

ISSN 2181-922X

LANGUAGE & CULTURE

UZBEKISTON

UZBEKISTAN

TIL VA MADANIYAT

**KOMPYUTER
LINGVISTIKASI**

2023 Vol. 3 (6)

www.compling.tsuull.uz

MUNDARIJA

Mavjuda Alimbekova

Abdurauf Fitrat mualliflik korpusini yaratishning ahamiyati.....6

Madinabonu Qodirova, Shahlo Hamroyeva

Zamonaviy dunyoda mashina tarjimasini tadriji:
tahlillar va natijalar.....22

Noila Matyakubova

"Aligner" dasturiy vositasi uchun o'zbek-ingliz tilida sifat va uning
darajalarining morfologik tahlili.....41

Mohiyaxon Uzoqova, Mansurbek Narzullayev

Sinonimayzer dasturida RoBERTaForMaskedLM modelidan leksik
sinonimlarni aniqlash uchun foydalanish.....54

Dlafroz Xudoyqulova

O'zbek-ingliz farmatsevtika terminlari korpusli lingvistik
ta'minotining milliy-madaniy asoslari.....69

Ruhillo Alayev, Gulshaxnoz Maxmudjonova

O'zbek tilidagi matnli hujjatlarda izlashni amalga
oshirishni takomillashtirish.....78

Sanjarbek Baxodirov

Tabiiy tilni qayta ishlashda matn tozalash tizimini
ishlab chiqish.....91

Azizaxon Raxmanova

Sun'iy intellekt yordamida o'zbek va ingliz tili lingvistik asoslarini
o'qitishning zamonaviy uslublari.....106

ZAMONAVIY DUNYODA MASHINA TARJIMASI TADRIJI: TAHLILLAR VA NATIJALAR

**Madinabonu Qodirova¹,
Ilmiy rahbar: Shahlo Hamroyeva²**

Annotatsiya. Bugungi kunda mashina tarjimasi sohasi jadal rivojlanib borayotganligi hech kimga sir emas. Bu borada dunyo ilm-fanida amalga oshirilayotgan tadqiqotlar, yaratilayotgan nazariyalar haqida to'xtalib o'tmaslikning iloji yo'q. Ushbu g'oyalarni o'rganish va lozim bo'lsa o'zbek mashina tarjimasi yo'nalishiga tatbiq etish bu sohaning jadalroq rivojlanishiga turtki bo'ladi, degan fikrdamiz. Ushbu maqolada bugungi kunda jahon mashina tarjimasi sohasida amalga oshirilgan ishlar, avtomatik tarjimani optimallashtirish uchun taklif etilayotgan nazariyalar, g'oyalar haqida fikr yuritamiz. Ularni o'zaro qiyoslab, tahlil qilamiz. Shuningdek, hozirda mashhur tarjimon dasturlarining ishlash prinsiplariga to'xtalib o'tamiz, natijalar olamiz.

Kalit so'zlar: *avtomatik tarjima, RBMT, SMT, neyron tizimlar, NMT, nurli qidiruv tizimi, dinamik yondashuv, qayta tuzilish.*

Kirish

Globalashuv va faol integratsiyalashuv kuzatilayotgan bugungi kunda davlatlar, millatlar o'rtasidagi aloqalar kuchayishi tabiiy. Keyingi yillarda ham bu jarayonlar yanada shiddatli tus olishi kutilmoqda. Bunday sharoitda tillararo to'siqlarni olib tashlash zamonaviy ilm-fanning dolzarb vazifalaridan biri sanaladi. Bu borada butun dunyo, xususan, Yevropa, Osiyo, Amerika olimlari muhim izlanishlarni amalga oshirmoqdalar. Shu o'rinda aytib o'tish kerakki, 1988-yildan buyon nashr etib kelinayotgan va 2009-yildan boshlab ochiq kirish imkonini taqdim etgan xalqaro "Kompyuter lingvistikasi" ("Computer linguistics³") jurnali ushbu tadqiqotlarni keng targ'ib qilinishi va ish-

¹Qodirova Madinabonu Murodjon qizi – Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti tayanch doktoranti.

E-pochta: begonammmm@gmail.com

ORCID: 0009-0005-8715-8074

²Hamroyeva Shahlo Mirdjonovna – filologiya fanlari doktori, dotsent. Alisher Navoiy nomidagi Toshkent davlat o'zbek tili va adabiyoti universiteti.

E-pochta: shaxlo.xamrayeva@navoiy-uni.uz

ORCID: 0000-0002-5429-4708

³<https://cljournal.org/>

lab chiqarishga joriy etishda muhim tayanch bo'lib xizmat qilmoqda. Bu jurnal hisoblash tilshunosligi va tabiiy tilni qayta ishlash bo'yicha tadqiqotlar uchun asosiy arxiv forumidir.

Shuningdek, AQSHdagi Shimoliy Kaliforniya universiteti, Germaniya IBM tadqiqot markazi, Germaniya Axen Universiteti, Google, Unbabel, Alibaba kabi yirik korporatsion tashkilotlar mashina tarjimasi rivojiga katta hissa qo'shmoqda.

Mashina tarjimasi g'oyalari bilan deyarli bir vaqtda tug'ilgan til qoidalariga asoslangan mashina tarjimasi (RBMT) yondashuvi mukammal tarjimani ta'minlash imkoniga ega emasligi tasdiqlandi. Hozirgi davr zamonaviy mashina tarjimasi sohasi bu vazifani amalga oshirishda statistik mashina tarjimasi (SMT) va, ayniqsa, 2016-yildan boshlab tarjima tizimiga joriy etish boshlangan neyron tizimlarga asoslangan mashina tarjimasi (NMT) yondashuviga katta umid bog'lamoqda. Zero, keyingi yillardagi avtomatik tarjima yo'nalishidagi tadqiqotlar asosan shu sohalarga qaratilganligiga guvoh bo'lamiz.

Statistik mashina tarjimasiga doir tadqiqotlar tahlili

Yevropa va AQSHda amalga oshirilgan tadqiqotlar.

2000-2016-yillarga oid tadqiqotlarning ko'pchiligi SMT yondashuviga asoslanadi. Xususan, Germaniyadagi Axen texnologiya universiteti olimlari Franz Josef Och va Herman Neylar "Statistik mashina tarjimasiga tizimli shablon yondashuvi" [Och, Ney, 2004. 55] nomli maqolalarini nashr ettirdilar. Unga ko'ra, tarjima modelida so'zlarning konteksti hisobga olinadi va so'z tartibidagi mahalliy o'zgarishlarni manba tildan maqsad tiliga yuqori aniqlikda olib o'tish mumkin. Model log-chiziqli modellashtirish usuli yordamida tasvirlanadi. Olimlarning fikricha, klassik statistik mashina tarjimasi tizimlariga qaraganda bu usul yordamida modelni kengaytirish osonroq bo'lishi kutiladi.

