

TERMINOLOGIK MA'LUMOT BAZALARINI YARATISHNING ME'YORLARI

Zaripboyeva Nafosat Maratovna

Alisher Navoiy nomidagi Toshkent
davlat o'zbek tili va adabiyoti universiteti
“Tarjima nazariyasi va amaliyoti fakulteti” talabasi
e-mail: nafosat2502@mail.ru

Annotatsiya: maqolada o'zbek tili ta'limiy korpusi uchun terminlar korpusi yaratish jarayoni muhokama etilgan. Terminlar bazasini shakllantirishning lingvistik asoslari tadqiq qilinib, terminlar bazasini tuzish tamoyillari ishlab chiqilgan hamda morfologik teg xususiyatlari tavsiflangan.

Til – o'zgaruvchan hodisa. Tildagi o'zgarishlar dastavval leksikologiya va terminologiyada o'z ifodasini topadi. Ilmiy, texnik, iqtisodiy va boshqa terminlar hamda terminologik ma'lumotlarning katta hajmini tartibga solish va qayta ishlash murakkab, uzluksiz jarayon bo'lib, bu jarayonni qo'lda amalga oshirish ancha mashaqqatli, shuningdek, ko'p vaqt sarflashni taqozo etadi. Axborot texnologiyalarining tilga ta'siri natijasida terminologiya sohasida ham kompyuter qurilmalari keng qo'llanila boshladi. Terminologik ishlarni mexanizatsiyalash va avtomatlashtirish termin ma'lumotlarini yig'ish, saqlash va qayta ishlash kompyuter yordamida amalga oshiriladigan terminologik ma'lumot bazalarini yaratishni ta'minladi. Terminologik ma'lumot bazalari XX asrning 60-yillarida rus tilining mashina fondini yaratish g'oyasidan yuzaga kelgan. Terminologik ma'lumotlar bazasi – turli ixtisoslikka oid terminlar haqidagi ma'lumotlarni o'z ichiga olgan, foydalanuvchilarning axborot ehtiyojlarini qondirish uchun xizmat qiladigan elektron tizimdir. U barcha mavzular bo'yicha terminologik ma'lumotlarni elektron shaklda yozib olish va doimiy ravishda yangilab borish imkoniyatini beradi. Ushbu ma'lumot bazasidagi har bir termin metama'lumotlar, ya'ni termin haqida qo'shimcha ma'lumotlar bilan ta'minlangan:

Terminning ma'nosi

Boshqa tillardagi ekvivalenti

Qisqa shakli

Sinonimi (maqbul va nomaqbul)

Qo'llash doirasi to'g'risidagi ma'lumot

Terminning qanday hujjat bilan tasdiqlanganligi (masalan, milliy yoki xalqaro standart) va h.k. Terminologik ma'lumot bazalariga, asosan, terminning kelib chiqishi, matnlarda ishlatilishi, nutqning bir qismini ko'rsatishi kabi ma'lumotlarni kiritish mumkin. Baza terminga so'z sifatida yondashib, uning o'ziga xos xususiyatlaridan kelib chiqqan holda tushuncha haqida ham ma'lumot beradi: 1) muayyan muallif asarining yaxlit matni; 2) muayyan muammolar bo'yicha barcha matnlar (matnlar parchalari); 3) tanlangan maxsus terminlar; 4) matnning butun so'z boyligi; 5) matnning chastotali lug'ati; 6) chizmalar, diagrammalar mavjudligi yoki yo'qligi 7) matnda terminlarning qo'llanilishi to'g'risida.

Terminologik ma'lumot banklari jiddiy ilmiy izlanishlar olib borishning kuchli vositasi bo'lib, ularni yaratilish maqsadiga qarab ikki guruhga bo'lish mumkin¹⁶:

a) ilmiy-texnik adabiyotlar va hujjatlar tarjimasini amalga oshirishga yo'naltirilgan ma'lumotnoma tipidagi banklar. Tarjimonlar uchun xoslangan terminologik ma'lumot bazalari ma'lum bir tushunchaning boshqa birliklar bilan munosabatini ko'rsatmasdan, uning o'ziga xos xususiyatlariga oid ma'lumotlarni ifoda etadi. Ular ma'lumotnoma tipidagi banklar deb nomlanadi. Tanlangan tushunchalarning o'zaro munosabati faqat belgilangan predmet kodi asosida bir doiraga mansub ekanligi bilan izohlanadi. Bu orqali hujjatlarga asoslangan tasnif bo'yicha terminlar ro'yxatini tuzish imkoniyati yaratiladi. Ushbu yondashuv

¹⁶ Булдакова Е.В. Компьютерная лексикография / конспект лекций. – Санкт-Петербург: Балт. гос. техн. ун-т, 2009

ma'lum bir foydalanuvchilar guruhi uchun qulaylik kasb etadi. Ular lug'atni shunday tuzadiki, har bir birlik o'xshash va farqli jihatlariga ega bo'lib, ko'rsatilgan kalit so'zlar haqida ma'lumotlar beradi. b) standartlashtirilgan va tavsiya etilgan atamalar to'g'risida ma'lumot berish uchun mo'ljallangan lug'at tipidagi banklar.

