



STATISTIKAAMET

Majandus-ja Kommunikatsiooniministeerium
ott.velsberg@mkkm.ee

Teie: 30.01.2024

Meie: 14.02.2024 nr 1.1-3/0082

Statistikaameti tagasiside andmete ja tehisintellekti
valgele raamatule ja selle tegevuskavadele

Statistikaamet peab väga oluliseks, et Majandus- ja Kommunikatsiooniministeerium on seadnud eesmärgi koondada tehisintellekti (edaspidi ka TI) ja andmevaldkonna arendamise ideed ja suundumused ühte dokumenti, mis sellisel kujul on omalaadseks seletuskirjaks lisades olevatele tegevuskavadele. Toome siiski välja rea soovitusi, mille arvesse võtmine teksti edasiarendamisel võimaldaks luua dokumendi, mis kajastaks valdkonda ja selle vajadusi-probleeme-lahendusi terviklikumalt ja tõhusamalt, kaasates suuremat ringi asjassepuutuvaid osalisi.

Andmete ja tehisintellekti valge raamat peaks ülevaatlilikult ja veenvalt esitama valdkonna arengu kitsaskohad ning väljakutsed ja neist lähtuvad tulevikusuunad, mis mõtestaks ja kujundaks Eesti tulevikku andmete ja tehisintellekti valdkonnas. Statistikaamet kui riikliku statistika tootja ja andmehalduse koordineerija esitab peamised kommentaarid valge raamatu osas lähtudes praktilisest kogemusest andmete korrastamisel, analüüsimisel ning andmeteenuste pakkumisel. Statistika tootmine ning mitmesuguste kvantitatiivsetel näitajatel põhinevate analüütiliste tööde teostamine on omamoodi mikromudel riigi andmete ökosüsteemist, mille toimimise käigus tehtud tähelepanekutest võiks kasu olla nii andme- kui ka tehisintellekti valdkonna kujundamisel. Samuti peame oluliseks märkida, et meie kogemuse järgi on nendes valdkondades nii Euroopa Liidu kui ka maailma tasemel ootus Eestile olla suunanäitajaks ja eestvedajaks.

Üldist

Valge raamatu tüüpi dokument peaks näitama lugejale veenvalt, millised võiksid olla valdkonna sihid, milliseid tegevusi võiks konkreetses valdkonnas ette võtta ning kuidas nende lahendusteni on jõutud. Just viimases osas saaks valget raamatut muuta oluliselt paremaks. Näiteks ei ole valges raamatus sageli viidatud esitatud andmestiku allikatele (nt lk 14 teine lõik, kus hinnatakse andmeturu mahtu ja selle kasvu või lk 16 ja edasi dokumendis Eesti avaliku sektori asutuste küsitluse tulemused, jmt) ning see ei võimalda lugejal veenduda kasutatud meetodika asjakohasuses ning kokkuvõttes juba esitatud andmete usaldusväärsuses.

Samuti ei tundu nende andmete põhjal tehtud järeldused tuginevat laiemale ekspertide konsensussele. Nii probleemide illustreerimiseks kasutatud andmestik kui ka andmete alusel tehtud järeldused peaksid saama kinnitust mitmest asjatundlikust allikast, sest nende alusel jõuavad autorid järeldustele, milles seisnevad valdkonna probleemid.

Probleemide ülevaade on omakorda aluseks arengusuundadele ja tegevuskavadele, mille rakendamisele tuleb kulutada riigi ressursse. Kui aga probleemipüstitus on väär, siis asume lahendama valet ülesannet ning seda isegi õigesti tehes ei saa tulemus olla ei tõhus ega mõjus. See on praegusel juhul arvestatav oht, sest käesolev versioon valgest raamatust ei esita üksnes visiooni, vaid püüab veenda kindlate lahenduste õigsuses ning kuna neile ei ole pakutud alternatiivi, siis võib jääda mulje, et alternatiive ei olegi ja toodud lahendus on ainumõeldav.