Axen texnologiyalar universitetining yana bir guruh olimlari statistik tarjimada dinamik dasturlash texnologiyalarini qo'llashni taklif etib chiqdilar. K. Tillman va H. Neylar o'zining "Statistik mashina tarjimasi uchun so'zlarni qayta tartiblash va dinamik dasturlash nurlarini qidirish algoritmi" [Tillman, Ney, 2003] nomli maqolasida o'z g'oyalarini bayon etganlar. Ushbu maqolada dinamik dasturlash asosida statistik mashina tarjimasi uchun samarali nurli qidiruv algoritmi tasvirlanadi. Ular samarali qidiruv algoritmiga erishish uchun manba va maqsad til o'rtasidagi qo'llash mumkin bo'lgan so'zlarni qayta tartiblashning yangi usulini taqdim etadilar. Tadqiqot

obyekti sifatida nemis tilidan ingliz tiliga tarjima yo'nalishi tanlab olingan. Unga ko'ra so'zlarni qayta tartiblashni nazorat qilish uchun to'rtta parametr to'plami kiritiladi, keyinchalik ularni yangi tarjima yo'nalishlariga osongina o'tkazish mumkin bo'ladi. Hermann Neyning tadqiqotlari keyinchalik ham davom etadi. U Sonya Niessen bilan birga "Morfo-sintaktik ma'lumotlardan foydalangan holda kam manbalar bilan statistik mashina tarjimasi" [Nießen, Ney, 2004] nomli tadqiqotini e'lon qiladi. Bu tadqiqot ishida u an'anaviy qoidalarga asoslangan tarjimani statistik tarjimadan butunlay ajratib qo'yishga harakat qiladi. Uning fikricha, statistik mashina tarjimasida manba va maqsad tildagi so'zlar o'rtasidagi muvofiqlik parallel korpuslar asosida tekshiriladi va ko'pincha asosiy modellarni tuzish uchun lingvistik bilimlardan foydalanilmaydi. Xususan, mashina tarjimasi uchun mavjud statistik tizimlarda ko'pincha bir lemmaning turli xil flektiv shakllari bir-biridan mustaqil so'zlar sifatida qaraladi. Shuningdek, ular so'zlarning ekvivalentlik sinflari asosida iyerarxik leksika modellarini qurishni taklif qiladilar. Bundan tashqari, bog'langan jumalardagi so'z tartibini o'zlashtirishga o'rgatilgan dasturiy tizim yaratish taklif qilinadi.

Hermann Ney ushbu fikrlarini keyinchalik "Mashina tarjimasi uchun so'z darajasidagi ishonchni baholash" [Ueffing, Ney, 2007] nomli tadqiqotida davom ettiradi. U mashina tarjimasida so'z darajalari orasida ishonchni baholashning bir necha usullari taqdim etadi. Ular avtomatik tarjimadagi har bir so'zni to'g'ri yoki noto'g'ri deb belgilash imkonini beradi. Maqolada aytilishicha, ishonchni baholash yondashuvlari so'zning keyingi ehtimolliklari (posterior ehtimolligi) ga asoslanadi. So'zning posterior ehtimoli haqida turli tushunchalar hamda ularni hisoblashning turli usullari taqqoslanadi. Ularni ikki toifaga bo'lish mumkin: dastur tomonidan qilingan tarjimani o'rganuvchi tizimga asoslangan usullar va tarjima tizimidan mustaqil bo'lgan to'g'ridan-to'g'ri usullar. Posterior ehtimollik tarjima qilish maqsad qilingan tildagi so'zlarni o'z ichiga olgan tarjima gipoteza maydonidagi jumalarning ehtimolliklarini yig'ish orqali aniqlanadi.

Axen universiteti tadqiqotchilari mashina tarjimasi natijalarini baholash va xatolarni tahlil qilish masalasiga ham alohida yondashganlar. Xususan, M.Popovich va H.Ney 2014-yilda "Mashina tarjimasi natijalarini baholash va xatolarni tahlil qilish" nomli maqolasini e'lon qildilar [Popovich, Ney, 2014]. Ushbu maqolada xatolarni avtomatik tahlil qilish va tasniflash uchun so'zdagi xatolik (WER) va pozitsiyaviylik xatosi (PER)ni aniqlash algoritmlaridan foydalangan holda haqiqiy noto'g'ri so'zlarni ajratib olish haqida so'z boradi. Bu

tadqiqot muayyan tarjima muammolari haqida aniqroq ma'lumot beruvchi avtomatik baholash choralari ishlab chiqish yo'lidagi dastlabki qadam bo'ldi. Taklif etilayotgan yondashuv tarjima xatolarini turli yo'llar bilan tasniflash uchun lingvistik bilimlarning har xil turlaridan foydalanish imkonini beradi. Ushbu ish beshta xato toifasiga qaratilgan: flektiv xatolar, so'z tartibining buzilishi, so'zlarni tushirib qoldirish, ortiqcha qo'llash va so'zni noto'g'ri tanlash kabilar. Shuningdek, ular avtomatik xatolar tahlili natijalarini inson xatolarini tahlil qilish natijalari bilan solishtirdilar va ijobiy natijalarga erishdilar.

Statistik mashina tarjima borasida amerikalik olimlar o'ziga xos maktab yaratdilar. Shimoliy Kaliforniya va Kembrij shaharlari tadqiqot markazi bo'lib xizmat qildi. Xususan, Shimoliy Kaliforniya universiteti olimlari Aleksandr Frazer va Daniel Marku statistik mashina tarjimasida ikki tildagi so'zlarni o'zaro moslashtirish jarayonining sifat darajasini aniqlashga bag'ishlangan "Statistik mashina tarjima uchun so'zlarni MOSLASHTIRISH sifatini o'lchash" [Fraser, Marcu, 2005] maqolasini e'lon qildilar. Ushbu tadqiqot institutida statistik tarjima bo'yicha Daniel Marku yirik tadqiqotlarni amalga oshirmoqda. Olim D.S.Muntenu bilan birgalikda 2006-yilda "Parallel bo'lmagan korpuslardan foydalanish orqali mashina tarjima samaradorligini oshirish" [Munteanu, Marcu, 2006] nomli tadqiqotini e'lon qildi. Maqolada taqqoslanadigan, parallel bo'lmagan korpuslarda parallel jummalarni kashf qilishning yangi usuli taqdim etiladi. Ular maksimal entropiya klassifikatoridan foydalanadilar. Unga ko'ra bir juft jumla berilganda, ular bir-birining tarjima yoki yo'qligini dastur ishonchli tarzda aniqlay oladi. Shuningdek, ular juda kichik parallel korpusdan (100 000 so'z) boshlab, katta va parallel bo'lmagan korpusdan foydalanish orqali yaxshi sifatli MT tizimini noldan qurish mumkinligini ko'rsatib beradilar. Shunday qilib, bu usul juda kam manbalar mavjud bo'lgan til juftliklari orasida qo'llanilishi mumkin bo'ladi.

