning maqsadga muvofiqligi va ulardan foydalanish omillari quyidagilar bilan belgilanadi:

- 1) korpus hajmining yetarlicha ulkanligi (ayniqsa, reprezentativli) va balanslashtirilganligi ma'lumotlar tipikligini kafolatlaydi va til hodisalarining butun spektrini aks ettirishning to'liqligini ta'minlaydi;
- 2) har xil turdagi ma'lumotlar korpusda tabiiy holda (kontekstli shaklda) joylashgan bo'lib, bu ularni har tomonlama va obyektiv o'rganish imkoniyatini yaratadi;
- 3) bir marta yaratilgan va tayyorlangan matnlar majmuasi (ma'lumotlar to'plami) turli tadqiqotchilar tomonidan turli maqsadlarda qayta-qayta foydalanilishi mumkin;
- 4) korpusda har xil uslubdagi materiallar bo'lgani bois nutqiy uslubga xos tadqiqotlarni pragmatik jihatdan o'rganish imkoni yuqori bo'ladi;
- 5) muayyan ijodkorga tegishli materiallarning tayyor elektron to'plami sifatida korpusdan ijodkorning lingvopoetikasi, yozish uslubi, so'z qo'llash mahorati, leksik birliklarini qo'llash statistikasini aniqlash mumkin bo'ladi;
- 6) muayyan leksik birlikning uzual (kontekstdan tashqari) va okkozional (kontekstdagi) ma'nolarini aniqlash imkoni oshadi.

Korpus leksik ma'lumotlar massivi sifatida muayyan davrda tilning leksik tarkibi o'zgarishi jarayonlari dinamikasini o'rganishga, turli janrlarda va turli mualliflarda leksik hamda grammatik xususiyatlarning qo'llanilishini tahlil qilishga imkon beradi. Korpuslar, shuningdek, turli xil tarixiy va zamonaviy lug'atlarni



tayyorlashda ko'p qirrali leksikografik ishlarning manbai va vositasi bo'lib xizmat qiladi.

Shuningdek, ushbu korpuslardan ta'lim jarayonida tilni o'rgatish uchun grammatikani qurish va takomillashtirishda foydalanish mumkin. Korpus matnlarni statistik tahlil qilish va lingvistik gipotezalarni tekshirish, qoidalarni asoslash yoki yuzaga kelish holatlarini tekshirish uchun foydalaniladigan tizim hisoblanishi bilan ahamiyatli. Korpus lingvistikasining obyekti tabiiy tildagi matnlar hisoblanadi.

Ammo KorpLingning ikki tomonlama tabiati (matn korpuslarini yaratish va ulardan foydalanish) obyektning ham ikki tomonlama xarakterini belgilaydi. Bir tomondan KorpLing va boshqa lingvistik fanlar uchun boshlang'ich nutq materialidir; ikkinchi tomondan, KorpLing faoliyatining natijasi.

Korpus lingvistikasining predmeti esa til korpuslari sanaladi. Aytish mumkinki, KorpLing o'z predmeti sifatida keng doiradagi foydalanuvchilarning manfaati uchun lingvistik tadqiqotlarga mo'ljallangan til ma'lumotlarining yirik majmualarini yaratish va ulardan foydalanishning nazariy asoslari hamda amaliy mexanizmlariga ega.

Korpus lingvistikasining asosiy maqsadi til tizimining lingvistik tavsifi sanaladi, bunda muloqotni o'rganish jarayonida olingan nutqiy materiallar matnlari korpusida aks etishi e'tiborga olinadi.

Korpus lingvistikasining vazifalari quyidagilar hisoblanadi:

- 1) ushbu sohaning nazariy asoslarini ishlab chiqish;
- 2) har xil turdagi korpuslarni yaratish va ulardan foydalanish tajribasini tahlil qilish;
- 3) korpusga qo'yiladigan umumiy talablarni shakllantirish;
- 4) turli xil tadqiqot va o'quv vazifalari uchun korpuslarni yaratish;
- 5) tilshunoslikning turli sohalarida matn korpuslarini qo'llashning samarali usullarini shakllantirish.

Vazifalarning bu tarzda shakllantirilishi bilan matnlar korpusining o'zi obyekt vazifasini bajaradi, ammo boshqa tomondan, u korpus tilshunosligining mahsuli va tilni o'rganish sohasidir.

Foydalanilgan adabiyotlar.

1. Баранов А.Н. Введение в прикладную лингвистику, Лидия Васильевна; государственный лингвистический университет. - Москва, 2004.

3. Л.Ю. Щипицина информационные технологии в лингвистике С.

4. Гатауллин Р.Р., Гильмуллин Р.А. Контекстные правила для разрешения морфологической многозначности в корпусе татарского языка / Р.Р. Гатауллин, Р.А. Гильмуллин // Открытые семантические технологии проектирования интеллектуальных систем OSTIS-2016 Open Semantic Technologies for Intelligent Systems: Материалы VI международной научно-технической конференции (Минск, 18-20 февраля 2016 года). - Минск: БГУИР, 2016. - С.

5. Невзорова О.А., Мухамедшин Д.Р., Билалов Р.Р. Корпус-менеджер для тюркских языков: основная функциональность // Труды международной конференции «Корпусная лингвистика - 2015». - СПб.: С.-Петербургский гос. Университет, филологический факультет, 2015. - С.

6. <http://thttps://www.sketchengine.eu/>