Oluline on ka see, et hea tava kohaselt esitatakse sellised dokumendid kooskõlastamiseks ja arvamuse avaldamiseks võimalikult laiale asjassepuutuvate asutuste ja isikute ringile ega piirduta mitteformaalsete konsultatsioonidega. Vastasel juhul ei jõua dokumendi koostajateni kõik olulised seisukohad, millega tuleks dokumentide täiendamisel arvestada. Kuna valge paberi tüüpi strateegiadokumente kasutatakse Eestis pigem

vähem (tüüpilised on valdkonna arengukavad, väljatöötamiskavatsused, tegevuskavad, seletuskirjad, õigusaktide eelnõud) ning nende menetlemine ei ole alati ühtne, siis peaks otsima parimat lahendust, et vältida siin riski, mis tuleneb probleemide väärast mõistmisest või ka nende suhtelise olulisuse kallutatud hinnangutest. Mainitud riski maandamiseks võiks püüda keskseid probleeme (NB! neist kõik ei ole koondatud SWOT-analüüsi, vaid on kirjeldatud dokumendis) eraldi konsultatsioonide käigus valideerida ning esitada ka lahenduste alternatiive, mis siis omakorda oleks võimalik tegevuskavades konkretiseerida.

Näiteks on meie hinnangul, vaatamata dokumendi püüdele haarata kõiki olulisi aspekte, jäänud kahetsusväärset käsitlemata olemasolevate meetmete ja valdkonna korralduse tulemuslikkus ning senine mõju. Valge raamat ei käsitleni tänast valdkonna korraldust (nt institutsionaalset, sh õiguslikku, ega ka keskvalitsuse ja kohaliku omavalitsuse tasandi erinevusi) ja sellest tulevaid probleeme (nt valdkondade ja valitsemisalade ülestevuste korraldamine Eesti avalikus halduses ja rahastamise kitsaskohad tegevuspõhises eelarvestamises). See lünk aga ei takista autoritel pakkumast dokumendis lahendusi ja tegevusi, mis võivad hoopis raskendada probleemide lahendamist või neid lausa süvendada (vt ka alapeatükk „Institutsionaalne raamistik“). See peaks muutma lugeja ettevaatlikuks pakutud tegevuskavade asjakohasuse suhtes.

Juhime tähelepanu ka asjaolule, et mitte üheski kontekstis ei ole dokumendis peetud vajalikuks käsitleda ega ka mainida ei riiklikku ega ametkondlikku statistikat. Selline analüütiline andmeteenus on laialt levinud ja see on oluline sisend andmepõhisele otsustamisele. Samuti on nende ülesannete täitmine seotud kuludega, mis üle kogu avaliku sektori võivad olla märkimisväärsed ning mõjutavad oluliselt ka asutuste juhtide otsuseid nii andmete korrastamise kui ka tehisaru rakendamise osas.

Samuti on just erinevat liiki statistika ja analüütika kvaliteet hea indikaator, mille probleemid viitavad andmehalduse probleemidele ning edusammud jällegi tulemuslikule andmekorraldusele. Seepärast teeme ettepaneku käsitleda andmete valges raamatus avaandmetele lisaks ka riikliku statistikat, mis on avalik hüve, mahukas ja keerukas andmeteenus ning riiklik ülesanne, mis vajab korraldamist nüüd ja ka edaspidi. Selle kvaliteet ja seos muu andmevaldkonnaga mõjutab otseselt andmepõhiseid otsuseid ja kujundab ka väga praktiliselt ja suurel määral andmevaldkonna arengut (nt registrite kaudu).

Statistikaameti ettepanekud valge raamatu täiendamiseks

Prioriteedid

Teeme ettepaneku valges raamatus selgelt välja tuua vähemalt esimese kahe aasta prioriteedid (samaks perioodiks on tehtud tegevuskavad ning kuna valdkond areneb väga kiirelt, on mõistlik olla paindlik ning vajadusel olukorra muutudes prioriteete ümber tõsta).

Samuti soovime kirjeldada protsessi, kuidas nende prioriteetide kaasajastamine võiks toimuda (nt kas sarnaselt selle valge raamatu koostamise konsultatsioonidele koos tavapärase kooskõlastamise menetlusega, või muus vormis).