Dunyoga mashhur Google kompaniyasi tadqiqotchilari monolingual statistik mashina tarjima yordamida so'rovlarni qayta yozish ustida tadqiqot olib bordilar [Riezler, Yi Liu, 2010]. Ularning fikricha, konyuktiv metod asosida ishlaydigan axborot texnologiyalarida birmuncha muddat avval kiritilgan va kam foydalanilgan so'rovlar tizim xotirasida saqlanmasligi mumkin. Shu sababli o'xshash sintaktik xususiyatlarga ega boshqa atamalarni ham qayta yozish orqali muammoni hal etishni taklif etadilar. Ular foydalanuvchi so'rovlari-ning manba tilidan veb-hujjatlarning maqsadli tiliga tarjima qilish

orqali so'rovlar va hujjatlar o'rtasidagi "leksik jarlik"ni o'rnatish istiqbollari qabul qilish orqali eng yaxshi natijalarga erishish mumkinligini ta'kidlaydilar.

AQShning Massachusetts shahrida joylashgan Rayton BBN Texnologiyalar universiteti tadqiqotchilari Libin Shen va Jinxi Xu statistik mashina tarjimasida matndan bog'liqlikka tamoyilini ilgari surdilar [Shen, Xu, 2010]. Ular jumllararo qaramlik algoritmini taklif etadilar. Ushbu algoritm an'anaviy n-gramm modeli bilan modellashtirish mumkin bo'lmagan so'z munosabatlarida foydalanish uchun dekodlashda maqsadli bog'liqlik tili modelidan foydalanadilar.

Shuningdek, 2010-yilda Kevin Nayt nomidagi axborot fanlari instituti va Daniel Marku Axborot fanlari instituti tadqiqotchilari "Sintaksisga asoslangan mashina tarjimasini uchun qayta tuzilish, qayta belgilash va qayta moslashtirish [Wang va boshqalar, 2010]" maqolasini e'lon qildilar.

Ushbu maqolada standart tahlil qilish va moslashtirish vositalari sintaksisga asoslangan statistik mashina tarjimasini tizimlari uchun maqbul emasligi isbotlab beriladi. Va ularning o'rniga zamonaviy sintaksisga asoslangan MT tizimining aniqligini oshirish uchun uchta modifikatsiya dasturi taqdim etiladi: qayta tuzilish - quyi tuzilmalardan qayta foydalanishni ta'minlash uchun matnning sintaktik tuzilishini o'zgartiradi; ikki xil til materialiga oid so'zlarni qayta kodlash va qayta moslashtirish – noto'g'ri tarjima qilingan so'zlarni olib tashlash va to'g'ri tuzilmalar uchun jumllar bo'ylab so'zlarni birlashtirish. Eng muhimi, bu ishlarning barchasini EHM tomonidan bajarilishiga erishiladi.

2009-yilda Jons Xopkins universiteti Inson tili texnologiyalari markazi amaliy tillarni o'rganish uchun yozgi lager tashkil etdi. 8 hafta davom etgan SCALE-2009 dasturi doirasida jalb etilgan olimlar semantik ma'lumotli mashina tarjimasini (SIMT) bo'yicha resurs va tizim yaratish harakatlarini amalga oshirishgan. Ular semantik ma'lumotli sintaktik MTda modallik va inkordan foydalanishda yangi modallik/inkor (MN) annotatsiya sxemasini va leksikon yordamida yaratgan ikkita avtomatlashtirilgan MN teggerini ommaga namoyish etdilar. Annotatsiya sxemasi modallik va inkorning uchta komponentini ajratib turadi: trigger (modallik yoki inkorni bildiruvchi so'z), maqsad (modallik yoki inkor bilan bog'liq harakat) va tayanch (modallik tajribasi). Ular strukturaga asoslangan MN teggeri standart LDC ma'lumotlar to'plamini teglash uchun taxminan 86% (janrga qarab) aniqlikka olib kelishini ta'kidlaydilar [Baker va

boshqalar, 2010].

Isroillik olimlar esa mashina tarjimasi sifatini oshirish uchun tarjima modellarini tarjima tiliga moslashtirishni taklif etib chiqdilar [Lembersky, Ordan, 2013]. Statistik mashina tarjimasi uchun ishlatiladigan tarjima modellari qo'lda tarjima qilinadigan parallel korpuslardan tuzilgan. Tarjimashunoslik sohasidagi ko'plab tadqiqotlar shuni ko'rsatadiki, tarjimaning yo'nalishi muhim, ammo tarjima tili (o'girish maqsad qilingan til) noyob xususiyatlarga ega bo'lishi tabiiy. Tarjima modelini tarjimaning maxsus xususiyatlariga moslashtirish orqali biz iboralar jadvallarini tuzishda tarjima yo'nalishiga oid ma'lumotlardan foydalanamiz. Va shuningdek, "to'g'ri" va "noto'g'ri" yo'nalishlarda tarjima qilingan matnlardagi iboralar jadvallarini interpolyatsiya qilish orqali aralash modelni yaratishni taklif qilishadi. Bu tarjima sifatining izchil, statistik jihatdan sezilarli yaxshilanishiga olib kelishini taxmin qiladilar.

Amerika, Yevropa, Sharq mamlakatlari olimlari statistik tarjima uchun ko'plab modellarni taklif etdilar. Zero, Chikago shtatida joylashgan Toyota texnologiyalar instituti olimi Kevin Gimpel va Karnegi Mellon Universiteti tadqiqotchisi A.Smitlar "daraxtdan daraxtga" kvazinxon xususiyatlariga ega iboralarga asoslangan mashina tarjimasi g'oyasini ilgari surdilar [Gimpel, Smith, 2016]. 2016-yilda Ispaniyaning Valensiya politexnika universiteti o'qituvchisi Daniel Ortiz-Martinez Statistik mashina tarjimasini onlayn o'rganish kursini taqdim etdi [Martinez, 2016]. Niderlandiyalik olimlar Arianna Bisazza va Marsello Federiko 2016-yilda "Statistik mashina tarjimasida so'zlarni qayta tartiblash bo'yicha tadqiqot: hisoblash modellari va til hodisalari" nomli tadqiqotini e'lon qildilar [Bisazza, Federico, 2016].

Osiyolik olimlar taqiqotlari tahlili. Osiyo mintaqasida ham bu borada salmoqli ishlar amalga oshirilgan. Chunonchi, G.Neubing, Taro Watanabe va Takayama-cho singari yaponiyalik olimlar statistik mashina tarjimasini optimallashtirishning bir qancha usullarini taqdim etdilar [Neubig va boshqalar, 2016].

Resurs kam bo'lgan muhitda mashina tarjimasi uchun manba tilini moslashtirish yondashuvlarini Qatar hisoblash tadqiqot markazi olimi Pidong Vang va Singapur milliy tadqiqot markazi ilmiy xodimi Hve Toular tadqiq etishdi [Wang va boshqalar, 2016]. Shuningdek, kam resursli mashina tarjimasini optimallashtirish borasida Edinburg tadqiqotchilari ham ish olib borganlar. Zero, 2016-yilda B.Haddov, R.Bavden va A.Michellar "Kam resursli mashina tarjimasini o'rganish" nomli yirik tadqiqot ishini e'lon

qildilar [Haddow va boshqalar, 2016].

