

Oskussõnadest valdkondadevahelisel duelliplatsil

Anto Veldre¹

Cybernetica AS tehniline toimetaja

Sissejuhatuse asemel: Miks segab vigane eesti keel minu tööd?

Õigustekstidel ja infoturbetekstidel on suur sarnasus. Avatasemel jagavad nad inimesele juhiseid ühes või teises olukorras käitumiseks, süvatasemel aga peavad terminid olema täpselt paigas, et teadlased ja praktikud saaksid aru, mis olukorraga on tegemist, ning suudaksid sel teemal arutleda. Kui olukord pole selge, siis pole võimalik ka sobivalt käituda (ei seaduskuulekalt ega turvaliselt). Kui *privacy* tõlgitakse ühes dokumendis 'privaatseks', teises aga 'eraelu kaitseks', siis on tegemist kahe täiesti erineva olukorraga.

Kui turvakaamera salvestab infobitte ehk informatsiooni, kuid salvestist ükski elusolend ei vaatle ega tõlgenda (ega ole kaameraserveris näotuvastust ja tehisintellektigi), siis jääb arusaamatuks, mil moel saaksid informatsioonist järsku tekkida (isiku)andmed ja edasi neist andmetest teave (isiku kohta). Tegelikult on asi väga lihtne: eesti keel teeb informatsiooni ja andmete vahel liiga täpset vahet, küll aga mitte indoeuroopa keeled (kus *antus* – ladina 'data', vene 'данные' – läheb sujuvalt üle mitteformeerituseks, *in-* + ladina 'forma'). Andmekaitse ja infoturbe leksikon AKIT² esitab inglise mõistele *information* kaks eestikeelset vastet: 'informatsioon' on miski, mis on esitatud masinale sobival kujul, ja 'teave' inimesele sobival kujul. Järeldus – usutavate kaamerakasutusjuhiste adekvaatne kirjeldamine eesti keeles polegi triviaalne ülesanne. Andmete ja informatsiooni üks võimalikke suhteid on kirjeldatud DKIW püramiidis³, kuid on võimalik ka tõlgendus, et informatsiooni (ja teadmust) saab agregeerimise teel tuletada (andmebaasis hoitavatest) andmetest.

Sotsiaalteadustes tuntakse terminit *schema*. Eesti keelde on see tõlgitud kui 'skeem', kuid praktikas kasutavad spetsialistid omavahel rääkides ikkagi 'skeema'. Aju talletab kogemusi skeemi kujul. Kui inimene tunneb skeemi ära (tuvastab selle), näiteks helistab talle „pangatöötaja“ ja palub nimetada pangakaardi PIN-koodi, siis oskab inimene õigesti käituda. Kui inimene skeemat ära ei tunne (sest seda on ähmastatud jutuga panga sisekontrollist⁴, vajadusest ebaausat töötajat rahakande abil kontrollida vms), siis ohver skeemi ei tuvasta, satub pättuse ohvriks ja jääb oma rahast ilma. Ehk siis, skeem on kognitiivne raamistik, mis aitab inimesel reaalsust mõista ja sellega suhestuda.

Mingil ajal tõdesin, et indoeuroopa keeltest tõlgitud turvajuhised ei olnud üheselt mõistetavad isegi mitte minule – tipptaseme spetsialistile –, saati siis lihtinimesele. Võtame näiteks ühe küberajastu esimestest asjalikest turvajuhenditest⁵: „8.61. Kas konfidentsiaalse teabe

¹ Olen infoturblane ja infosüsteemide analüütik ning viimased kümme aastat on minu töö keelega tihe seos – parandan, toimetan, kirjutan ja tõlgin tehnilisi kirjeldusi ja käsulaudu sisaldavaid tekste.

² Cybernetica AS. [Andmekaitse ja infoturbe leksikon](#), 2011–2023.

³ Vt [DKIW pyramid](#).

⁴ Eesti Pangaliit. [Sergei lugu](#).

