

Tõlketööde hankimise parimad praktilad ning põhimõtted

juuni 2022

Sisukord

Terminoloogia	3
Juhendmaterjalist	4
Tõlkimisest üldiselt.....	5
Tõlke tellimine.....	6
Tõlkemälude haldamine	7
Tõlkeabiprogrammide kasutamine	8
Tõlkevaliteedi tagamine	9
Soovitused hangete korralduse kohta	10
Keeleabi viited.....	11

Terminoloogia

Alljärgnevalt on välja toodud olulisemad valdkonna terminid.

Stiiljuhend – stiiljuhendis kirjeldatakse eelistatud stiili ja reegleid, millest tõlkimisel tuleks juhinduda. See aitab ühtlustada asutuse dokumentide keelekasutust.

Terminibaas – terminibaas on terminite andmebaas, enamasti mitmekeelne. Terminibaas pakub välja oskussõna tähenduse, definitsiooni, kasutusnäiteid, kasutusvaldkonna, allikad ja tõlked teistesse keeltesse (mitmekeelse terminibaasi puhul).

Tõlkeabiprogramm (*CAT-tool* ehk *computer assisted translation*) – tõlkeabiprogramm on tarkvara, mis hõlbustab tõlkimise protsessi. Näiteks võimaldab salvestada tõlkimise käigus lähtek testi kui ka vastava tõlgitud teksti andmebaasi ning võimaldab edaspidi seeläbi teksti automaatset tõlkimist või tõlkimisest kõrvale jäämist. Lisaks võimaldavad tõlkeabiprogrammid ühtlustada terminoloogia kasutamist, luua terminisõnastike, parandada grammatikat ja saada ülevaadet tõlgitavast tekstist.

Tõlkemälu (*Translation memory*) – andmebaas, kus on talletatud nii lähtetekst kui ka sellele vastav tõlgitud tekst. Tõlkemälu kasutatakse koos tõlkeabiprogrammiga ja seda on võimalik tõlkeabiprogrammist eraldada (eksportida) ning mõnda teise tõlkeabiprogrammi paigaldada (importida). Tõlkemälu aitab hankival asutusel tulevikus vähendada tõlkekulusid ja panustada täpsema masintõlke arengusse. Tõlkemälu hõlbustab mälus salvestatud lausetega sarnaste lausete tõlkimist ja tõlgete ühtluse tagamist. Tõlkemälu võimaldab tõlkijal sõnu või fraase otsides teada saada, kuidas neid varem tõlgitud on.

Juhendmaterjalist

Avalike ülesannete täitmisel tõlgitakse pidevalt tekste, mis korduvad dokumendist dokumenti ja kasutatakse samatähenduslikke mõisteid ning termineid. Samas puuduvad keskne tõlkemälu ja automaattõlke vahendid, mis võimaldaksid oluliselt vähendada tõlkimisele kuluvat aega ja raha. 2016. aastal hindas Riigi Tugiteenuste Keskus, et tõlkemälu kasutamisega, mis eeldab ka kogumist, kaasnev rahaline kokkuhoid jääb 20% piiridesse. Samas tõdeti, et arvutus on hinnanguline ja vajaks põhjalikumat analüüsi. Sellegipoolest suureneb tõlkemälu mahu kasvades ja selle pikaajalisel kasutamisel ka korduvtõlgete osakaal.

Kahjuks puudub täna asutustes tõlkemälu kasutamise võimalus, kuna hankepartnerid ei anna enamikel juhtudel loodud keeleressursse tellijale üle. Asutustes, kus tõlkemälu küsitakse või ise luuakse, puudub ühtne tõlkemälu, kuna töötatakse eraldi. See tingib olukorra, kus korduva teksti tõlkimise eest makstakse uuesti, kasutatakse erinevaid termineid, tõlked pole ühtlase kvaliteediga ja tõlkimisprotsess võtab kaua aega. Teisalt ei osata alati tõlkehankeid läbi viia, pakkumisi hinnata ja hiljem tehtud tööd tagasisidestada. Kõik see tingib olukorra, kus tõlkimiseks kulub palju aega, tõlgete kvaliteet ei vasta ootustele ning eestikeelse keeletehnoloogia areng on aeglane. Käesoleva juhendi eesmärgiks on anda asutustele soovitusi, kuidas tõlkehankeid läbi viia ning panna paika põhimõtted, mida kõik avaliku sektori asutused järgima peaksid.