Xitoy olimlari Dun Deng va Nianwen Xue Mashina tarjimasida xitoy-ingliz tilidagi tarjima farqlari mavzusida empirik tadqiqotlar olib bordilar [Deng, Xue, 2017]. Qatar olimi Shafiq Joty va F.Guzman esa mashina tarjimasida nutq tuzilmasini baholash metodlarini yaratdi [Joty, Guzman, 2017]. N.Durrani va H.Schmidtler bo'lsa N-gram modeli va iboraga asoslangan statistik tarjimani birlashtirib operatsiyalar ketma-ketligi modelini yaratdilar [Durrani va boshqalar, 2021]. Hind olimi Payal Khullar 2021-yilda "Mashina tarjimasida uchun ellipslar muhimmi?" nomli maqolasini e'lon qildi [Khullar, 2021]. Ellips - bu lingvistik hodisa bo'lib, unda jumlaning qismlari tushiriladi va ular nutq yoki real dunyo kontekstidan olinishi kerak. Tadqiqot ishida ellipsning muhimlik va ayrim hollarda qo'llab bo'lmaslik holatlarini dalillar bilan isbotlab berdi. Kanada milliy tadqiqot kengashi olimi Saif M.Muhammad esa tuyg'ularni avtomatik aniqlash texnologiyasini taqdim etdi [Muhammad, 2022].

2000-2016-yillarni tom ma'noda statistik mashina tarjimasining gullab-yashnash davri deb atash mumkin. Bu sohada ko'plab asarlar, monografiyalar, tadqiqot ishlari, ilmiy maqolalar e'lon qilindi. 2009-yilda K.Gutte, N.Kanseda, M.Dimetman va Jorj Fosterlarning "Mashina tarjimasini o'rganish" [Gutte va boshqalar, 2009] 2010-yilda Filipp Kohning "Statistik mashina tarjimasida [Koehn, 2010]", 2016-yilda P.Villiams, R.Senrich, F.Kohlarning "Sintaksisga asoslangan statistik mashina tarjimasida [Williams va boshqalar, 2016]", 2017-yilda L.Spesy, K.Skarton va G.Genriklarning "Mashina tarjimasida sifatini baholash" [Specia va boshqalar, 2017] kabi o'nlab asarlar e'lon qilindi.

Neyron tizimlarga asoslangan mashina tarjimasida

2016-yilda neyron tarmoqlarni mashina tarjimasiga joriy etilishi haqidagi g'oyalar paydo bo'lishi bu sohada tub burilishlar bo'lishiga olib keldi. Bu borada Google kompaniyasi tashabbus ko'rsatib neyron tarmoqlarga asoslangan 20 tilda tarjimani amalga oshirdi. Keyingi yillardagi tadqiqotlarning deyarli barchasi neyron tizimlarga qaratildi. Yaponiyalik olimlar bunga katta qiziqish bildirib kelmoqdalar. Masalan, asli yaponiyalik Microsoft tadqiqotchisi Akiko Eriguchi, Tokio axborot va kommunikatsiyalar iqtisodiyoti universiteti professori Yoshimasa Tsuruoka va Kazuma Hashimotolar neyron mashina tarjimasiga manba tomonidagi iboralar tuzilmalarini kiritish borasida ish olib bordilar [Eriguchi va boshqalar, 2019]. Neyron mashina tarjimasida (NMT) bir nechta tillarda an'anaviy

statistik mashina tarjimasini modeliga yangi muqobil sifatida katta muvaffaqiyat qozondi. Dastlabki NMT modellari ketma-ketlikni o'rganishga asoslangan bo'lib, u vektor fazosiga manba so'zlar ketma-ketligini kodlaydi va vektordan maqsadli so'zlarning boshqa ketma-ketligini hosil qiladi. Ushbu NMT modellarida jumlar ichki tuzilishga ega bo'lmagan so'zlar ketma-ketligi sifatida ko'rib chiqiladi. Ushbu olimlar esa jumlarning sintaktik tuzilishiga e'tibor qaratdilar va NMT daraxtidan ketma-ketlik modeli deb alatadigan yangi sintaktik NMT modelini taklif qiladilar. Ular taklif etgan model dekoderga tarjima qilingan so'zni yaratish va uni iboralar bilan oson moslashtirish imkonini beruvchi mexanizmga ega. Eksperimental natijalar shuni ko'rsatadiki, sintaktik tuzilmadan foydalanish ma'lumotlar to'plami kichik bo'lganda foydali bo'lishi mumkin, lekin ikki yo'nalishli kodlovchidan foydalanish kabi samarali emas. Ma'lumotlar to'plamining hajmi oshgani sayin, sintaktik daraxtdan foydalanishning afzalliklari pasayadi.

Xitoy Fanlar Akademiyasi Hisoblash Texnologiyalari Institutining Intellektual Axborotni qayta ishlash laboratoriyasi tomonidan avtoregressiv bo'lmagan neyron mashinasi tarjimasini uchun ketma-ketlik darajasini sinovdan o'tkazish ishlari olib borilmoqda [Key va boshqalar, 2021]. Olimlar NMT uchun moslashtirilgan bir nechta yangi kuchaytirish algoritmlari asosida ketma-ketlik darajasini optimallashtirish uchun yangi model taklif qiladilar. Bu model gradientlarni baholashda ularning farqini kamaytirish xususiyati orqali an'anaviy usuldan ustun turadi.

Xitoyning Alibaba Group korporatsiyasi tomonidan olib borilayotgan tadqiqotlar qisqa matnlar bilan ishlashdagi muammolarga qaratilmoqda [Wan va boshqalar, 2021].

Qisqa matnlar turli ko'rinishlarda, jumladan, so'rov, dialog va xabar shakllarida mavjud. Neyron mashina tarjimasini (NMT) bo'yicha ko'plab tadqiqotlarning aksariyati qisqa gaplarga emas, balki uzun jummalarga oid ochiq muammolarni hal qilishga qaratilgan. Buning sababi shundan iboratki, insonda o'rganish va qayta ishlashga nisbatan qisqa ketma-ketliklar odatda oson misollar sifatida qabul qilinadi. Tadqiqotchilar eksperimentlar yordamida aslida bunday emasligini isbotlab beradilar. Ularning ta'kidlashicha, jumla qanchalik uzun bo'lsa uning mazmuni aniqlashadi va jumalararo variantlarni kamaytiradi, kontekstli ma'lumotlarning yetishmasligi NMT-ni qisqa jumlar bo'yicha ma'lumotlar noaniqligiga sabab bo'ladi va shuning uchun NMT modeli noto'g'ri tarjimaga olib keladi.