Andmepoliitika ja seosed teiste arengukavadega

Teeme ettepaneku pöörata valdkonna hetkeseisu kirjeldamisel tähelepanu asjaolule, et Eestis puudub ühtne andmepoliitika, nii riigi sees kui ka Euroopa Liidu suunas. Selline poliitika võiks olla nt selle valge raamatu lisa, milles on lühidalt kirjeldatud nii meie peamised riigisisised kui ka rahvusvahelised positsioonid (nt toetame andmete ühekordset küsimist ja mitmekordset kasutamist; peame oluliseks andmete tasuta taaskasutamist avaliku hüve loomisel jmt). Mõningaid selliseid keskseid põhimõtteid leiab ka valgest raamatust, ent nt Euroopa Liidu poliitikate osas on Eestit kirjeldatud kui passiivses rollis olijat, kes järgib juhiseid, kuid ei võta poliitikakujundamises eestvedaja rolli, mida samas edukalt digiriigilt võiks oodata ja mida meilt sageli ka tegelikkuses oodatakse.

Kuigi rõhutatakse, et kõiki seotud arengukavasid ei mainita, siis teeme ettepaneku tuua arengukavade plokis välja siiski ka halduspoliitika programm 2024–27, mis on riigivalitsemist suunav strateegiline dokument ning seetõttu peaks valge raamatu eesmärgid ja tegevused sellega sobituma. Riigireformi tulevik on meie hinnangul väga suuresti seotud just andmete valdkonna ning tehisaru arengutega.

Institutsionaalne raamistik

Andmepõhine riigikorralduse ja majanduse peatükk alapeatükk „Ökosüsteem“. Eesti valdkonna hetkeseisu kirjelduses ei ole eraldi käsitletud olemasolevat andmete ökosüsteemi, täpsemalt institutsionaalset raamistikku. Ökosüsteemi plokis on kirjeldatud, milline näeks ökosüsteem välja aastaks 2030, sh ühe tegevusena mainitakse ökosüsteemi kontseptsiooni loomist. Tegevuskavas on kompetentsikeskustele ette nähtud ligikaudu 7 miljonit eurot, mis viitab, et dokumendi koostajatel on oma ettekujutus olemasoleva

korralduse probleemidest ning samuti ka visioon, milline tulevane ökosüsteem välja peaks nägema. Eestis on juba praegu olemas mitu asutust, mille ülesannete hulka kuulub andmehalduse, andmeteenuste, andmekorralduse jne koordineerimine, korraldamine ja pakkumine. Neist asutustest Statistikaamet täidab andmehalduse koordineerija rolli baasrahastusega. Sellest tulenevalt teeme ettepaneku valges raamatus täpsemalt kirjeldada, milline on valdkonda kujundav institutsionaalne raamistik, mille kontseptsiooni hakatakse uuendada, millised on nende asutuste võtmeülesanded ning rahastuse tagamise mehhanism.

Institutsionaalse raamistiku sees vajaks eraldi tähelepanu kompetentsikeskuste käsitus. Valges raamatus on mitmes kohas viidatud nii andmehalduse kui ka valdkonna kompetentsikeskustele, kuid pole välja toodud, milliseid asutusi silmas peetakse. Lisaks mainitakse uue kompetentsikeskuse loomist olukorras, kus ühe nõrkusena on dokumendi alguses välja toodud valdkonnas tervikuna puuduolevad andmehalduse, andmeanalüüsi ja andmeteaduse jm kompetentsus ja oskused. Arvestama peab ka valges raamatus esitatud eesmärkide ambitsioone (haldusalad peavad looma tugiüksused; kõik avaliku sektori organisatsioonid peavad koostama pikaajalise andmete ja TI strateegia ning tegevuskavad ning tagame ressursi nende efektiivseks rakendamiseks) ning kõrvutame seda praeguse reaalsusega, kus vähesed asutused on seni endale vajalike oskustega töötajad leidnud, sh pole seni ka palju õppevõimalusi, sest eksperte on vähe ja valdkond uus. Seetõttu teeme ettepaneku keskenduda seniste valdkonda edendavate asutuste tugevdamisele ja baasrahastuse tagamisele, et koostöös teisi asutusi koolitada, nõustada ning tugevdada. See on eriti oluline lähiaja riigieelarve olukorda vaadates, kus peame keskenduma eelkõige vältimatutele kuludele ja investeeringutele.