⁵ J. Bumgarner, S. Borg. [US-CCU küberturbe kontrollküsimustik](#). U.S. Cyber Consequences, 2007.

edastamise piiramiseks ettevõttevälistele isikutele kasutatakse e-posti filtreid, välja arvatud juhul, kui see on lubatud ja krüpteeritud?“ Asesõna puudutava lausestusvea tõttu pole aru saada, kas lubatud ja krüpteeritud on teave ise, e-post, filter või kasutamine. Olukorrast pole ingliskeelsesse originaali kiikamata võimalik aru saada, skeemat ei moodustu. Järelkult ei saa korralduste subjekt neid juhiseid ka täita. Inglise keeles saab, kuivõrd neil reegel ütleb, et asesõna asendab lause esimese nimisõna – aluse.

Töötasin kolm aastat säärasel unikaalsel ametikohal nagu infoturbeekspert riigiameti PR-osakonnas. Minu peamine ülesanne oli väljuvate infoturbealaste riigisõnumite filtreerimine: et teadmatusest ei valetataks, et kogemata oskussõnu sassi ei aetaks (poliitikutel ikka juhtub – meenutame interneti otsalõppemist ja elektrimolekule) ning et resulteeriv iva/sõnum oleks Harju keskmisele eestlasele selgesti mõistetav. Praegu olen ma kindel, et infoturbeteksti kvaliteetne, skeemipõhine eestindamine maksab tõlkebüroo umbes kümnekordset letihinda. Keegi peab tõlkes kirjeldatud tehnilise situatsiooni pulkadeks lahti võtma ning veenduma, et see oleks isomorfne⁶ lähtekeeles kirjeldatuga ning et sisse ei oleks libisenud soovimatuid kõrvaltähendusi.

Inimese võime sõnalisest kirjeldusest skeemat taastada sõltub otseselt tema emakeele iseärasustest. USA politseil on näiteks juhis selle kohta, kuidas inglise ja hispaania kultuurikeskkonnast pärit inimesi erineval moel üle kuulata – kuna nende tajul on iseärasused. Hiinlaste kohta on teada, et nad tajuvad kontrafaktuaalseid väiteid (“kui oleks olnud, siis poleks”) teisiti kui eurooplased⁷. Erineb rahvuste ja kultuuride värvitaju ning isegi tõe ja vale kriteeriumid. Eesti infoturbestandardit luues selgus, et korraldusi tuleb eestlasele jagada hoopis teises keeleregistris kui indoeurooplastele.

Terminoloogis-semantiline kurbmäng avalikus ruumis

Lahkaksin siin mõnda konkreetset näidet. Kõigepealt, eesti keele puhul ei seisne küsimus enamasti selles, kuidas leida võõrale mõistele see ainuvõimalik terminivaste. Vaid üliharva juhtub, et vasteks pole pakkuda ühtki sõna. Sama harva juhtub, et vaste leitakse kiiresti. Eesti keel erineb indoeuroopa keeltest nii põhjalikult, et tema semantilised ressursid võimaldavad sealse mõistega vastavusse panna mitu vastet, kusjuures ükski neist pole täpne. Võib arutleda ka sedasi, et võõra keele semantika abil sõnastatud mõiste on meie keele seisukohast küsitav juba algusest peale ega saagi eesti keeles taanduda ühele ainukesele selgele vastele. Näiteks *access* – tähendab nii ’pöördus’ kui ka ’juurdepääs’, lisaks on nüanss, et eestikeelne ’pääs’ on oma olemuselt mõlemasuunaline ning tuleb alati täpsustada, kas sissepääs või väljapääs. Siin ja edaspidi käib jutt mitte üldkeelest, vaid teadus- ja tehnikakeelest ning toodud näited saab üle kontrollida AKITist.

Teadjamatel on kindlasti näiteid varasemastki ajast, kuid üks klassikalisi kolmikuid moodustub mikroampermeetrist, galvanomeetrist ja EKG/EEG seadmest. Tehniliselt on tegu küllalt sarnaste esemetega, mida aga eri valdkondades nimetatakse erinevalt. Kuna need valdkonnad

⁶ **Isomorfism** (algebra termin) – vabas sõnastuses: selline sarnasus kahe struktuuri vahel, kus esimese elemendid on üks ühele teise omadega vastendatavad.