Antud juhend on loodud koostöös erinevate osapooltega ning sellesse saavad kõik huvipooled ka tulevikus panustada! Läbi selle tagame, et valdkonna arenedes saame omavahel uusi kogemusi ja teadmisi paremini jagada. Käesolevate põhimõtete ja soovituste osas on suure panuse andnud Haridus- ja Teadusministeerium, Eesti Tõlkemagistrite Liit, Luisa Tõlkebüroo, Tõlkebüroo Interlex, Avatar Tõlkebüroo ja Neuma Tõlkebüroo. Suur aitäh panuse eest Kadri Varele ja Kristiina Vilbastele.

Ott Velsberg

Riigi andmete juht

Tõlkimisest üldiselt

Tõlkimine on lähteteksti sisu, vormi ja stiili võimalikult täpne edasiandmine sihtkeeles. Samal ajal ei tohi tõlgitud tekst mõjuda tõlkena, vaid loomuliku sihtkeelse tekstina, st ei tõlgita mitte sõnu, vaid mõtet, andes selle edasi sihtkeelele omasel viisil. Hea tava kohaselt peaks tõlkega töötama alati kaks inimest: tõlkija, kes tõlgib, ja toimetaja, kes korrigeerib teksti keeleliselt ja võimalusel ka sisuliselt. Toimetaja peaks olema sihtkeelt emakeelena kõnelev isik.

Tõlkemahtu arvestatakse lähteteksti sõnade või lehekülgede arvu järgi. Keskmine tõlkimiskiirus on 1 lehekülge ehk 250 sõna tunnis, st umbes 8 lehekülge ehk 2000 sõna päevas. Kiirtõlgete puhul võiks arvestada kuni 500 sõna päevas. Täpne päevane tõlkemaht sõltub teksti olemusest ja keerukusest.

Fraaside ja lausete eest, mille vasted on osalise kattuvusega (*fuzzy matches*, st mõningad sõnad erinevad), tuleks maksta vähem. Tõlkemälu kasutamisel võiks hinnasoodustuse arutamisel kasutada järgnevat mudelit:

100% kattuvad vasted, sh täielikud vasted, kordused: tasu täishinnast 20%

95%-99% kattuvad vasted: tasu täishinnast 30%

85%-94% kattuvad vasted: tasu täishinnast 50%

75%-84% kattuvad vasted: tasu täishinnast 60%

50%-74% kattuvad vasted: tasu täishinnast 100%

100% kattuvad vasted peaksid olema tasustatud, sest tõlkija/toimetaja peab need siiski üle kontrollima. Tõlkemällu võib kõigest hoolimata vigu tekkida/jääda ja terminoloogia võib aja jooksul muutuda. Peale selle võib sama sõna/fraas nõuda sõltuvalt kontekstist eri kohtades erinevat tõlget. Kui tellija ei soovi 100% kattuvate vastetega teksti eest maksta, ei vastuta tõlkija 100% kattuvate vastetega tõlke õigsuse eest. Tõlkemälust tuleneva kasu arvestamiseks tuleks võrrelda tehtud pakkumisi, arvestades kokkulangevustega ja nende hinnastamisega.

Teksti korrigeerimine on tõlke ülelugemine teise isiku poolt ning õigekirjavigade parandamine. Keskmine korrigeerimiskiirus on 36 lehekülge ehk 9000 sõna päevas. Teksti toimetamisel arvestada keskmiselt 4000 sõna päevas.

Teksti toimetamine on teksti õigekirja ja stiili parandamine ning võimalusel tõlke sisu võrdlemine lähtetekstiga. Keskmine toimetamiskiirus on 24 lehekülge ehk 6000 sõna päevas.

Tõlke tellimine

Kirjalikule ja suulisele tõlkele on soovitus korraldada eraldi hanked. **Tooge tõlke hankimisel välja järgnev: tõlke eesmärk ja kontekst, keelesuunad, maht (lehekülgede ja sõnade arv), kuidas haldate tõlkemälu ja kes on selle omanik. Stiiljuhise ja terminibaasi olemasolul tuleks ka need välja tuua.**

Tõlketeenuse pakkujale tuleks anda järgmine taustinfo: mis tekstiga on tegu ning miks ja kelle jaoks seda tõlgitakse. Tõlke eesmärk ja kontekst aitavad tõlkijal paremini mõista, kus ja millisel eesmärgil tõlget kasutatakse. Näiteks: Mis: krattidealase info levitamiseks loodud turundusmaterjal. Eesmärk: mõeldud kodanikele, kellel puudub valdkonnaga varasem kokkupuude. Lisainfo: tõlgitavat materjali ei pea rangelt järgima, see tähendab sõna-sõnalt tõlkimine pole vajalik. Tähtis on lihtsasti mõistetav tekst, mis ei tundu lugejale tõlkena.

