

# Kokkuvõte

## Uuring „Lingitud Eesti“

Uuringu autorid:

Peep Küngas  
Hele-Mai Haav  
Jaak Tepandi  
Lauri Tepandi  
Mihkel Lauk  
Triin Tars

Uuringu algatasid Majandus- ja Kommunikatsiooniministeerium ning Eesti Infotehnoloogia ja Telekommunikatsiooni Liit. Uuring viidi läbi Riigikantselei tarkade otsuste fondi ja Euroopa Sotsiaalfondi toel.

Käesolev dokument on lühikokkuvõtte pikemast ning mahukamast uuringust. Kokkuvõttes toodu mõistmiseks soovitame lugeda uuringu lõppraportit.

Tervikaruanne: ISBN 978-9949-33-722-4



# Kokkuvõte

**„Lingitud Eesti“ all mõistetakse IKT-lahenduste toel loodud kogu Eesti inforuumi hõlmavat hajutatud isikute, organisatsioonide, tarkade seadmete, teadmiste, infosüsteemide ning andmete koosvõimelist võrgustikku.**

---

Linkandmed on struktureeritud andmed, mis on kirjeldatud ja avaldatud viisil, mis lubab neid automaatselt seostada. Linkandmete tehnoloogia võimaldab seni hajutatult eksisteerivad andmebaasid, sotsiaalmeedia, dokumendid, faktid, seadmed, teadmised jms integreerida ühtseks tervikuks ning avab seeläbi riikidele, era- ja vabasektorile seni võimatud olnud ehk põhimõtteliselt uued võimalused uut tüüpi teenuste loomiseks. Kuid selle võrgustiku loomine eeldab läbimurret mitmetes suundades nagu ülikiire andmesidevõrk ja targad seadmed, koosvõimelised infosüsteemid ja teadmusvõrgustikud, semantilised võrgustikud – kõige olulisemana aga linkandmete kasutuselevõttus.

Linkandmete kontseptsiooni järgi saavad lisaks dokumentidele otselingitavaks ka andmed ise ja kõik füüsilised kui ka abstraktsed „asjad“. Kui praegu saavad brauserite abil teavet lugeda eelkõige inimesed (andmed on inimloetavad), siis kasutades linkandmeid toetavaid tehnoloogiaid, saavad arvutid andmeid lugeda – ning mis olulisemgi, eri kohtades paiknevaid andmeid siduda, nendes otsida ja neid töödelda. Linkandmete tehnoloogia võimaldab loobuda traditsiooniliste andmekogude loomisest ja nende seostamisest: kasutades linkandmete tehnoloogiaid on see funktsionaalsus andmete avaldamisega vaikumisi saavutatav.

**Linkandmete kasutuselevõtu initsiatiiv edendab läbipaistvust, koostööd riikide tasandil ja laiemalt, võimaldab luua uusi, uuenduslikke teenuseid ning tõsta otsuste tegemise kvaliteeti.**

---

Linkandmed kui tehnoloogiline suundumus on tihedalt seotud selliste ülemaailmsete trendidega nagu tulevikuinternet (*Future Internet, FI*) ja semantiline veeb (*Semantic Web*), asjade Internet (*Internet of Things*) ja suurandmed (*Big Data*), avaandmed (*Open Data*) ja koosloome (*Crowdsourcing*).

Eestis puudub terviklik, Eesti konteksti ja vajadusi arvestav uuring linkandmete kasutusvõimaluste osas, millel võiks põhineda linkandmete kasutuselevõtu poliitika. Käesoleva uuringu tulemused on otseseks sisendiks riigi IKT-poliitika kujundamisele: järgmiste aastate infoühiskonna arengukava rakendusplaanide koostamisele, struktuurifondide rahastamisvoorude ettevalmistamisele ja koosvõime raamistiku dokumentide täiendamisele.

Uuringu eesmärgiks on loodava „Infoühiskonna arengukava 2020“ eesmärkide saavutamiseks vajalike rakendusplaanide jaoks sisendi andmine lingitud andmete loomise ja kasutamise osas ning vastavate tegevuste efektiivseks elluviimiseks vajalike muudatuste tegemiseks Eesti õigusruumis ja poliitilistes meetmetes.