Ilmiy tadqiqotlar uchun, birinchi navbatda, predmetlar sohasi, uni tashkil etuvchilarning konseptual tuzilishi muhimdir. U tushunchalarning o'zaro bog'liqligini, ekstralingvistik ma'lumotlar mazmunini, jumladan, terminlar olingan manbalarni batafsil aks ettirib turadi. Leksik tipdagi banklar konseptual tizim va tushunchalar o'rtasidagi munosabatni yoritish maqsadida tuziladi va terminologik lug'atlar uchun asos bo'lib xizmat qiladi. Terminologlar, avvalo, yirik korxonalar, muassasalar va ixtisoslashgan tashkilotlarda yagona yoki ko'p tilli terminologik ma'lumotlar banklarini tashkil qiladi. Terminologik ma'lumotlar bazasida katta korxona yoki muassasa ishi uchun zarur bo'lgan barcha maxsus ixtisoslashtirilgan kengaytirilgan tushunchalar mavjud bo'lib, ularga ta'riflar berilgan.

Terminologik ma'lumot bazalarini yaratishda turli me'yorlarga amal qilinadi. Bular quyidagilar:

- 1) Bazadan foydalanish imkoniyatiga ko'ra: a) ochiq; b) cheklangan;
- 2) Tegishlilik darajasiga ko'ra: a) jamoat mulki; b) xususiy mulk;
- 3) Kirish imkoniyatiga ko'ra: a) pulli; b) bepul;
- 4) Yaratilishiga ko'ra: a) mutaxxassislar tomonidan tuzilgan; b) ochiq tarmoq hamjamiyati tomonidan tuzilgan;
- 5) Mavzu ko'lamiga ko'ra: a) tematik; b) global;
- 6) Ma'lumotlar tiliga ko'ra: a) bir tilli; b) ko'p tilli.

Ko'p tilli terminologik ma'lumot bazalarini yaratish har qanday yirik ko'p tilli loyihalar uchun odatiy ishlarning bir qismidir, u yangi terminlarni ishlab chiqishda

yagona terminologiyani saqlashga va nomuvofiqliklarning oldini olishga yordam beradi. Mukammal darajada yaratilgan terminologik ma'lumot banklari sifatli tarjimaning kafolatidir. Hozirgi vaqtda turli mamlakatlarda bir necha o'nlab yirik terminologik ma'lumot bazalari mavjud. Rossiyada eng katta terminologik ma'lumotlar bazalari standartlashtirish va sifat bo'yicha keng qamrovli ma'lumotlarning Butunrossiya ilmiy-tadqiqot instituti hamda Butunrossiya ilmiy-texnik axborot institutida joylashgan. Xalqaro standartlashtirish tashkiloti (ISO – International Standartization Organization) tomonidan ham konseptual terminologik ma'lumot bazalari yaratilgan va bu baza barcha ISO standartlari uchun konseptual terminologik tarkibni topish, ishlab chiqish va saqlash uchun izchil platformani taqdim etadi, shuningdek, uchta muhim toifadagi qidiruvlarga imkon beradi:

Terminlar va ularning ta'rifi

Grafik belgilar

Kodlar (mamlakatlar, valyutalar, tillar va boshqalar)¹⁷

Baza rivojlanib borishi bilan uning qo'shimcha imkoniyatlari doirasi ham kengayib boraveradi. Shuningdek, Yevropa Ittifoqining barcha terminologik manbalari uchun veb-interfeysni yaratish, ma'lumotlardan foydalanish ko'lamini oshirish va Yevropa Ittifoqi organlarida standartlashtirishni ta'minlash maqsadida Yevropa Ittifoqi hujjatlarida ishlatiladigan terminlarning ko'p tilli ma'lumotlar bazasi – IATE (Interfaol Terminology for Europe)¹⁸ ishlab chiqilgan. IATE 2004-yildan Yevropa Ittifoqi organlari va tashkilotlarida qo'llanila boshlandi. Ochiq foydalanuvchi interfeysi 2007-yilda sinov uchun e'lon qilindi va rasmiy ravishda ishga tushirildi. IATE Yevropa Ittifoqi tarjima xizmatlarining barcha mavjud terminologik bazalarini taxminan 1,4 million ko'p tilli maqolalarni o'z ichiga olgan

¹⁷ <http://cdb.iso.org/>.

