

**RAQAMLI ASRDA TIL VA TEXNOLOGIYA: KOMPYUTER
LINGVISTIKASINING EVOLYUTSIYASI**

SADINOV OTABEK BAXODIROVICH

Navoiy davlat universiteti Aniq fanlar fakulteti talabasi.

Annotatsiya: *Ushbu ishda kompyuter lingvistikasi fanining shakllanish tarixi, rivojlanish bosqichlari hamda uning raqamli texnologiyalar bilan integratsiyasi yoritilgan. Ayniqsa, inson tilini avtomatik tarzda tahlil qilish, tarjima qilish, tushunish va sun'iy intellekt tizimlarida qo'llash masalalariga alohida e'tibor qaratilgan. Maqolada kompyuter lingvistikasi nafaqat texnik fan sifatida, balki tilshunoslik, informatika va matematika kesishgan nuqtadagi muhim yo'nalish sifatida baholanadi. Shuningdek, o'zbek tilida olib borilayotgan tadqiqotlar va zamonaviy yondashuvlar haqida ham ma'lumotlar keltiriladi. Mazkur mavzu ta'lim jarayonida til va texnologiyani uyg'un holda o'rgatishga xizmat qiladi.*

Kalit so'zlar: *Mashina tarjimasi, mashinali o'rganish, avtomatik tahrir, Lingvistik modellashtirish, kompyuter lingvistikasiga e'tibor, qozoq kompyuter lingvistikasi, o'zbek tilining milliy korpusi, Til va texnologiya integratsiyasi.*

Bugungi global raqamli muhitda axborot texnologiyalari va kompyuter tizimlari barcha sohalarda muhim o'rin egallab bormoqda. Kompyuter vositalarining hayotimizning turli jabhalariga jadal kirib borishi tilshunoslik sohasiga ham bevosita ta'sir ko'rsatdi. Aynan shu integratsiya natijasida yangi tarmoqlardan biri — kompyuter lingvistikasi shakllana boshladi. Aytish joizki, ushbu atama dastlab paydo bo'lganida darhol bugungi nomi bilan yuritilmagan. XX asrning 1950-yillaridan e'tiboran tilshunoslik fanida “mashinali tarjima” va “matn lingvistikasi” kabi iboralar qo'llanila boshlandi.

Mashinali tarjima deganda, biror tilga oid matnni boshqa tilga kompyuter

yordamida qisqa vaqt ichida avtomatik tarzda tarjima qilish jarayoni tushuniladi. Bunday jarayon davomida tarjimaning grammatik to'g'riligi hamda semantik aniqligini ta'minlash ustuvor vazifa sifatida qaraladi. Aynan shu davr mobaynida lingvistik bilimlar bilan kompyuter texnologiyalarining o'zaro aloqasi bosqichma-bosqich shakllana boshlagan.

1954-yil 7-yanvar sanasida Jorjtaun universiteti hamda IBM kompaniyasining hamkorligida ishlab chiqilgan IBM II mashina tarjima tizimi ilk bor sinovdan o'tkazildi. Ushbu tajriba kompyuter lingvistikasi fanining shakllanishida muhim tarixiy bosqichlardan biri bo'lib xizmat qildi. Dastlab bu yo'nalish turli nomlar bilan atalgan bo'lsa-da, Ruslan Mitkovning qayd etishicha, "Computational linguistics" (kompyuter lingvistikasi) atamasi 1960-yillarda AQSh Milliy fanlar akademiyasi qoshidagi "Tilni avtomatik qayta ishlash bo'yicha maslahat qo'mitasi" a'zosi David Hays tomonidan ilmiy muomalaga kiritilgan.

Avtomatik tarjima tizimi kompyuter lingvistikasining rivojlanib, taraqqiy etishida muhim bosqich bo'ldi. Chunki, bu tizimning ham kamchiliklari bor edi. Asta-sekin tabiiy tilni qayta ishlash tizimlari va formal grammatika nazariyalari ishlab chiqila boshladi. Chunki, bu tizimning ham kamchiliklari bor edi. Ya'ni so'zlar avtomatik tarjima jarayonida to'g'ridan-to'g'ri tarjima qilinadi. Bunda ko'p ma'nolilik, omonimlik va iboralardagi ma'no noto'g'ri tarjima qilingan. Shuning uchun tarjima grammatik qoidalarga rioya etilmagan tarzda amalga oshirilgan va matnni uslub jihatdan tahrirlab chiqish zarur bo'lgan. Bu bo'yicha tadqiqotlar hali ham davom etib kelmoqda. Jahon axborot tizimining rivojlanishi bilan tarjima tizimlari internetda ham yo'llga qo'yilgan. Biroq tarjima tizimida hali ishlanishi kerak bo'lgan ishlar juda ko'p. Sababi, o'zbek tilidagi ma'lum bir matnni biror xorijiy tilga o'girmoqchi bo'lsak yoki buning teskarisi bo'lsa, yuqorigidagiday xatoliklarni kuzatishimiz mumkin. Kompyuter lingvistikasi bo'yicha jahon mamlakatlarida juda ko'plab tadqiqotlar olib borilmoqda. Jumladan, matn tahriri, avtomatik tarjima, kompyuter leksikografiyasi, information qidiruv tizimi, sun'iy intellekt, til o'qitishni avtomatlashtirish yo'nalishlari bo'yicha salmoqli ishlar amalga oshirilgan. AQSHda kompyuter lingvistikasi bo'yicha tematik to'plamlar

va hatto shu nomdagi “Computational linguistics” jurnali ham chop etiladi. Kompyuter tilshunosligi assotsiatsiyasi tuzilgan bo‘lib, bu uyushma