So'nggi paytlarda tabiiy tilni qayta ishlash va sun'iy intellektning

boshqa sohalarida chuqur neyron tarmoqlarning muvaffaqiyatiga qaramay, ularni izohlash qiyinligicha qolmoqda. Dunyoning turli mamlakatlari olimlari neyronli mashina tarjimasini (NMT) modellari tomonidan o'rganilgan tasvirlarni turli darajalarda tahlil qildilar va ularning sifatini tegishli tashqi xususiyatlar orqali baholashga harakat qilmoqdalar. Xususan, ular quyidagi savollarga javob izlamoqdalar: (I) Morfologik jihatdan boy tillarni tarjima qilishda muhim jihat bo'lgan ifodalar doirasida so'z tarkibi qanchalik to'g'ri saqlanadi? (II) Tarjimani aks ettiruvchi mexanizmlar uzoq masofali bog'liqliklarni qamrab oladimi va sintaktik jihatdan farq qiluvchi tillarni samarali boshqaradimi? (III) Aks ettiruvchi mexanizmlar leksik semantikani qamrab oladimi? Xususan, Massachusetts kompyuter texnologiyalari va sun'iy intellekt laboratoriyasi Instituti tadqiqotchisi Yonatan Belinkov va Jeyms Glans, Qatar kompyuter tadqiqotlari Instituti olimlari Nadir Durrani va Fahim Dalvi, Garvard Universiteti professori John F. Paulsonlar Neyron mashina tarjimasini modellarining lingvistik aks ettirish kuchi haqida tadqiqot ishlarini olib bordilar va quyidagi yo'nalishlar bo'yicha batafsil tekshiruv o'tkazdilar [Belinkov, Paulson va b., 2020]:

(I) tarjima arxitekturasining qaysi qatlamlari ushbu lingvistik hodisalarning har birini qamrab oladi?

(II) Tarjima birligini tanlash (so'z, belgi yoki pastki so'z birligi) asosiy tasvirlar tomonidan olingan lingvistik xususiyatlarga qanday ta'sir qiladi?

(III) Koder va dekoder tarjima obyektini mustaqil ravishda o'rganadimi?

(IV) Ko'p tilli NMT modellari tomonidan o'rganilgan tasvirlar ularning ikki tilli muqobillari bilan bir xil miqdordagi lingvistik ma'lumotlarni qamrab oladimi?

Ushbu tadqiqotlarning natijasi o'laroq hozirgi kunda yuzlab tarjimon dasturlar yaratilgan bo'lib, ular tarjimaning qaysidir usuliga asoslanadi. Quyida dunyoning eng yuqori reytinglarini egallab turgan onlayn tarjimon dasturlarining xususiyatlari va o'zbek-ingliz tarjima yo'nalishida qilgan tarjimalari tahliliga to'xtalib o'tamiz.

Zamonaviy tarjimon dasturlarining qiyosiy tahlili.

Google – neyron tizimlarga asoslangan mukammal onlayn tarjimon. Interfeys sodda va yuzdan ortiq tillarga tarjimani qo'llab-quvvatlaydi. Lug'atlar va mavjud funksiyalar soniga ko'ra hozirgi kunga qadar eng funksional va ko'p qirrali xizmat sanaladi. Undan maqolalar, hujjatlar, veb-sahifalarni tarjima qilishda, shuningdek,

PPT, PDF va boshqa fayl formatlarida foydalanish mumkin.

Tarjima qilingan jumla va maqolalarni Google xotirada saqlab qolish imkoni yaratilgan. Agar siz alohida so'zlarni tarjima qila boshlasangiz xizmat avtomatik ravishda onlayn lug'at rejimiga o'tadi. Har bir so'zning qisqacha tavsifi bilan muqobil variantlari taklif qilinadi. Transkripsiya va transliteratsiya qo'llab-quvvatlanadi. Shuningdek, ovoz bilan ishlash xususiyatiga ega: ovozdan matnga, matndan ovozga texnologiyasi, tarjima natijasini ovozli formatda qaytarish xususiyatlariga ega. Biroq, tarjima uchun matn kiritish hajmi 5000 belgi bilan cheklangan bo'lib, ingliz-o'zbek tarjima yo'nalishida mazmuniy tushunarsizliklar ko'p uchraydi.

Tebranib yonayotgan sham zo'r mo'jizaday hammaning diqqatini jalb qilgan edi gapi google translatorning o'zbek-ingliz yo'nalishida quyidagicha tarjima qilindi:

The flickering candle attracted everyone's attention like a miracle.

Endi ushbu tarjimani ingliz-o'zbek tarjima rejimida o'girsak, quyidagi natijaga ega bo'lamiz:

Miltillovchi sham hammaning e'tiborini mo'jizadek o'ziga tortdi.

Bu yerda *miltillamoq* so'zi *tebranib yonmoq* birikmasining aniq tarjimasini chiqarib bera olmaydi.

Yandex translator – Statistik va neyron tarjima tizimlarining aralashuvidan hosil bo'lgan gibrid tarjima modeliga asoslanuvchi yuqori sifatli onlayn tarjimon tizim. 90 dan ortiq til juftliklari orasida tarjimani amalga oshiradi va kiritish tilini avtomatik ravishda o'zi aniqlaydi. Tarjima bazasi statistik yondashuvga asosan o'z-o'zini rivojlantirib boradi. Onlayn tarjimon tizim negizida foydalanuvchi o'zining lug'at bazasini yaratishi mumkin. Dastur uning shaxsiy terminologiyasiga asoslangan holda tarjimani amalga oshiradi.

Yuqoridagi gapni yandex yordamida tarjima qilamiz:

Tebranib yonayotgan sham zo'r mo'jizaday hammaning diqqatini jalb qilgan edi.

Va quyidagi natijaga ega bo'lamiz:

The vibrating burning candle was the perfect miracle that attracted everyone's attention.

Endi ushbu gapni ingliz- o'zbek tarjima yo'nalishi bo'yicha o'giramiz. Natija quyidagicha:

Vibratsiyali yonayotgan sham barchaning e'tiborini tortgan mukammal mo'jiza edi.

Guvohi bo'lganimizdek, gap strukturasi va mazmunida ham

o'zgarishlar kuzatildi.

Prompt.One - bu neyron tarmoq yondashuvi va qoidalarga asoslangan tarjima tizimi, ya'ni Rule-Based Translation (RBMT) ni birlashtirgan yangi gibrid texnologiya bo'lib, asosan, Yevropa va Osiyo tillari uchun avtomatik tarjima xizmatlarini taqdim etadi. Ushbu mashina tarjimasi xizmati rus kompyuter tilshunosligida yaratilgan eng mashhur onlayn tarjimonlardan biridir. Xizmat yuqori sifatli tarjimani kafolatlaydigan o'ziga xos lingvistik texnologiyalariga asoslangan. PROMTning neyron algoritmlari matnni oldindan tahlil qiladi va matnning ma'lum bir qismini tarjima qilish uchun qaysi texnologiya eng mos kelishini hal qiladi. Tizim 22 ta tilni qo'llab-quvvatlaydi. Uning quyidagi afzalliklarga ega ekanligi guvohi bo'lamiz:

- so'zlarning imlosini tekshiradi;
- taklif etilayotgan tarjimani baholash imkoniyati mavjud;
- kiritish tilini avtomatik aniqlash funksiyasiga ega;
- Dasturchilarga tarjimaning o'z versiyasini yuborish

imkoniyati bor.