⁷ **H. Eifring**. [The Chinese Counterfactual](#). – Journal of Chinese Linguistics 1988, nr 16, 2, lk 193–218.

ei puutu kokku, siis nende vahel aktiivset terminoloogilist võitlust ei toimu. Pisut uuema mõistega *scanner* on aga teisiti. Kord räägitakse skaneerimisest, kord skännimisest, vahel aga suisa skännerdamisest. AKIT räägib infoturbe kontekstis ründest ja Militerm rünnakust, inglise keeles on mõlemad *attack* ning nende kahe vaste ühtlustamiseks puudub igasugune vajadus peale esteetika.

Ent olgu, absoluutsel enamikul juhtudest leidub terminivasteid mitu, tüüpiliselt kolm kuni neli, vahel enamgi. Ka on Eestis mitu terminoloogilist koolkonda (vrd raal, kompuuter, arvuti). Samuti on suur tõenäosus, et eri ülikoolides valitakse samale mõistele erinev vaste.

Eesti Keele Instituudil (EKI) leidusid hoovad varasema trüki- ja telemeedia mõjutamiseks, kuid paraku Internet nende kontrollile ei allu. EKI tegutsemisalaks on jäänud üldkeel ning neil puudub praegu võimekus oskuskeelte (või kitsamalt tehnikakeele) terminoloogiatööd koordineerida ja suunata. Enam ei pääse võidule kõige parem termin (derivatiivne, kõrvaltähendusteta, lühim ja veel kuus tunnust Tiiu Ereli „Terminiõpetusest“⁸). Praegu pääseb võidule termin, mille taga on laialdase sotsiaalse haarde ja meediavõimalustega inimene või organisatsioon.⁹ Termin kvaliteet kahjuks enam ei määra, määrab võimekus seda erinevates keskkondades levitada (sh õigusaktides, mida hakatakse kohe lääne tõlkerobotitele etalonkeelekasutusena ette andma) ning suutlikkus konkurente edestada. Kohati osutub määravaks ametkondlik häbi. Kuidas vaadata ülemusele silma pärast seda, kui oled sunnitud konkurendi terminoloogiale üle minema?

Tuntuimad kaksikud infoturbes on *turvalisus* ja *julgeolek* – paralleelvastad inglise *security*'le. Kõiki rahuldav kompromiss leiti pärast 15 aastat kestnud vaidlusi. Turvalisus on infoühiskonna ja siseministeeriumi teema, julgeolek on kaitse- ja välisministeeriumi teema. Puhuti häirib seda tasakaalu anakronistlik *sisejulgeolek*. Vähem tuntutest mainiksin *juurdepääsu* ja *ligipääsu*. Infoturbe (st AKIT) tunnistab neist esimest, sest vara ohustamiseks tuleb pääseda mitte lihtsalt ligidale, vaid ikka kohe päris vara juurde. On põnev veebis uurida, kuidas osa ametkondi pääsevad *ligi* ja teised *juurde*. Toodetele ja teenustele ligipääsetavuse seadus¹⁰ kahjuks rikkus selle kena tasakaalu, ajendiks vist soov iga nimetamiskorra pealt kaks tähemärki kokku hoida.

Suurima mõjuga skandaaliks infoturbes tuleb pidada isikuandmete kaitse üldmääruse (GDPR)¹¹ väärtõlget. Nimelt, terminitega *risk* ja *threat* läks tõlkes miskit valesti ja nii räägib GDPR meile mõlemas kontekstis ohust ja ohtudest ning ei kordagi riskist. See on tõeline katastroof, kuivõrd riskihalduse terminid on ISO standardites jäigalt paigas ning *risk* ja *oht* on tegelikult erinevad mõisted (riskiks muutub rakendumiskoha, tõenäosuse ja kahjuhinnanguga varustatud oht). Tulemusena pole GDPR eestikeelne tõlge kasutatav. Kuniks jaanalinnu pea püsib kindlalt betoonis, ei ole parandatud versiooni europarlamentidele läbikoputamiseks veel esitatud. Saksa

⁸ T. Ereli. Terminõpetus. Tartu: Tartu Ülikooli Kirjastus, 2007.

⁹ A. Veldre. [Kust on pärit zuumeri eesti keel?](#) – Sirp, 02.09.2022.

¹⁰ RT I, 15.06.2022, 1.