Õigusraamistik

Teeme ettepaneku pöörata põhjalikumalt tähelepanu väljakutsetele õigusraamistiku osas. Õigusraamistik on kohati andmevaldkonna arengu takistuseks, selle reeglid on tihti killustunud, tekitavad ebavõrdsust era- ja avaliku sektori vahel ning ei suuda ajaga kaasas käia, probleeme ennetada ega mõistlikult lahendada. Nt riikliku statistika teenuse vaatest takistab valdkonna praegune regulatsioon eraandmete laiemat kasutuselevõttu ning selle lahendus eeldab mitme õigusakti muudatusi, millele seisavad vastu nii eraettevõtted kui ka mõned riigiasutused.

Teeme ka ettepaneku käsitleda valges raamatus avaliku- ja erasektori erinevaid andmevajadusi ja võimalikke huvide konflikte. Hetkel on esiplaanil riigi roll teenuse pakujana, kuid piisavat tähelepanu pole saanud riigi roll võimaldajana ja õigusraami loojana (sh ka Euroopa Liidu õiguses), mis peaks parimal moel toetama ja tagama andmemajanduse arengut ning andmete rist- ja taaskasutamist mitte ainult era- vaid ka avalikus sektoris.

Õigusraamistik peab toetama ka eelpool nimetatud andmepoliitikat. Heaks näiteks on siin Euroopa Liidu andmehalduse määruse rakendamise küsimus, kus Eesti peaks meie hinnangul olema ambitsioonikas ning rakendama määrust moel, mis võimaldab suurimat võimalikku andmete taaskasutamist ning andmemajanduse kasvu.

Tehnoloogiline mahajäämus ja baasrahastus

Valges raamatus ei ole meie hinnangul piisavalt tähelepanu pööratud andmevaldkonna tehnoloogilisele mahajäämusele ja selle valdkonna rahastuse puudustele, mis seostub ka eespool kirjeldatud institutsionaalse raamistiku problemaatikaga, nt juhime tähelepanu, et siia maani on riigieelarvest keskpikas vaates rahastamata Statistikaameti kui andmehalduse koordinaatori roll, mis peab kindlustama juhiste andmise ning asutuste nõustamise andmekirjelduste, standardite, andmekvaliteedi jm osas ning vastavate infotehnoloogiliste rakenduste arendamise. Seda probleemi tuleks käsitleda koos valges raamatus tehtud tõdemusega, et andmete kvaliteet on madal ning soovitusena suunata riigiülesele märkimisväärselt enam ressursse (investeeringud, personal) andmete kvaliteedi haldamisse.

Haridus ja oskused

Valges raamatus rõhutatakse meie hinnangul õigesti valdkonna oskuste puudumisega seotud erinevaid aspekte. Kuigi kompetentsi ja hariduse teemades on kajastatud, et tarvis on pakkuda õpet erinevatel andmeerialadel, siis teeme ettepaneku lisada vajalike erialade hulka sõnaselgelt ka TI valdkonda tundvad õigusekspertid ja üleüldiselt andmeõiguse ekspertid.

Terminoloogia ja standardid

Mitmed valges raamatus kasutusel olevad terminid võiksid olla raamatu alguses lisaks olemasolevatele definitsioonidele samuti defineeritud, sh andmemajandus; andmekorraldus, andmeüksus, andmealtruism, riigi andmearhitekt.

Tehisintellekti definitsiooni puhul teeme ettepaneku kasutada Euroopa Liidu tehisintellekti definitsiooni, eriti oluline on fraas: *"given set of human-defined objectives"*, mis on kõigis varasemates TI määratlustes ja mille OECD on välja jättnud.