- matn mavzusini aniqlay olish qobiliyati mavjud;
- so'zning qo'llanish variantlarini o'z ichiga olgan lug'at

tizimiga ega mukammal tizim.

Tebranib yonayotgan sham zo'r mo'jizaday hammaning diqqatini jalb qilgan edi jumlasini dastur yordamida tarjima qilib quyidagi natijaga ega bo'lamiz:

Lightning candle attracted the attention of everyone.

Ushbu tarjimani qayta o'zbek tiliga o'girsak, gap mazmuni quyidagicha bo'ladi:

Cho'qqon shami barchaning e'tiborini o'g'irladi.

Bing Microsoft translator. Microsoft kompaniyasi 2000-yildan boshlab statistik mashina tarjimasiga asoslangan tarjimon xizmatini taqdim etib kelgan. SMT bir nechta so'zlarning kontekstini hisobga olgan holda so'zning eng yaxshi tarjimalarini baholash uchun ilg'or statistik tahlildan foydalanadi. 2010-yil boshidan beri esa sun'iy intellektning yangi texnologiyasi, chuqur neyron tarmoqlarining joriy etilishi tarjima sifatini yaxshilash imkonini berdi. Microsoft kompaniyasi neyron tarmog'ini tarjima qilish algoritmi bosqichlari quyidagicha:

- Har bir so'z, aniqrog'i uni ifodalovchi belgilar 500 o'lchovli vektor "neyronlar"ning birinchi qatlamidan o'tib, uni jumladagi boshqa so'zlar kontekstida so'zni ifodalovchi 1000 o'lchovli vektorga (b) kodlaydi.

- barcha soʻzlar ushbu 1000 oʻlchovli vektorlarga bir marta kodlangandan soʻng, jarayon bir necha marta takrorlanadi, bu esa har bir qatlamda soʻzning 1000 oʻlchovli tasvirini yaxshiroq sozlash imkonini beradi.

- Yakuniy chiqish matritsasi dasturiy taʼminot algoritmi tomonidan ishlatiladi, bu avval tarjima qilingan soʻzlarning asl jumladan qaysi soʻz keyingi tarjima qilinishi kerakligini aniqlash uchun amalga oshiriladi. Shuningdek, bu jarayonda maqsadli tildagi keraksiz soʻzlar olib tashlanadi.

- Dekoder qatlami tanlangan soʻzni (yoki aniqrogʻi, toʻliq jumla kontekstida ushbu soʻzni ifodalovchi 1000 oʻlchovli vektorni) maqsadli tilning eng mos ekvivalentiga aylantiradi. Ushbu oxirgi qatlamning (c) chiqishi, soʻngra asl jumladan qaysi soʻz keyin tarjima qilinishi kerakligini hisoblash uchun tekshiruv qatlamiga qaytariladi.

- Bu dekodeer tarjima darajasiga yetganda mos tarjima uzatiladi.

Ushbu yondashuv tufayli yakuniy natija koʻp hollarda SMT-ga asoslangan tarjimaga qaraganda ravonroq va inson tarjimasiga yaqinroq boʻladi.

Tarjima sifatini tekshirish uchun yuqoridagi gapni tarjima qilamiz:

The vibrating burning candle had attracted the attention of everyone like a great miracle.

Aynan shu gapni yana oʻzbek tiliga tarjima qilamiz:

Tebranayotgan yonayotgan sham barchaning eʼtiborini buyuk moʻjizadek oʻziga tortgan edi.

| № | Machine Translator | Tebranib yonayotgan sham zoʻr moʻjizaday hammaning diqqatini jalb qilgan edi. | |
|----|---------------------------|---|--|
| | | In English | In Uzbek |
| 1. | Google | <i>The flickering candle attracted everyone's attention like a miracle</i> | <i>Miltillovchi sham hammaning eʼtiborini moʻjizadek oʻziga tortdi</i> |
| 2. | Yandex translator | <i>The vibrating burning candle was the perfect miracle that attracted everyone's attention</i> | <i>Vibratsiyali yonayotgan sham barchaning eʼtiborini tortgan mukammal moʻjiza edi.</i> |
| 3. | Prompt. One | <i>Lightning candle attracted the attention of everyone</i> | <i>Choʻqqon shami barchaning eʼtiborini oʻgʻirladi</i> |
| 4. | Bing Microsoft Translator | <i>The vibrating burning candle had attracted the attention of everyone like a great miracle.</i> | <i>Tebranayotgan yonayotgan sham barchaning eʼtiborini buyuk moʻjizadek oʻziga tortgan edi</i> |

Koʻrib turganingizdek, gap mazmunida ham, strukturasi ham katta oʻzgarish kuzatilmadi. Demak, badiiy matnlarni tarjima qilishda oʻzbek-ingliz tarjima yoʻnalishi uchun eng mukammal tizim Microsoft Bing tarjimon dasturi sanaladi.

Xulosa

Hozirgi kunda dunyoga mashhur yana yuzlab tarjima dasturlari mavjud. Masalan, SYSTRAN, Babylon, Deeple, Translatedict, Translate.net, Reverso kabilar mukammal tarjima xizmati va foydalanuvchilarga taqdim etib kelayotgan keng imkoniyatlari bilan mashhur. Biroq, bu tizimlar tarkibi o'zbek tilini o'z ichiga olmaganligi uchun ulardan foydalanish imkonsiz. Umuman olganda, mashina tarjimasi tizimlari kundan kunga takomillashib bormoqda. Bu kundan kunga ularga bo'lgan talabning ortishi bilan belgilanadi. Tarjima jarayonida mutarjim ishini osonlashtirish bugungi mashina tarjima sohasining asosiy vazifasi sanaladi.

Mashina tarjimasi bugungi globallashuv davrida hayotning ko'plab jabhalarida zaruratga aylanib bormoqda. **Birinchi**dan, dunyodagi biznes va ilmiy matnlar miqdori tobora ortib borayotgan bu davrda ilmiy va biznes domenlarida mavjud bo'lgan kitoblar va maqolalarni o'zga tillarda nashr etish zaruriy ehtiyojga aylanib bormoqda. Xalqaro aloqalar va hamkorlik, huquqiy shartnomalar, texnik hujjatlar, ko'rsatmalar, e'lonlar, reklamalar va boshqa matnlarni tarjima qilish zarurati paydo bo'lishi tabiiy.