¹¹ Euroopa Parlamendi ja nõukogu määrus (EL) 2016/679 füüsiliste isikute kaitse kohta isikuandmete töötlemisel ja selliste andmete vaba liikumise ning direktiivi 95/46/EÜ kehtetuks tunnistamise kohta ([isikuandmete kaitse üldmäärus](#)).

keeles on olukord eriti huvitav, seal on meie *ohul* tervelt kolm vastet: *Bedrohung* (*threat*), *Gefahr* (*threat*) ja *Gefährdung* (*applied threat*).

Isomorfisuse puudumise näitena taanduvad inglise tüved *manage*, *control* ja *govern*, *drive* ja *lead* (vt AKIT) eesti keeles sageli pelgale juhtimisele. Sõnal *control* on võimalikke eestikeelseid vasteid vähemalt kümme, kõik oleneb kontekstist. On hästi näha, kuidas eesti ja inglise semantiliste ressursside vahel valitseb pigem M : M (mitu mitmele) suhe kui 1 : 1 (üks ühele) suhe. Lisaks sõnale *control* eksisteerib IT-alal terve nimekiri ingliskeelsetest terminitest, mis eesti keelde toomisele kuigi hästi ei allu, näiteks sõnad *pivot*, *function*, *facility*, *space*, *scope*, *credential(s)*, *baseline*, *provision*, *to infer*, *pod/pad* jne. Vastupidises tõlkesuunas on teada abiverbistunud *panema*, mis idioomides ja koos erinevate lisasõnadega annab vähemalt 150 ingliskeelset tõlkevastet. Tõlkeroboteid saab kimbatusse ajada verbidega *leilitama* ja *kiibitsema*.

Mõned terminoloogilised probleemid pole aga üldse ilmsed. Näiteks inglise *security* tähendab eesti keeles ühtaegu mitut asja: tegevusena on see turvamine ehk turve, aga seisundina on see turvalisus. Sarnaselt ei selgu hädaolukorra seadusest¹², kas analüüs on tegevus või dokument. Ent skeema mõttes on väga oluline vahe, kas asju tuleb ka analüüsida sisuliselt ja pidevalt või tuleb lihtsalt dokumendiplank ära täita ja sellega ongi just kui kõik tehtud. Inglisekeelne *-ing* märgib vaheldumisi kaht erinevat grammatilist konstruktsiooni (*gerundium* vs. *present participle*), samas kui eesti *des*-vorm nõuab tegija nimetamist ilmsi (vähemasti tehnilises ja juriidilises kirjanduses).

Eesti keelele on omane sõnu lühendada – kui mitte muuks, siis liitsõnade moodustamiseks (näiteks *turvamine* ja *turve*). Tekib küsimus, millal öelda *turva-* (näiteks *turvamees*) ja millal *turbe-* (näiteks *turbemeetod*). Tegelikult tuleks kasutada *turbemeest* (omastav sõnast *turve*, lühike vorm *turvamise mehest*). Tänu laulusõnadele on aga väär vorm meemina kinnistunud, sellega võitlemiseks puudub mõte. Kuid kas õige oleks *infoturvapoliitika* või *infoturbepoliitika*? Poliitika puudutab kindlasti turvalisust kui saavutatavat seisundit ja mitte niivõrd turvamist kui tegevust (*turvet*, omastavas *turbe*). Seega on tegemist info turvalisuse poliitikaga, lühidalt infoturvapoliitikaga. Rääkimata pisiasjast, et sõnu *bio*, *geo*, *krüpto* ja *info* korrektsetes (st AKITile tuginevas) eesti IT-keeles ei eksisteeri, need vormid on lubatud üksnes liitsõna esiosana. Üldkeelt kirjeldav ÕS neid nüansse ei kajasta.¹³ Siit paistab, kuidas eesti keelele omane lühendamine tekitab kohati palju segadust.

Lisaks segadusele tekitab lühendamine ka otseseid probleeme, seda siis, kui lühendamise käigus toimub tegijale märkamatu tüvemuutus.

- ***avaandmed*** – kas need on august/avausest läbitiritavad andmed või hoopis avalikud andmed? *Ava* ehk siiski ei teki *avaliku* lühendamisest? Tähelepanelik lugeja kindlasti märkas, et sellegi artikli sissejuhatavas lõigus on kasutatud sõna *avatase*.