Uuringus pakume välja lähenemise linkandmete tehnoloogiate laiemaks kasutuselevõtuks Eestis, tuues välja puudused, mis pärsivad lingitud andmete efektiivset loomist ja kasutamist. See hõlmab soovitusi avaliku sektori, erasektori ja vabasektori andmekogude, veebisaitide ja muude inforessursside ümberkorraldamiseks ning neid toetavateks poliitikameetmeteks riigi poolt. Järgnevas toome välja ettepanekud ja poliitikasoovitused. Nende põhjendused ja täpsem kontekst on esitatud uuringu täispikas aruandes.

## Valdkonna globaalne areng

Kokkuvõtvalt võib öelda, et linkandmetega seotud tehnoloogiatest on avaliku sektori avaandmed ja NoSQL andmebaasisüsteemid (siia alla kuuluvad ka RDF andmebaasid) ajaliselt kõige lähemal produktiivsuse platoole (2-5 a). Suurandmete hõive ja analüütika tehnoloogiad on kasutuselevõtu faasi alguses. Gartneri

2012. a suurandmetega seotud tehnoloogiate kasutuselevõtu graafikus oli MapReduce jt suurandmete tehnoloogiad kasutuselevõttust 2-5 aasta kaugusel. Teised seonduvad tehnoloogiad nagu semantiktehnoloogiad ja asjade internet on produktiivsuse platoost rohkem kui 10 aasta kaugusel.

Avaandmete arengul on otsene mõju avalike linkandmete kasutuselevõtule ja linkandmete kasutuselevõtule üldse, sest ka erasektor saaks oma andmetega linkida avalikke linkandmeid ja seega linkandmete mõju veelgi suurendada. Tuleb tähtsustada rahvusvaheliste ja rahvuslike initsiatiivide ja tegevuste mõju avaliku sektori andmete avalikustamisel ja avalike linkandmete edendamisel.

2013. a oli avaandmete seisukohast väga tähtis aasta: avaldati G8 riikide avaandmete harta ja USA Presidendi otsekorralduse riigivalitsemise uue digitaalstrateegia kohta, W3C avaldas linkandmete platvormi parimad praktikad ja juhised, ODI laienemine (Prantsusmaale, Dubaisse, Kanadasse, USA-sse, Argentiinasse ja Venemaale) ja avaldati Eesti avaandmete rohelise raamatu mustand.

Nii OECD kui EK oma peadirektoraatide kaudu on pühendanud palju initsiatiive avalikele avaandmetele ja nende linkimisele ning seotud tehnoloogiate toetamisele. Eraldi võib märkida OECD avaandmete projekti ja avaandmete indikaatorite liitumist OECD 2013. a ülevaatesse „*Government at a Glance*“. Euroopa Komisjoni tegemistest võib esile tõsta DG CONNECT EL digitaalarengu tegevuskava 2020, DIGIT töötas välja vabavara strateegia perioodiks 2011-2013 ja haldab ISA programmi.

Euroopa Komisjon on toetanud tervet rida teadus- ja arendusprojekte oma FP7 IKT tööprogrammi raames selleks, et vähendada nii linkandmete loojate kui kasutajate sisenemistõkkeid valdkonda. Nende projektide hulgas on suurema mõjuga LOD2, LATC, PlanetData, TELEIOS, BIG ja IKS projektid. EL uus teaduse ja innovatsiooni raamprogramm Horizon 2020 perioodiks 2014-2020 rahastab aastatel 2014-2015 palju ava- ja linkandmetega seonduvaid teadus- ja arendusprojekte. Põhiline fookus on selles suurandmete teema sidumisel avalike avaandmetega. Eriti palju tähelepanu saavad 2014-2015 tegevusprogrammis tulevikuinterneti ja asjade interneti teemad.

Ava- ja linkandmete rakendamisel on tähtis roll rahvusvahelistel standarditel ja soovitel, mida annab välja rahvusvaheline veebikonsortsium W3C. Linkandmete rakendamisega seoses on tähtsad rahvusvahelised sõnastikud ja ontoloogiad. Praegu on W3C poolt soovitatud ontoloogiate hulgas juba inimeste ja organisatsioonide standardontoloogiad.