Ikkinchidan, chet tilidagi har bir so'z madaniyatlar chorrahasidir, bu madaniyatlararo muloqot amaliyotidir, chunki ular begona olam va begona madaniyatni aks ettiradi: har bir so'z ortida milliy ong bilan shartlangan dunyo g'oyasi turibdi. Tarjima o'zining ko'p qirraliligi bilan dunyoni to'g'ri idrok etish va madaniyatlararo muloqotda muhim ahamiyatga ega. Ammo turli mamlakatlardan kelgan mutaxassislar o'rtasidagi muloqotda hali ham til to'siqlari yoki tarjima qiyinchiliklari mavjud.

Kompyuter fanlari va sun'iy intellektdan foydalangan holda mashina tarjimasini qo'llashning bir qancha afzalliklari mavjud. Ulardan biri - mashinaning ish **tezligidir**. Xususan, jumla matni kompyuter yoki smartfonda nusxalash orqali darhol tarjima qilish mumkin. Albatta, natija yuz foiz to'g'ri bo'lmasligi mumkin, lekin aniq faktlarga asoslangan, qolipli jumalarda foydalanilsa dolzarblikni to'liq qondira oladi. Masalan, xarid qilish saytlari, ob-havo ma'lumotlari kabi. Shuningdek, sayohat, ma'muriy hujjatlar va davlat hujjatlari kabi qat'iy jumlar ko'p bo'lgan matnlarni mashina tarjimasi yordamida tezda tarjima qilish mumkin.

Ikkinchisi - **ko'p tilli muvofiqlik**. Cheklangan elektron lug'atlardan farqli o'laroq tillar soni, eng zamonaviy avtomatik tarjima ilovalari bir necha tillar bilan ishlashi mumkin. Jumladan, Google tarjimoni bir tildan ko'p tillarga tarjima qilishi va chiqish tilini

tezda o'zgartirish xususiyatiga ega. Elektron lug'atlar esa ma'lum tillar doirasidagina chegaralanib qoladi.

Uchinchi – makon va zamonda **qulaylik**. Texnika asrida foydalanuvchilar uchun eng zarur narsa ham aynan qulaylikdir. Mashina tarjimasidan har qanday vaziyatda, murakkab mashina bo'lmagan taqdirda ham smartfon yordamida osonlik bilan foydalanish mumkin. Uning tezligi va ko'p tilli interfeysga egaligi cho'ntak lug'atlardan samaraliroq ekanligini ko'rsatadi. Shuningdek, zamonaviy mashina tarjimon dasturlarining ko'pchiligi nutqni aniqlash va matnga aylantirish tizimiga egaligi har qanday makon va zamonda foydalanish imkonini beradi.

Mashina tarjimasi kompyuter tilshunosligining eng muhim yo'nalishlaridan biri bo'lib, barcha til darajalarida nutqni qayta ishlashning barcha muammolarini o'z ichiga oladi.

Foydalanilgan adabiyotlar

- Baker K., Bloodgood M., Dorr B. J., Burch Ch., Filardo N. W., Piatko Ch., Levin L., Miller S. 2012. Use of Modality and Negation in Semantically-Informed Syntactic MT // Johns Hopkins University, Moulton Street, Cambridge.
- Belinkov Y., Paulson J., Durrani N., Dalvi F., Sajjad H. Glans J. 2020. On the Linguistic Representational Power of Neural Machine Translation Models // Qatar Computing Research Institute HBKU Research Complex. – Doha.
- Bisazza A., Federico M. 2016. A Survey of Word Reordering in Statistical Machine Translation: Computational Models and Language Phenomena, Amsterdam.
- Deng D., Xue N. 2017. Translation Divergences in Chinese-English Machine Translation: An Empirical Investigation // Department of Chinese Languages and Literature, Tsinghua University, Beijing, China.
- Durrani N., Schmid H., Fraser A., Koehn P., Schutze H. 2015. The Operation Sequence Model — Combining N-Gram-Based and Phrase-Based Statistical Machine Translation // Ludwig Maximilian University – Munich.
- Khullar P. 2021. Are Ellipses Important for Machine Translation? // India, Hyderabad.
- Eriguchi A., Hashimoto K., Tsuruoka Y. 2019. Incorporating Source-Side Phrase Structures into Neural Machine Translation // The University of Tokyo Department of Information and Communication Engineering – Tokyo.

- Fraser A., Marcu D. 2007. Measuring Word Alignment Quality for Statistical Machine Translation. University of Southern California. USA.
- Gimpel K., Smith N. A. 2014. Phrase Dependency Machine Translation with Quasi-Synchronous Tree-to-Tree Features // Carnegie Mellon University. – Germany.
- Goutte C., Cancedda N., Dymetman M., Foster G. 2009. Learning Machine Translation Institute for Information Technology, National Research Council// MA: The MIT Press, Cambridge.
- Haddow B., Bawden R., Miceli A. V., Birch A. 2022. Survey of Low-Resource Machine Translation // University of Edinburgh School of Informatics - Edinburgh.
- Joty Sh., Guzman F. Marquez L., Nakov P. 2017. Discourse Structure in Machine Translation Evaluation // School of Computer Science and Engineering Nanyang Technological University - Qatar Foundation.
- Koehn P. Statistical Machine Translation (book). Cambridge University Press// Edinburg, 2010
- Lembersky G., Ordan N. Wintner S. Improving Statistical Machine Translation by Adapting Translation Models to Translationese // University of Haifa, Israel, 2013
- Martinez D.O. Online Learning for Statistical Machine Translation// Universitat Politecnica de Valencia, 2016.
- Muhammad S. M. Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis. National Research Council Canada, 2022.
- Munteanu D.S., Marcu D. Improving Machine Translation Performance by Exploiting Non-Parallel Corpora. Information Sciences Institute University of Southern California. - USA, 2006.
- Neubig G., Watanabe T. Optimization for Statistical Machine Translation: A Survey // Tokyo, Japan, 2016.
- Nießen S., Ney H. Statistical Machine Translation with Scarce Resources Using Morpho-syntactic Information // Computer Science Department, RWTH Aachen, Germany, 2004.
- Och F. J. Ney H. 2004. The Alignment Template Approach to Statistical Machine Translation. Computer Science Department, RWTH Aachen–University of Technology, Ahornstr. 55, 52056 Aachen, Germany.
- Popovic M., Ney H. 2011. Towards Automatic Error Analysis of Machine Translation Output// RWTH Aachen University – German Research Centre for Artificial Intelligence, Alt-

Moabit Berlin, Germany.