¹² RT I, 09.08.2022, 24.

¹³ Vt ÕS-ist märksõna *info*.

- **avaleht** – kas see on auguga leht või esimesena avanev leht? Lauserütmi ja soome keelegraafika riimuse poolest sobib hoopis paremini *avakuva*.
- **toimepidevus** – kas see on toimingute pidevus või mingi mõjuri katkematu toime (*effect*)? Talitluspidevusega (*business continuity*) säärast probleemi ei esine, kuivõrd seda ei anna teiseks tüvisõnaks (*talipidevuseks*) lühendada. Toimingu lühendamine ei tohiks meid kuidagi viia eraldi tähendust omava toimeni.
- **eriliigiga** on asjad üldse pahasti – *erinev* ja *eriline* ei eristu. Praegu tundub, et *eriliigilised* (mitmuses) on asjad siis, kui nad omavahel erinevad (näiteks eriliigiliste teede ristmik), aga *eriliiki* (ainsuses) on ohtlikud asjad, mida tuleb hallata vastutustundega. Sellegipoolest räägitakse läbisegi nii isikuandmete eriliigist, eriliiki isikuandmetest kui ka isikuandmetest. See teema vajaks hädasti keelekorraldaja abi.

Liiklusseaduse¹⁴ vastuvõtmisel saabus selgus, et *eri-* on pelk superklass ('eriülesandega salk'), selmet osutada kahe objekti erinevusele. Paraku, *toimepidevus* on hädaolukorra seaduses kasutusel senimaani nagu ka *avaanne*¹⁵ valitsemiskontseptsioonis.

Kõige tõhusam on korraldada sõnaus. Sõnause tulemus ei pea vastama ühelegi „Terminiõpetuse“¹⁶ üheksale kriteeriumile (sest *vox populi*). Näiteks ingliskeelsele sõnale *resilience* on teada hulk vastekandidaate: elastsus, kerksus, paindlikkus, säilendõtkus, vastupidavus, vastupanuvõime. Hädasti oleks vaja arbiitrit, kes neist julgeoleku ja infoturbe valdkonnale ühe välja valiks. Selline kogus sünonüüme ühe võtmetermi jaoks on pisut liiga palju. Praegu on lihtsam valida *ad hominem* tee, kui et tegeleda terminivastega.

Mõne oskussõna saamisluug on aga oluliselt huvitavam kui lihtsalt koolkondade võitlus. Tänapäeval tuleb arvestada, et IT- ja infoturbe terminid käivad sageli ajast ees ning näitavad teed aladele, kus IT-le leitakse rakendus märkimisväärselt hiljem. Nii näiteks tegi AKIT aastal 2017 pärast ID-kaardi „teoreetilist turvaviga“ vahe sisse ingliskeelses sõnas *token* sisalduvatele eri mõistetele. *Tõuken*, mida lingvistid Pythoniga parsivad, on 'sõne' (mõni ütleb ka stringi kohta 'sõne', kuid sõne on ikkagi vaid 'täenduslik string'). Too teine *tõuken*, mis antakse olemise volituste tõestamiseks, on *token* ('isikutõend, volitustõend'). Siis on veel too kolmas *tõuken* ('tingnimi'), mida kasutatakse mõne andmeelemendi varjamiseks. On veel mõiste *security token*, mida saab tõlkida kord pääsmikuks, kord identsustõendiks (ehk tookeniks) ja lõpuks ISO/IEC 10181¹⁷ kontekstis isegi turvatalongiks. Itimehed pruugivad suulisel kõnes varianti *tõuken*, mida formaalsed reeglid väidetavalt eesti keelde tuua ei luba, sest täht *õ* inglise keeles puudub. Selgitus ei veena, kuid on olnud tõhus.

Siis aga saabus aasta 2020 ning seoses krüptorahanduse edendamise tekkis vajadus mõnede terminite toomiseks meie õigusaktidesse. Rahandusministeeriumil, kes seda tööd vedas, polnud

¹⁴ RT I, 20.06.2022, 70.

