

ÕIGUSKEEL = KEEL + ÕIGUS. Teekond sõnast õigusaktini ja õigusaktist inimeseni uute keeletehnoloogiate ajastul

1. osa – Euroopa keel on tõlkimine (*Umberto Eco*). Uued keeletehnoloogiad ja õiguslikud väljakutsed

Väitlusjuht: *PhD* Heiki Pisuke

Euroopa Komisjoni kirjaliku tõlke peadirektoraadi (DGT) eesti keele osakonna juhataja, Tartu Ülikooli külalisprofessor

MPhil Merit-Ene Ilja

DGT direktor, endine Eesti Õigustõlke keskuse direktor

Tänapäeval on tehniline ja õigustõlge mõeldamatud ilma keeletehnoloogia toeta. Juba aastaid on tõlketekstide taaskasutus tõlkemälu abil olnud tõhusa kvaliteettõlke alustala. Ka masintõlge on teatud keelesuundades juba mõnda aega kasutuses. Keeletehnoloogia kiire areng viimasel ajal on kaasa toonud kvalitatiivse hüppe masintõlkes. Neuromasintõlge ehk närvivõrkudel põhinev masintõlge (NMT) on asendamas statistilist masintõlget ja pakub juba head tulemust ka keeltele, mille puhul statistilised süsteemid ei töötanud.

Euroopa Komisjoni kirjaliku tõlke peadirektoraat teeb esimesi NMT katsetusi inglise-eesti õigustõlkes, st ELi õigusaktide tõlkimisel eesti keelde. Kas 5–10 aasta perspektiivis tõlgib õigusakte ainult masin ja inimese roll tõlkeprotsessis keskendub pigem kvaliteedikontrollile ja andmehooldusele? Praeguse teadmise alusel võib öelda, et mida standardsem on õigusakti keel, seda kvaliteetsem on masintõlke väljund. Missugused võiksid olla automaatse taaskasutuse võimalused õigusloomes?

Madis Vunder

Euroopa Kohtu mitmekeelsuse peadirektoraadi direktor

Ajalugu on näidanud, et lisaks igapäevase kõnekeele olemasolule on ühe keele säilimiseks pikemas perspektiivis hädavajalik ka erialase oskuskeele olemasolu ja areng. Juriidiline erialakeel kujuneb õigusloome ja õiguse kohaldamise (sh kohtumenetluse) käigus.

Kuivõrd ELi õigus on siseriikliku õiguse osa, kujuneb osa juriidilisest eesti keelest ELi õigusloome- või kohtumenetluses, praktikas eesti keelde tõlkimise käigus. Eesti juristkonna ning laiemalt võttes kogu eesti keelekogukonna huvides on, et see osa ei mõjaks mitte võõrapärase ja kunstliku, vaid omasena ning vastaks juba loomise käigus õigusselguse (arusaadavuse, lihtsuse ja täpsuse) nõuetele.

Keelemõjutused on vastastikused ja mitmekülgsed (Tallinnas ja Tartus loodav õiguskeel mõjutab pidevalt Brüsseli ja Luksemburgi oma ja vastupidi), oluline on nende mõjutuste teadvustamine ning ühine keelehooldus.

Ann Stolfot

Euroopa Liidu Nõukogu õigusteenistuse õiguskeele ekspert

Kuigi ELi õiguses kehtib keelte võrdsuse põhimõte, koostab Euroopa Komisjon õigusakti ettepaneku ühes keeles, milles Euroopa Liidu Nõukogu seda tavaliselt ka menetleb. Viimase kümne-viieteistkümnepäevase aasta jooksul on rõhk järjest enam nihkunud inglise keelele. Samas tuleb tagada, et lõpptulemuseks olev õigusakt oleks kõigis keeltes sama õigusvõimega ja võrdselt kohaldatav. Selleks et tõlked erinevates keeltes oleksid samaväärsed, peab algtekst olema selge, arusaadav ja üheselt mõistetav. ELi õigusakti tõlge peab keeleliselt sobituma ka selle riigi õigusesse, kus seda kohaldatakse. Siinkohal on asendamatu liikmesriikide ekspertide abi, kes tunnevad asjaomase valdkonna terminoloogiat ja tegelevad õiguse praktilise poolega. Terminite puhul tuleb aga jälgida, et sisu oleks sama, mis liidu õiguses ette nähtud.