Koos tõlgitava materjaliga tuleks kaasa anda abimaterjalid: näiteks tõlkemälu, kui see on samalaadsete tekstide kohta olemas, sõnastik, terminibaas, seotud või sarnased tekstid ja nende tõlked (näiteks samadel või sarnastel teemadel või potentsiaalselt osaliselt kattuvad), stiiljuhend, kontaktisik(ud), kelle poole tõlketeenuse pakkuja saaks terminite kooskõlastamiseks või muude teksti sisu puudutavate küsimuste selgitamiseks pöörduda, ja muud materjalid, millest võib tõlkijale abi olla. Tõlkijale oleks lisaks veel kasulik välja tuua, millal eelistatakse väiketähtede kasutamist, kuidas peaks tõlkima ametiasutuste nimesid, kas tuleks kasutada ametlikku kirjaviisi, kuidas tehnilisi termineid tõlkida ning kellelt või kust asjakohast nõu saab.

Samuti tuleks teada anda võimalikest eelistustest tõlkimisel, näiteks kas ja kuidas tõlkida lühendeid või kuidas tõlkida teatud termineid või millist tõlkeabiprogrammi tuleb tõlkimisel kasutada.

Tõlkimise ajal tuleks tõlketeenuse pakkujale vajaduse korral tuge pakkuda, näiteks kui tekib küsimusi terminoloogia, teksti sisu vms kohta.

Pärast tõlke üleandmist tuleks tõlketeenuse pakkujale anda tagasisidet tõlke kohta. Kui tõlkes tehakse muudatusi, tuleks ka need tõlketeenuse pakkujale edastada, et ta saaks parandused vajaduse korral tõlkemällu sisse viia ja nendega tulevastes töödes arvestada.

Tõlkemälude haldamine

Tõlketeenuse tellijal oleks hea luua asutusesisene terminibaas ja kasutada tõlgete tegemiseks alati tõlkemälu. **Tõlkemälu (ja terminibaasi) on oluline hooldada, st viia tõlgetes tehtud parandused alati ka tõlkemällu.** Terminibaasi loomise korral tuleks kokku leppida, kellel on õigus terminibaasis muudatusi teha juhul, kui terminoloogia muutub või avastatakse vigu. **Hankijal on alati soovitatav nõuda endale tõlkemälu. See tähendab, et tõlkemälu omanik peaks olema tõlget hankiv asutus. Alustada võiks valdkonnapõhiste tõlkemälude koostamisest ja kasutamisest.**

Tõlkemälu on väga kasulik töövahend, kuid seda ainult siis, kui see on hooldatud ja usaldusväärne. Tõlkemälu kasutamise eesmärk on vähendada tõlkekulusid, sest korduva teksti ja osaliste vastete eest tuleb vähem maksta.

Tõlkemälu tuleb ülevõtmisel kontrollida, et oleks tagatud tõlkemälu kvaliteet. Selleks oleks hea asutustes või osakonna tasemel määrata inimene, kes jooksvalt tegeleb tõlgete ja tõlkemälude haldamisega. Kui tõlkes viiakse sisse parandusi ja muudatusi, siis tuleks need sisse viia ka tõlkemälusse. Muidu võivad sarnased vead tulla sisse ka tulevastes tõlgetes. **Tõlkemälu küsimisel nõudke avatud ja platvormist sõltumatut vormingut.**

Kui asutuses ei ole kasutusel tõlkeabiprogramme, kuid on olemas tõlked ja nende lähtetekstid, on võimalik nende tekstide alusel luua tõlkemälu (seda nimetatakse tekstide joondamiseks). Kui seda teenust soovitakse tõlketeenuse pakkujalt, tuleb arvestada

lisatasuga, sest tekstide joondamine tõlkemälu saamiseks võtab samuti aega. Kui on valmisolek seda ise luua, siis on selleks võimalik kasutada tõlkeabitööriistad, aga ka vabavaralisi veebipõhiseid lahendusi, näiteks Matecati „Alignerit¹“.