¹⁵ Anne pole siiski pelk sõnamäng. Ainsuse järele on vajadus olemas, kunstsõna *annes* kahjuks ei juurdunud. Soome keeles on näiteks *informatsioon* – 'tieto', *andmed* aga 'tietot', *faiil* on 'tiedosto' ja *teadmus* on 'tiedämys'.

¹⁶ T. Erelt (viide 8).

¹⁷ ISO/IEC 10181-3:1996. Information technology – Open Systems Interconnection – Security frameworks for open systems: Access control framework.

ettekujutust, et a) AKIT on selle terminoloogia ammu eestindanud ning b) krüptoraha alusterminid on määratud infotehnoloogia ja krüptograafiaga, mitte finantsturgude või näiteks merendusega. Siinsamas Õiguskeeles pakuti välja uhiuus vaste terminile *token* – selleks sai *token*.¹⁸ Tõrge selle sõna vastu tekib foneetilistel põhjustel. Otsisin eesti keele 10 000-sõnalisest sagedussõnastikust kahesilbilisi sõnu, mille tüvivokaaliks oleks *o* ja sellele järgneks sulghäälik ("^.oke" või ".ope" või ".ote"). Mitte ühtegi säärast sõna ei leidunud. Leidusid küll sõnad *tootev* ja *totter*. Kokkuvõttes tähendab see, et *tõuken* ja isegi *tokken* sobiksid eesti keelde foneetiliselt paremini kui *token* ja foneetiliselt välistatud *token*. Ei läinud just kõige paremini, sest IT-alal kujunes vasteks kahe *o*-tähega *tooken* (hääldusega [tõuken]) ja rahanduses ühe *o*-tähega *token* (hääldusega [togen]). Eestikeelsete terminite säärane diversifitseerumine on küll vaevalt vajalik.

Esineb seiklusi määratluse *smart* tõlkimisel. Kohati tõlgitakse see 'nutiks', kohati 'targaks'. Ent tegemist on esemetega, mitte elusolenditega. Eesti keeleloogika kohaselt puudub esemel teadvus, järelikult ka tarkus. Iga kord kui asi/ese on tark, tõuseb programmeerija huulile muie. Suurim mure on siiski *smart contract*'i tõlkimisega. Definitsiooni kohaselt on tegu mingis programmeerimiskeeles (näiteks Solidity) kirjutatud seireprogrammiga, mis jälgib kontraktis märgitud väärtuste muutumist ajas vastavalt *IF* või *CASE* lause tingimustele. Küsimus on filosoofiline – kas programmi lähtekood ikka saab talitleda juriidilise lepinguna? Palju arusaadavam oleks öelda, et see programm on lepingumonitor. Vastuargumendina lisati terminoloogiaõpetusse omavoliline kümnes tunnus – meie rannakülas (või kõrvalkabinetis) nii ei räägita.

Peame rääkima ka digikukrust. Nimelt on inglise keeles tore sõna *wallet*, millele siis lisandub erinevaid prefikseid, näiteks *identity wallet*, *cryptowallet*, *eurowallet*. Küsimus on, mis sõnaga *wallet*'it tõlkida. Osa arvab, et see on 'rahakott'. Selle variandi puudus on, et *identity wallet* peaks siis olema 'identiteedikott' (ülemmõiste tähendus erineb). Teised usuvad, et tõlkida tuleks 'rahataskuks'. Probleem, mis tekib, on sama – identiteedi puhul peaks tegu olema identiteeditaskuga. Seevastu *kukkur* üldmõistena on äärmiselt derivatiivne, võimaldades kõiki kirjeldatud liitumisi. Ent tundub, et *kukkur* alamsaksi sõnana oli kasutusel vaid Eesti teatud regioonides. Kui sõna kasutuse üle otsustaja on pärit mujalt, siis pole kindel, kas *kukkur* kasutusse jõuab.

Hea termini (Tiiu Erelti sõnastatud) üheksa tunnust kardetavasti Eestis praeguseks enam ei kehti. Oleks hädasti vaja, et EKI talitleks valdkondadevahelistes terminoloogiaküsimustes arbiitrina, seda teenust vajatakse ametkondadevahelises suhtluses väga.

¹⁸ K. Leppik. [Krüptovaradest](#). – Õiguskeel 2020, nr 2.