Mida teha siis, kui Euroopa Liidu Teatajas avaldatud õigusakti mõnes keeleversioonis või kõigis keeleversioonides on viga? Kui see takistab õigusakti kohaldamist, algatatakse parandusmenetlus. Seadusandliku tavamenetluse raames ja nõukogu poolt vastu võetud õigusaktide puhul algatavad parandusmenetluse nõukogu õiguskeele eksperdid.

Dr.iur. Kadri Siibak

Rahandusministeeriumi finantsturgude poliitika osakonna peaspetsialist

Keel on juristi töö peamine instrument. Õigusakti tõlkes kasutatud terminite erinevast sisustamisest tulenevad probleemid võivad olla sügavamad, kui esmapilgul võiks arvata (finantsvaldkonna karistusõiguse näitel).

Eesti õiguskeelt peavad ELi õigusnormide tõlkijad ja valdkonna eksperdid arendama koos, tihe koostöö on väga oluline. Tulenevalt ühiskonnas asetleidvatest kriisidest või kiirest arengust, võetakse ELis spetsiifilistes valdkondades kasutusele täiesti uusi õiguslikke lahendusi ja mõisteid. Näiteks finantsvaldkonnas esineb ELi õigusakti originaaltekstis küllaltki sageli termineid, millele Eestis sisuline vaste puudub ja tuleb luua uusi eestikeelseid termineid. Vahel on uute terminite teemal ette tulnud päris teravaid diskussioone ja erimeelsusi.

Eesti ametnikud peaksid riigisisese õiguse eelnõude kirjutamisel direktiivide tõlgetesse kriitiliselt suhtuma. Kasuks tuleb eri keeleversioonide võrdlemine.

PhD Mark Fišel

Tartu Ülikooli arvutiteaduse instituudi keeletehnoloogia dotsent

Automaattõlke süsteemid kindlasti ei asenda lähitulevikus inimtõlkijaid. Juriidiliste tekstide puhul ei saa anda garantiid, et automaatselt tõlgitud teksti sisu on sama mis originaalis, või et tõlke kvaliteet on üle seatud lävendi, või isegi seda, et kogu lähteteksti sisu on tõlgitud. Samas aga teevad masintõlke süsteemid inimtõlkijate töö oluliselt kiiremaks, seda on eriti näha tehniliste ja ka juriidiliste tekstide tõlkimisel.

PhD Arvi Tavast

qlaara asutaja

Masintõlge on üks keeletehnoloogia tuntumatest rakendustest. Selle kõrval on massidesse jõudmas ka inimkeelsed kasutajaliidesed ehk võimalus suhelda mitmesuguste seadmetega loomulikus kirjalikus või kõneldud keeles (Siri, Alexa jne). Keeletehnoloogia kasutusvõimalusi on veel palju, näiteks võiks mainida selle rakendusi kohtuekspertiisis: hääle või isegi kirjaliku teksti järgi on võimalik isikut tuvastada, telefonikõnede salvestusi on võimalik automaatselt transkribeerida ja otsitavaks muuta, suhtlusvõrgustikes tuvastada vihakõnet või terrorismile viitavaid tekstilõike jne. Veel üks lihtne, kuid praktiline rakendus on riigikogu ja kohtuistungite stenogrammide automaatne koostamine.

Selliste lahenduste plahvatuslik levik viimasel ajal on võimalikuks osutunud tänu statistiliste meetodite ja masinõppe arengule. Traditsioonilise programmeerimise jaoks, kus arvutile kirjutati käsitsi üksikasjalikke juhiseid igas olukorras käitumiseks, ei olnud inimkeel oma keerukuses jõukohane. Tänapäeval inimkeelset suhtlust vahetult programmeerida enam ei üritata. Selle asemel programmeeritakse masin õppima, misjärel ta saab keele ära õppida põhimõtteliselt samamoodi nagu laps - teiste kõnet kuulates ja jäljendades. Isegi meetodid on teatud määral sarnased, nimelt on närvivõrkude kui praegu populaarseimate masinõppealgoritmide inspiratsioon algselt pärit aju neuronite tööpõhimõttest.