¹⁸ <http://iate.europa.eu>

idoralararo ma'lumotlar bazasiga birlashtirdi. Germaniyaning “Siemens” kompaniyasining terminologik ma'lumot bazasi dunyodagi eng yirik ma'lumot qobig'i bo'lib, sakkizta dunyo tillarida taqdim etilgan 2,5 millionga yaqin terminologik birliklardan tarkib topgan. Shunday nufuzli Yevropa terminologik ma'lumot banklaridan yana biri – TermPlusdir. Ushbu tizim dunyoga mashhur DANTERM (Daniya terminologiya markazi) ma'lumotlar bazasidan kelib chiqqan bo'lib, 1978-yilda Kopengagen Iqtisodiyot maktabi asosida shakllana boshladi. TermPlus – bu maxsus foydalanuvchilar guruhiga mo'ljallangan, yangilangan va takomillashtirilgan DANTERMdir. TermPlus ma'lumotlar banki mantiqiy konseptual sxemaga asoslangan bo'lib, u tizimga yangi terminlar qo'shibgina qolmay, balki barcha mavzu bo'yicha ma'lumotlarni har xil tillarda import qilish imkoniyatini beradigan tarzda ishlab chiqilgan. TermPlusning asosiy maqsadi tillarni keng qamrab olish bo'lgani uchun u asosiy Yevropa tillarida keltirilgan iqtisodiyot, tibbiyot, mashinasozlik, axborot texnologiyalari, kimyo, biologiya kabi fanlarning tushunchalarini birlashtiradi va butun dunyo bo'ylab foydalanuvchilar uchun ochiq. TermPlusdagi termin ta'rifi ekspertlar tomonidan tasdiqlangan va lug'at tarkibi ixcham shaklda berilgan, istalgan foydalanuvchi qo'shimcha havolalarga murojaat qilib, termin to'g'risida batafsil ma'lumot olishi mumkin. TermPlus ixtiyorida yetti til mavjud (ingliz, nemis, fransuz, daniya tili, ispan, italyan va rus). Terminologik ma'lumot banklari taraqqiyotida kompyuterning dasturiy ta'minoti va undagi ma'lumotlar asosiy rol o'ynaydi. Hozirgi vaqtda ma'lumotlar bazasi bilan ishlashga mo'ljallangan ko'plab tizim mavjud: SQL, MySQL, Oracle, Access. Katta hajmdagi ma'lumotlar bilan ishlash doim ham qiyin kechadi, vaholanki, har bir tizim o'ziga xos afzallikka ega. Terminologik banklarni yaratish va yangilash maqsadida maxsus Multiterm dasturi yaratildi, u SDL Trados (Germaniyaning Trados GmbH kompaniyasi tomonidan ishlab chiqilgan avtomatlashtirilgan tarjima tizimi)da terminologiyani boshqarish uchun



ishlatilmoqda¹⁹. Tizimning asosiy maqsadi manba matnini tayyorlash bosqichidan tarjimaning oxirigacha mazmunan izchillik va yuqori sifatni ta'minlashdir. MultiTerm terminologiyani boshqarishga semantik yondashadi. Bitta lug'at maqolasida bitta tushuncha bilan bog'liq barcha terminlar, shuningdek har qanday tavsiflovchi va tasniflovchi ma'lumotlar mavjud. Ushbu yondashuv ko'p tilli terminologiya asoslarini yaratishga, shuningdek lug'at maqolasida terminning hayotiy siklini aks ettirishga imkon beradi: yangi yoki eskirgan, maqbul yoki nomaqbul kabi. Bu to'g'ri terminni tanlashni ancha osonlashtiradi. Arxivlar bo'yicha Xalqaro Kengash (ICA) arxiv terminologiyasining ko'p tilli ma'lumotlar bazasini²⁰ yaratish loyihasi rahbari Luciana Duranty ushbu terminologik ma'lumot bazalarining zaruratini quyidagicha bayon etadi: “Terminologiya til bilan birgalikda mavjud va doimiy ravishda o'zgaradi. Bu jarayon muammoni birdaniga hal qiladigan yakuniy, mukammal lug'at yaratishga qaratilgan, ko'p mehnat talab etadigan loyihalarni muvaffaqiyatli yakunlashga monelik qiladi. Chunki bizning sohamizda kuzatilayotgan o'zgarish va o'sish sur'atlarini hisobga olgan holda, natija nashrlarda allaqachon eskirgan bo'lar edi. Shuning uchun men terminologik ma'lumotlar bazasini raqamli tarmoqlar hamda arxivchi va hujjat yuritish bo'yicha mutaxassislar ishtirokidagi ko'ngillilar mehnati (crowdsourcing)dan foydalanadigan dinamik resurs deb bilaman.” Darhaqiqat, terminologik ma'lumot bazalari turli soha mutaxassislari ehtiyoji-talabiga javob beradigan keng ko'lamli ma'lumotlar majmuidir. U ilmiy va texnik adabiyotlarning mashina tarjimasida, avtomatlashtirilgan axborot tizimlarini lingvistik qo'llab-quvvatlashda, terminologiyani tartibga solishda, lug'atlar yaratishni avtomatlashtirishda muhim ahamiyat kasb etadi. O'zbek tilshunosligida terminologik ma'lumot bazalari hali yaratilmagan. Bugungi kunda terminologik ma'lumot bazalari yaratishga doir dastlabki ilmiy izlanishlar N.Abdurahmonovanning ishlarida ko'zga tashlanmoqda:

¹⁹ <https://www.tra-service.ru/multiterm>

²⁰ <http://www.ica.org/14282/multilingual-archival-terminology/multilingual-archival-terminology.html>

“Termin sifatida yangi kirib kelgan so‘zlarning tarjimasini to‘g‘ri tushunishda hamda muayyan terminning bir necha sohada qo‘llanib, turli ma‘no kasb etgani bois ularning lingvistik bazasini yaratish uning nechog‘lik dolzarbligini ko‘rsatadi”²¹. Ochiq manbali terminologik ma‘lumotlar bazasini loyihalashtirishda quyidagi ishlarni bajarishimiz kerak bo‘ladi:

1. Fan va sohalarga oid terminologik adabiyotlardan foydalangan holda boshlang‘ich terminologik ma‘lumotlar bazasini yaratib olishimiz lozim;
2. Internetda onlayn qidiruv tizimidan foydalanib ma‘lumotlar bazasini boyitib borishni yo‘lga qo‘yish;
3. Fan va sohalar mutaxassislari o‘z sohasining yangi terminlarini bazaga kiritish imkoniyatini yaratilishi;
4. Terminologik ma‘lumotlar bazasini ochiq manbali qilish imkoniyatlarini ko‘rib chiqish;
5. Ochiq manbali terminologik ma‘lumotlar bazasini aniqliligini oshirish va uni himoyalash. Fan va sohalarni o‘rganishda terminologik ma‘lumotlarning umumiy bazasi yo‘qligi qator qiyinchiliklarga olib keladi. Terminologik ma‘lumotlar bazasi fanlarni mukammal o‘rganishda, ya‘ni terminlar ma‘nosi va mohiyatini tushunishga imkoniyat yaratadi.

Фойдаланилган адабиётлар рўйхати:

1. Кутузов А.Б Корпусная лингвистика. – (Электрон ресурс): Лицензия Creative commons Attribution share-Alike 3.0 Unported (Электрон ресурс) - //lab314.brsu.by/kmp-lite/kmp-video/CL/CorporeLingva.pdf.
2. Булдакова Е.В. Компьютерная лексикография / конспект лекций. – Санкт-Петербург: Балт. гос. техн. ун-т, 2009
3. Abduraxmonova N. Ochiq manbali terminologik ma‘lumotlar bazasini loyihalashtirish/ ARES: Vol. 2, Issue 1, 2021. – 4-10 b.

²¹ Abduraxmonova N. Ochiq manbali terminologik ma‘lumotlar bazasini loyihalashtirish/ ARES: Vol. 2, Issue 1, 2021. – 4-10 b.



4. Хамроева Ш.М. Корпус лингвистикаси атамаларининг қисқача изоҳли луғати. – Т.: Globe Edit, 2020. – 96 б.
5. Mengliyev B.R., Karimov R.A. Theoretical fundamentals of Uzbek-english parallel corpus./ilmiy maqola, Journal of Critical Reviews, 2020
6. Karimov R.A., Mengliyev B.R. The Role of the Parallel Corpus in Linguistics, the Importance and the Possibilities of Interpretation./ilmiy maqola, International Journal of Engineering and Advanced Technology, 2019
7. Гарабик Р., Захаров В.П. Параллельный русско-словацкий корпус // Труды международной конференции "Корпусная лингвистика – 2006". – СПб., 2006. –81-82 с.