PwC EU Services avaliku sektori lingitud avaandmete ärimudelite alase uuringu järeldus oli, et avalikud linkandmed liiguvad edasi tehnoloogia kasutuselevõtu elutsüklis olles näiteks raamatukogunduse/arhiivinduse alal jõudnud juba varajase kasutuselevõtu faasi. Lõppjäreldusena toodi uuringus välja, et kuna põhilised andmete pakkujad on tootmas suures mahus avalikke linkandmeid, siis on ette näha, et ka nende andmete taaskasutajad leiavad lähitulevikus oma võimaluse innovatsiooniks.

Kui poliitiline motiiv avaandmete osas oli elavdada soiku jäänud majandust, siis suurandmete osas on poliitika stimuleerida andmeanalüüsil põhinevat otsustusprotsessi ning kasutada mittestruktureeritud veebiandmeid (põhiliselt sotsiaalmeediast) riiklike missioonikriitiliste rakenduste loomiseks.

Kasulikkuse aspektist on suurandmete tehnoloogia oma kasulikkust teatud määral juba tõestanud. Suurandmete tehnoloogiate kasutuselevõtt ja selle majanduslik kasu ettevõtluses ja avalikus sektoris kasvab vastavalt sellele, kuidas ollakse võimelised neid tehnoloogiaid kasutama keeruliste ülesannete lahendamiseks. Siin ei ole kitsaskohaks niivõrd tehnoloogia kui andmetöötajate analüütiline kompetentsus. Just personali suurandmete analüütika alast ebakompetentsust peetakse põhiliseks suurandmete tehnoloogiate kasutuselevõtu piirajaks.

SAS-i uuringu „*Big-Data-in-Big-Companies*“ järgi on selge, et suured tööstusettevõtted lähevad üle suurandmete töötlemise ja analüütika tehnoloogiatele kuigi see võtab aega ja nõuab täiesti uut tehnoloogilist lähenemist, muutuvat organisatsiooni ja juhtimisstruktuuri ning uusi oskusi. Suurandmete peamine väärtus tuleb just nende töötlemisest ja analüüsist ja toodetest ning teenustest, mis lähtuvad selle analüüsi tulemustest.

Kuigi on üsna selge, et suurandmete tehnoloogiate kasutuselevõtt võimaldab suurendada majanduslikku kasvu, peab siiski pöörama tähelepanu ka selle sotsiaalsetele aspektidele. Nendeks on autoriõiguste tagamine, turvalisus ja privaatsus. Seega saab suurandmete tehnoloogiate kasutuselevõtu kriitiliseks eduteguriks balansseerimine kättesaadavate andmete ja nende üha suureneva linkimise ning analüüsimise soovi ning võimaliku sotsiaalse häirumuse vahel.

Kaubamärkide omajad lasevad praegu iga nädal välja tooteid, teenuseid ja sisu, mis on inspireeritud ja loodud internetikogukondade poolt. Aastal 2014 loodetakse näha koosloome efektiivsuse tõestust turunduse ja innovatsiooni valdkondades. Gartner ennustab, et koosloomet kasutavate firmade kasum tõuseb 1% võrreldes firmadega, kes koosloomet ei kasuta ja seda juba aastaks 2015.

Praegu on murranguline periood interneti arengus, sest kommunikatsioonile orienteeritud internet peaks vastavalt muutunud nõuetele ja vajadustele muutuma sisule orienteeritud internetiks. Seoses sellega muutub aktuaalseks uute internetiarhitektuuride arendus. Pragmatilisest vaatepunktist lähtuvalt on praegu maailmas tähtis üleminek uuele internetiprotokollile IPv6, sest IPv4 aadresside hulk on ammendumas. IPv6 võimaldab kasutusele võtta uue aadressruumi ja lülitada internetti üha rohkem asju s.o sensoreid jms.