Üleminek masinõppele on aga toonud kaasa muutuse ressurrssides, mida rakenduste loomiseks vaja läheb. Nagu laps, vajab ka masinõppesüsteem rääkimaõppimiseks palju keelelist sisendit, kus "palju" tähendab tänapäeval miljardeid sõnu ehk rohkem kui üks inimene oma eluea jooksul lugeda ja kuulda jõuab. Konkreetselt eesti keeletehnoloogia arendamiseks on seega tarvis palju eestikeelset teksti, mille peal süsteeme treenida. Eri rakenduste jaoks läheb vaja eriliigilisi tekste, sh kirjalikke ja suulisi, toimetatud ja spontaanseid, üldisi ja erialaseid, tõlkeid ja originaale jne. Mõne masinõppetehnoloogia (juhendatud õppimise) korral tuleb treeningmaterjal või osa sellest ka eelnevalt märgendada ehk varustada inimese poolt "õigete vastustega", mida masin siis saaks jäljendama hakata.

Keeleressurrsside kogumisel on mitmeid puhttehnilisi piiranguid: aega ja tööd võtab tekstide leidmine või hankimine, valimine, puhastamine, sobivasse vormingusse teisendamine jne, märgendamisest rääkimata. See on loomulik ja vältimatu. Lisaks

aga tuleneb praegusest autoriõiguse ja isikuandmete kaitse regulatsioonist juriidilisi takistusi, mille otstarbekus ei ole päriselt väljaspool kahtlust ja mis võivad olla vastuolus ka eesti keele põhiseadusliku kaitse vajadusega.

PhD Aleksei Kelli

Tartu Ülikooli intellektuaalse omandi õiguse professor, CLARIN ERIC õiguskomitee esimees

Keeleressursside loomise ja keeletehnoloogia arendamise üldisemaks eesmärgiks on eesti keele digitaliseerimine. Digitaliseerimine aitab kaasa eesti keele säilimisele, mille vajadust rõhutab ka põhiseadus. Teemale lisab olulisust asjaolu, et eesti keele majanduslik potentsiaal ei ole võrreldav suurte keelte – näiteks inglise keele – omaga.

Keeletehnoloogia arendamine tugineb keeleressurssidele, mis on eesti keele digitaliseerimise vundament. Keeleressursid on andmebaas, mis võib sisaldada autoriõiguslikult kaitstud teoseid, kaasnevate õiguste objekte (nagu esitused) ja isikuandmeid. Nende loomine saab tugineda isiku nõusolekule või seadusest tulenevatele eranditele, mis lubavad kasutada intellektuaalse omandi objekte ja töödelda isikuandmeid ilma nõusolekuta. Üldjuhul on selleks teadustöö erand. Õiguslikud probleemid võivad tekkida asjaolust, et keeleressursside loomise õiguslik alus võib mõjutada nende kasutust. Tekib küsimus, kas teadustöö erandi alusel loodud keeleressursse võib kasutada äriliste rakenduste loomiseks ning mida üldse mõista ärilise rakenduse all?

Suulise kõne kasutamine keeletehnoloogilisteks arendusteks on õiguslikult üks komplekssemaid küsimusi. Isiku hääl on käsitatav isikuandmete osana. See, mis isik räägib, võib olla autoriõiguslik teos ning lisaks eksisteerivad ka esitaja õigused.

Kvaliteetsete keeletehnoloogiliste rakenduste loomine eeldab kvaliteetset sisendit (keeleressursse). Esindatud peab olema võimalikult mitmekesine keel. See tähendab, et peab salvestama nii laste kui ka vanurite kõnet, kõnedefekti ja vaimse puudega isikute kõnet. See kõik tekitab täiendavaid õiguslikke probleeme.

Eksisteerib suur keeleandmestik, mis on kogutud nõukogude perioodil (rahvaluule ja -pärimuse salvestised jms). Seda saaks kasutada keeleressursside loomiseks ja keeletehnoloogia arendamiseks, kuid selle kogumisel ei ole järgitud tänapäeval kehtivaid autoriõiguse ja isikuandmete kaitse nõudeid.