Tõlkemälude tasuta loomist pakub ka ELRC², aidates luua kõigile vabalt kasutatavad paralleelcorpuseid. ELRC teenus sobib seega kõigile asutustele, kelle dokumendid pole mõeldud ainult sisemiseks kasutamiseks.

Tõlkemälu tuleks avalikustada avaandmete teabevärravas (avaandmed.eesti.ee), kuna see aitab kaasa eestikeelse keeletehnoloogia arengule ja aitab vähendada avalikus sektoris tõlgetele kuluvat aega ja raha

Tõlkeabiprogrammide kasutamine

Tõlkeabiprogrammid salvestavad tõlgitud teksti tõlkemällu, võimaldades tõlkijal lihtsamini leida varem tõlgitud termineid ja fraase ning tagada korduva teksti ühesugune tõlkimine. Tõlkeabiprogramm võimaldab seega sama või piisavalt sarnast lauset automaatselt tõlkida kui see järgmine kord tekstis esineb. Tõlkija võib kas pakutud tõlke olemasoleval kujul kasutusele võtta, seda muuta või kõrvale jätta. Lisaks on tõlkeabiprogrammidega võimalik tõlkemälusid hallata – mälusid omavahel liita, tõlkevasteid muuta jne.

Kui tõlketeenuse tellija kasutab asutusesiselt tõlkeabiprogramme, oleks hea luua asutusesisene terminibaas ja kasutada tõlgete tegemiseks alati tõlkemälu.

Tõlkeabiprogramme on turul väga palju, mõned neist on omavahel ühilduvad, mõned mitte. Tõlkemälusid on võimalik enamikust tõlkeabiprogrammidest eksportida universaalses vormingus, mida saab kasutada ka teistes tõlkeabiprogrammides.

¹ <https://www.matecat.com/plugins/aligner/index#/>

² <https://www.lr-coordination.eu/>

Hanke tingimustes võiks olla määratletud, kas tõlkemälu kasutatakse, kes seda hooldab ning kellele kuuluvad selle edasise kasutamise õigused.

Turul on saadaval nii tasulisi kui ka vabavaralisi lahendusi. Vabavaralistest on olemas näiteks:

OmegaT - <https://omegat.org/>

MateCat - <https://www.matecat.com/>

Terminibaasi loomine ja haldamine

Esmalt tuleks lähtuda EKI terminibaasidest (<https://termin.eki.ee/esterm/>). Asutusepõhisel terminibaasil on mõtte ainult nende terminite osas, mis on nii spetsiifilised, et neid EKI allikatest ei leia või mujal asutustes ei kasutada.

Ühtse terminibaasi loomine on oluline, kuna see aitab tagada tõlgete kvaliteeti ja ühtlustada erinevates valdkondades kasutatavat terminoloogiat.

Tõlke kvaliteedi tagamine

Tõlke kvaliteedi tagamiseks tuleks pärast tõlkimist tekst keeleliselt toimetada. Tõlget ja toimetamist peaksid tegema erinevad inimesed. Tõlge peaks tavalugejale olema lihtsalt loetav ja arusaadav. Toimetaja peaks sihtkeelt rääkima emakeelena. **Tõlkemälude ja terminibaasi kasutamine ja hooldamine aitab kaasa parema kvaliteedi saavutamisele.** Tellija peaks alati tõlke üle kontrollima ja tõlkijat puudustest teavitama.

Kasutage võimalusel ainult testitud ja hinnatud tõlkijaid. Selle nõude tagamiseks oleks soovituslik nõuda pakkujalt Euroopa kirjaliku tõlke kvaliteedistandard EN ISO 17100:2015 olemasolu. ISO nõuetele vastamine (sertifikaadi olemasolu) aitab vältida olukorda, kus proovitöö tehakse kogunud tõlkija poolt, kuid hanke teostamisel kasutatakse algajaid tõlkijaid.

Soovitused hangete korralduse kohta

Kui hankega leitakse endale mitu partnerit, siis on soovitatav koostada pingerida valdkonniti või keelekombinatsiooniti ning esitada tellimusi järjekorras esimesele, selle hõivatuse korral järgmisele. See maandab hankija riske ja võimaldab vajaduse korral valida konkreetseks ülesandeks sobivaima teenusepakkuja. Kahjuks on levinud praktika, kus raamlepingud sõlmitakse majanduslikult soodsaimate pakkujatega, kuid pärast lepingu sõlmimist korraldatakse uus vähempakkumine igale tööle. Pakkujad pakuvad hankijale oma parima hinna, milles on arvestatud hanke kogumahuga. Kõikide lepingupartnerite võistupakkumised iga tellimuse eel tekitavad põhjendamatu halduskoormust nii hankijale kui ka pakkujale.