Ko‘plab ilmiy ishlarni amalga oshiradi. AQSHning o‘ndan ortiq universitetlarida kompyuter tilshunosligi fakultetlari mavjud bo‘lib, ular kompyuter lingvistikasining muammoli masalalarini o‘rganadilar. Rossiyada har yili kompyuter lingvistikasiga oid “Dialog” konferensiyasi o‘tkaziladi.

Kompyuter lingvistikasining mashina tarjimasi bo‘yicha qilingan ishlar bugungi kunda barcha kompyuterlarda mavjud bo‘lgan ruschadan boshqa tillarga tarjima qiluvchi dasturlarning yaratilishiga asos bo‘ldi. Xususan, L.L.Nelyubin va uning shogirdlari tomonidan ishlab chiqilgan so‘zma-so‘z tarjima qilishning ilmiy strategiyasi o‘sha davr jahon injener lingvistikasi talablari va ilg‘or tajribasiga monand bo‘lgan edi.³ Nelyubin iboralarning to‘g‘ri tarjimasini amalga oshiruvchi algoritmlar ishlab chiqqan va ingliz tilidagi rasmiy hujjatlarni rus tiliga tarjima qiluvchi dastur ishlab chiqqan.

Avtomatik tahrir bo‘yicha ham bir qator rus olimlari ilmiy tadqiqot olib borganlar. V.E.Berzon, V.G.Britvin, I.A.Melchuk, I.S.Duganova kabilarni keltirish mumkin. Matematik lingvistika Qozog‘istonda izchil rivojlangan. Dastlab, statistik metodga katta e‘tibor qaratilgan. Keyinchalik boshqa yo‘nalishlarda ham keng ko‘lamli ishlar olib borilgan. Qozoq kompyuter lingvistikasi shakllanishida Q.Bektaevning xizmatlari katta. Uning boshchiligida M.Avezovning 4 kitobdan iborat “Abay yo‘li” romani tilining chastotali lug‘atini 50 soat mobaynida tuzib chiqqanlar. Mazkur lug‘atda 20.000 leksema va 60.000 so‘z shakli bo‘lib, u 466.000 marta qo‘llangan. Agar ushbu ish qo‘lda bajarilsa albatta uzoq vaqt olishi tayin.

O‘zbek kompyuter lingvistikasida ham bir qancha olimlarimiz ilmiy izlanish olib borganlar va bu hozirgacha davom etib kelmoqda. S.Rizayev, S.Muhammedov va boshqalar dastlabki tadqiqotlarni olib borgan olimlar hisoblanadi. Ular asosan, statistik tahlil va avtomatik tahrir yo‘nalishlari bo‘yicha ishlarni amalga oshirganlar. O‘zbek tilidagi matnlar uchun mo‘ljallangan dasturlar ham yaratilgan.

Lotin alifbosidagi matnlarni kirill yozuviga va lotin alifbosidan kirillga o'tkazuvchi "Spellchecker", matnlarga qayta ishlov beruvchi "Uzlington" kabilar bunga misol.

Kompyuter lingvistikasining yurtimizda rivojlanishi bilan o'zbek tilining milliy korpusi hozirda qisman yaratildi. Uzbekcorpus.uz havolasi orqali biz ba'zi so'zlarning grammatik tavsifini bilib olishimiz mumkin. Shuningdek, lingvistik lug'atlarning imkoniyatlaridan foydalanishimiz mumkin. Lekin, shuni ta'kidlash kerakki, milliy korpus hali to'liq shakllangani yo'q. Chunki, unda, asosan, mahalla va oilaga oid bo'lgan so'zlarga bor, xolos. Buning boisi, korpus "Mahalla va oilani qo'llab quvvatlash vazirligi" tomonidan moliyalashtirilgan. Umid qilamizki, yaqin kelajakda boshqa turli sohalarga oid so'zlar ham milliy korpusimizga qo'shiladi va yurtimizda kompyuter lingvistikasining rivojlanishida asosiy tayanch bo'lib xizmat qiladi.

FOYDALANILGAN ADABIYOTLAR:

1. Po'latov A., Muhamedova S., Kompyuter lingvistikasi. – Toshkent, 2007.
2. Nelyubin L. "Перевод и прикладная лингвистикаю".
3. The Oxford handbook of computational linguistics. R.Mitkov. – 2003.
4. Grishman R. Computational linguistics. 1994