Standardite kättesaadavaks tegemine ja kommunikeerimine asjaosalistele on kahtlemata vajalik, kuid standardite tähendus ja vajalikkus, töö standardiga, vastavus Eesti oma standardi ja rahvusvahelise standardi

vahel, standardi rakendamine asutustes, standardi arendamine-ajakohasus (sh mis rakenduses) jne vajab omaette tähelepanu ja sisustamist kasutaja jaoks. Andmete tegevuskavas ettenähtud tegevused ei kata meie hinnangul (praktilisel kogemusel põhinedes) ära kogu süsteemi vajadusi. Nt andmete tegevuskavas on kirjas üksnes asutuste keskne toetamine avaandmete avalikustamisel, samas on asutusel oluliselt enam andmeid, mis vajavad ka haldust ja seepärast toetamist ning nõustamist.

Mõõdikud

Andmehalduse, aga ka teiste osade mõõdikute metoodika on ebaselge ning mõnes kohas on mõõdik alles väljatöötamisel. Teeme ettepaneku pöörata mõõdikutele täiendavalt tähelepanu ning täpse metoodika koos asjasse puutuvate osapooltega kas koos välja töötada või vähemalt kooskõlastada.

Mõõdikute probleemi illustreerimiseks on näiteks andmehalduse plokis välja toodud tulemusena „Lõppkasutajate (kodanike) halduskoormuse vähendamine vähemalt 70% võrrelduna 2023. aastaga.“ Siinkohal oleks vaja täpsemat infot, milliste andmete põhjal on algseis fikseeritud, eriti olukorras, kus samas dokumendis viidatakse, et halduskoormust paljudes asutustes süsteemselt ei hinnata.

Eraldi rõhutame, et igasugune andmemajanduse sidumine SKP arvestusega (st suhtarvude kasutamine, nt andmemajandus moodustab teatava protsendi SKP-st või panustab SKP-sse mingi teatava summa või protsendi) on väga oletuslik, metoodiliselt veel lõpuni välja arendamata ning sellise hinnangu lisamist rahvamajanduse arvepidamisse ei peeta veel piisavalt küpseks.

Andmekirjaoskus ja ühiskondlik valmisolek

Andmekirjaoskus ei ole pelgalt tehniliste oskuste pagas, vaid see sisaldab ka üldisemaid hoiakuid ning oskust andmeid ning andmeteenuseid kriitiliselt hinnata ja tarbida. Kuna oma töös puutub Statistikaamet tihti kokku andmekirjaoskuse lünkadega ning nende oskuste edendamine on üks Statistikaameti fookussuundi, siis leiame, et valges raamatus tuleb pöörata tähelepanu meetmetele, mis kujundaks ühiskondlikke hoiakuid andmete töötlemise kohta ning rõhutaks ühiskonna mõistmise ja toetuse olulisust andmete kasutuse juures. Kogu valge raamat on kantud eeldusest, et ühiskond tuleb kergesti nii kirja pandud eesmärkide kui ka tegevustega kaasa, kuid meie senine kogemus viitab pigem, et need eeldused pole enamjaolt iseenesestmõistetavad. Nende eelduste loomise nimel tuleb palju ja sihipäraselt veenmistööd teha. Seepärast teeme ettepaneku põhjalikumalt käsitleda andmekirjaoskuse teemat ka n-ö ühiskondliku kokkuleppe võtmes. Seda selleks, et praegusest visioonist ei kujuneks kitsast ametkondlikku ambitsiooni, millega ei suuda end seostada ei inimesed ega ettevõtted.

Lõpetuseks märgime, et hindame Majandus- ja Kommunikatsiooniministeeriumi pingutusi andmete ja tehisintellekti valge raamatu koostamisel ning peame väga oluliseks valge raamatu ja tegevuskavade sisulist kvaliteeti, täpset probleemipüstitust, läbikaalutud ambitsioone ning kõigi huvitatud osapoolte arvamuse sisulist arvestamist. Loodame, et meie esitatud märkustest on nende dokumentide täiendatud versioonide kirjutamisel abi ning oleme valmis sellesse protsessi igati panustama.

Lugupidamisega

(allkirjastatud digitaalselt)

Urmet Lee
peadirektor

Liina Osila 625 9313
liina.osila@stat.ee