Soovitame prognoosida hanke töömahtu võimalikult täpselt, et hankedokumentides väljakuulutatatu ei erineks oluliselt lepingu täitmise tegelikust mahust. Levinud on hanked, kus hankija prognoosib väga suurt mahtu, nõuab lepingupartnerilt selleks valmisolekut, kuid tellib väga väikeses mahus. Sellistel puhkudel kui eeldatakse iga päev mitme tõlkija valmisolekut, kuid vajatakse teenust vaid harva, tuleks sõlmida leping mitme partneriga, kes saavad vajaduse korral üksteist täiendada.

Hanke ettevalmistamisel soovitame mõelda läbi, kes on hankija vastutav isik ning kes osaleb töö vastuvõtmisel. **Nii hankija kui ka pakkuja peaksid tagama, et vastutav isik on kogu projekti ajal kättesaadav.** Mõistlik on kokku leppida aeg, mille jooksul peaksid osapooled üksteise küsimustele vastama. Tõlgitavates dokumentides peaks vältima faktivigu (nt ettevõtete nimed, standardite numbrid, sama nimega erinevad dokumendid, erineva nime ja sama sisuga dokumendid).

Hanke võitja määramisel tasuks arvesse võtta nii kvaliteeti kui ka hinda. Kvaliteedi hindamiseks võib tellida lühikese (tasustatud) proovitöö. Kui tõlkeid soovitakse mitmesse keelde, võiks kaaluda igale keelesuunale eraldi hanke korraldamist, sest näiteks ettevõtte, kes suudab pakkuda head EN-ET tõlget, ei pruugi mõnes teises keelekombinatsioonis konkurentidest parem olla. Kui siiski soovitakse valida hanke võitjaks ainult üks ettevõtte, võiks pakkujate tulemused järjestada paremuse järgi kõigepealt keelekombinatsioonide kaupa ning

lõpliku pingerea moodustamisel võtta arvesse töömahtude osakaalu keelekombinatsioonide kaupa: näiteks kui 60% tõlgetest toimub EN-ET suunal, 30% ES-ET ja 10% DE-ET suunal, loeks kõige rohkem EN-ET suuna tulemused, seejärel ES-ET ja alles siis DE-ET suunal saavutatud tulemused.

Keeleabi viited

<http://termin.eki.ee/esterm/> – Eesti Keele Instituudi mitmekeelne terminibaas
<https://ekilex.ee/> - Eesti Keele Instituudi terminihalduskeskkond
<https://sonaveeb.ee/> - Eesti Keele Instituudi sõnaraamatute ühtne päringusüsteem
<https://akit.cyber.ee/> - Andmekaitse ja infoturbe leksikon
<https://stats.cyber.ee/> - Tarkvaratehnika sõnastik
<https://www.eki.ee/dict/ametnik/> - Ametniku soovitus sõnastik
<http://www.keelevaab.ee/> - Keelevaab
http://www.filosoft.ee/html_speller_et/ - Eesti keele speller
<http://keeleabi.eki.ee/index.php> - Eesti Keele Instituudi keelenõuanne
<http://www.eki.ee/dict/gs2006/lyhendid.html> - Lühendivalimik
<http://www.eki.ee/knab/knab.htm> - EKI kohanimeandmebaas (KNAB)
<http://www.ut.ee/taimenimed/> - Eestikeelsete taimenimede andmebaas
<http://www.elus.ee/imetajad/> - Eestikeelsete imetajanimetuste andmebaas
<https://akit.cyber.ee/> - Andmekaitse ja infoturbe leksikon
<http://datasci.ee/masinoppe-sonastik/> - Närvivõrkude ja masinõppe sõnastik
<https://www.eestipank.ee/veebisonastik> - Pangasõnastik
<http://publications.europa.eu/code/et/et-5000500.htm> - Riikide, territooriumite ja vääringute nimistu
<https://iate.europa.eu/home> - IATE European Union Terminology
<https://inboxtranslation.com/resources/glossaries/> - 3000+ Translation Glossaries