- Riezler S., Yi Liu 2010. Query Rewriting Using Monolingual Statistical Machine Translation. Zurich, Switzerland.
- Shen L., Xu J. Weischedel R. 2010. String-to-Dependency Statistical Machine Translation Raytheon BBN Technologies. Moulton Street, Cambridge.
- Specia L., Scarton K., Paetzold G.H. 2017. Quality Estimation for Machine Translation University of Sheffield and Federal University of Technology, Parana.
- Tillmann Ch., Ney H. 2003. Word Reordering and a Dynamic Programming Beam Search Algorithm for Statistical Machine Translation. Computer Science Department, RWTH Aachen, Germany.
- Ueffing N., Ney H. 2007. Word-Level Confidence Estimation for Machine Translation RWTH Aachen University, Germany.
- Wang P., Nakov P., TouNg H. 2016. Source Language Adaptation Approaches for Resource-Poor Machine Translation // National University of Singapore.
- Wang W., May J., Knight K., Marcu D. 2010. Re-structuring, Re-labeling, and Re-aligning for Syntax-Based Machine Translation Los Angeles, USA.
- Wan Y., Yang B., Wong D. F., Chao L. S., Zhang H. 2021. On the Linguistic Representational Power of Neural Machine Translation Models // China: Alibaba DAMO Academy.

THE PROCESS OF MACHINE TRANSLATION IN THE MODERN WORLD: ANALYSIS AND RESULTS

**Madinabonu Qodirova¹,
supervisor: Shahlo Hamroyeva²**

Abstract. Today, it is no secret that the field of machine translation is rapidly developing. In this regard, it is impossible not to dwell on the researches and theories being created in world science. We believe that studying these ideas and, if necessary, applying them to the field of Uzbek machine translation will stimulate the rapid development of this field. In this article, we will reflect on the work carried out in the world of machine translation today, the theories and ideas proposed for the optimization of automatic translation. We compare and analyze them. Also, we will touch on the working principles of currently popular translator programs and get results.

Key words: *automatic translation, RBMT, SMT, neural systems, NMT, beam search system, dynamic approach, reconstruction.*

References

- Baker K., Bloodgood M., Dorr B. J., Burch Ch., Filardo N. W., Piatko Ch., Levin L., Miller S. 2012. Use of Modality and Negation in Semantically-Informed Syntactic MT // Johns Hopkins University, Moulton Street, Cambridge.
- Belinkov Y., Paulson J., Durrani N., Dalvi F., Sajjad H. Glans J. 2020. On the Linguistic Representational Power of Neural Machine Translation Models // Qatar Computing Research Institute HBKU Research Complex. – Doha.
- Bisazza A., Federico M. 2016. A Survey of Word Reordering in Statistical Machine Translation: Computational Models and Language Phenomena, Amsterdam.
- Deng D., Xue N. 2017. Translation Divergences in Chinese–English Machine Translation: An Empirical Investigation // Department of Chinese Languages and Literature, Tsinghua University, Beijing, China.

¹*Qodirova Madinabonu Murodjon qizi* – PhD student of Tashkent State University of Uzbek Language and Literature named after Alisher Navo'i.

E-pochta: begonammm@gmail.com

ORCID: 0009-0005-8715-8074

²*Hamroyeva Shahlo Mirdjonovna* – doctor of philological sciences, associate professor, etc. Alisher Navo'i Tashkent State University of Uzbek Language and Literature.

E-pochta: shaxlo.xamrayeva@navoiy-uni.uz

ORCID: 0000-0002-5429-4708

- Durrani N., Schmid H., Fraser A., Koehn P., Schutze H. 2015. The Operation Sequence Model — Combining N-Gram-Based and Phrase-Based Statistical Machine Translation // Ludwig Maximilian University – Munich.
- Khullar P. 2021. Are Ellipses Important for Machine Translation? // India, Hyderabad.
- Eriguchi A., Hashimoto K., Tsuruoka Y. 2019. Incorporating Source-Side Phrase Structures into Neural Machine Translation // The University of Tokyo Department of Information and Communication Engineering – Tokyo.
- Fraser A., Marcu D. 2007. Measuring Word Alignment Quality for Statistical Machine Translation. University of Southern California. USA.
- Gimpel K., Smith N. A. 2014. Phrase Dependency Machine Translation with Quasi-Synchronous Tree-to-Tree Features // Carnegie Mellon University. – Germany.
- Goutte C., Cancedda N., Dymetman M., Foster G. 2009. Learning Machine Translation Institute for Information Technology, National Research Council// MA: The MIT Press, Cambridge.
- Haddow B., Bawden R., Miceli A. V., Birch A. 2022. Survey of Low-Resource Machine Translation // University of Edinburgh School of Informatics - Edinburgh.
- Joty Sh., Guzman F. Marquez L., Nakov P. 2017. Discourse Structure in Machine Translation Evaluation // School of Computer Science and Engineering Nanyang Technological University - Qatar Foundation.
- Koehn P. Statistical Machine Translation (book). Cambridge University Press// Edinburg, 2010
- Lembersky G., Ordan N. Wintner S. Improving Statistical Machine Translation by Adapting Translation Models to Translationese // University of Haifa, Israel, 2013
- Martinez D.O. Online Learning for Statistical Machine Translation// Universitat Politecnica de Valencia, 2016.
- Muhammad S. M. Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis. National Research Council Canada, 2022.
- Munteanu D.S., Marcu D. Improving Machine Translation Performance by Exploiting Non-Parallel Corpora. Information Sciences Institute University of Southern California. - USA, 2006.
- Neubig G., Watanabe T. Optimization for Statistical Machine Translation: A Survey // Tokyo, Japan, 2016.

- Nießen S., Ney H. Statistical Machine Translation with Scarce Resources Using Morpho-syntactic Information // Computer Science Department, RWTH Aachen, Germany, 2004.
- Och F.J. Ney H. 2004. The Alignment Template Approach to Statistical Machine Translation. Computer Science Department, RWTH Aachen–University of Technology, Ahornstr. 55, 52056 Aachen, Germany.
- Popovic M., Ney H. 2011. Towards Automatic Error Analysis of Machine Translation Output// RWTH Aachen University – German Research Centre for Artificial Intelligence, Alt-Moabit Berlin, Germany.
- Riezler S., Yi Liu 2010. Query Rewriting Using Monolingual Statistical Machine Translation. Zurich, Switzerland.
- Shen L., Xu J. Weischedel R. 2010. String-to-Dependency Statistical Machine Translation Raytheon BBN Technologies. Moulton Street, Cambridge.
- Specia L., Scarton K., Paetzold G.H. 2017. Quality Estimation for Machine Translation University of Sheffield and Federal University of Technology, Parana.
- Tillmann Ch., Ney H. 2003. Word Reordering and a Dynamic Programming Beam Search Algorithm for Statistical Machine Translation. Computer Science Department, RWTH Aachen, Germany.
- Ueffing N., Ney H. 2007. Word-Level Confidence Estimation for Machine Translation RWTH Aachen University, Germany.
- Wang P., Nakov P., TouNg H. 2016. Source Language Adaptation Approaches for Resource-Poor Machine Translation // National University of Singapore.
- Wang W., May J., Knight K., Marcu D. 2010. Re-structuring, Re-labeling, and Re-aligning for Syntax-Based Machine Translation Los Angeles, USA.
- Wan Y., Yang B., Wong D. F., Chao L. S., Zhang H. 2021. On the Linguistic Representational Power of Neural Machine Translation Models // China: Alibaba DAMO Academy.