Praeguseks olemegi juba jõudnud olukorrani, kus rohkem seadmeid (asju) kui inimesi on liitunud internetiga ja ennustatakse, et internetti ühendatud asjade hulk ületab peagi inimpopulatsiooni arvu, moodustades 20-50 biljonit ühendatud tarka seadet. See tähendab ka ühtlasi seda, et interneti protokoll IPv4, mis on mõeldud 4 biljoni aadressi jaoks, ammendub.

Aastaks 2020 ennustatakse Internetti lülitatud seadmete arvuks 50 biljonit kuni 1 triljon. Sensor- ja võrguvõimekuse lisamine asjadele on kaotanud oma tehnoloogilised ja hinnabarjäärid. Praegu toimub asjade interneti tehnoloogiate läbimurre tarbijaturule ja toodetakse üha rohkem tarku laiatarbetaoteid. Personaalsete jälgimisvahendite ja kantava tehnoloogia turg peaks kasvama ülikiiresti just sel aastal. Personaalsete jälgimisvahendite hinnad langevad ja erinevad targad kellad, tervise ja inimkeha seisukorra näitajate jälgimise vahendid jms tooted tulevad turule. Asjade interneti laialdasel kasutuselevõtul on ka varjuküljed. Võrgustatud asjad võivad saada spioneerimise objektideks. Samas saab seda tehnoloogiat kasutada ka vastuluureks ja militaarsetel eesmärkidel.

Maailma juhtivad riigid nagu UK ja USA on oma riiklike informatsiooniliste infrastruktuurilahenduste alal üle minemas pilvearvutuse infrastruktuurile ja linkandmete kasutamisele info integreerimiseks. Vastavad strateegilised dokumendid anti välja põhiliselt 2012. aastal. Tulemused on juba praegu näha.

Uusimad trendid USA-s ja EL-is on aju-uuringutel põhinevate biokiipide loomine ja nende ühendamine võrkudeks, mille produktiivset kasutuselevõttu ennustab Gartner 10 aasta pärast. Järgnevad aastad on ilmselt murrangulised vaadeldud tõusvate tehnoloogiate osas, muutes oluliselt maailma meie ümber ja sees.

## *Linkandmete kasutuselevõtmine*

### *Linkandmete kasutamise eesmärgid ja tulemused*

Linkandmete tehnoloogiad võimaldavad seostada omavahel isoleeritud andmeid ja seeläbi luua uusi teenuseid ja teadmisi ning edendada innovatsiooni. Näiteks on uuringutes toodud positiivsete võimalustena välja valitsemise läbipaistvus ja efektiivsus, innovatsioon, 21. sajandi infrastruktuur. Valitsustele annavad linkandmete tehnoloogiad võimaluse nutikamate ja efektiivsemate avaliku sektori teenuste ja rakenduste väljatöötamiseks. Seega **linkandmete tehnoloogiate kasutuselevõttu tuleb soodustada**.

Linkandmete kasutamise eesmärkide ja tulemuste hulka kuuluvad riigi ja kodaniku vahelise suhtluse paranemine (nt LEO, Läbipaistev Eesti Omavalitsus); „nähtamatud“ teenused, millele inimesele on seadusandlik õigus ning mida ta peab saama sõltumatult sellest, kas ta tellib teenust või mitte (nt lapse riikliku sünnitoetuse teenus või riiklik vanaduspension); andmete kiirem analüüs (valedetele otsustele on võimalik kiiremini reageerida ning uusi otsuseid langetada); monitooringu võimaldamine elutähtsate

teenuste osutajate jaoks (lubab saada oluliselt kiiremini infot, kui elutähtis teenus on häiritud); avaliku sektori kulude vähenemine, sh seoses teabenõuete vastamise vajaduse vähenemisega ja kodanike kaasamisega (nt tänavaaukude kaardistamine mobiilirakendusega); elanikkonna elukvaliteedi tõstmine ja tervishoiu valdkonna kulude vähendamine (nt geenivaramu andmete sidumisel e-Tervise andmetega).

Linkandmete ning nendega seotud tehnoloogiate temaatika tuleks ajakohastatult ja täiendatult lülitada riigi infosüsteemi koosvõime raamistikku ja sellega seotud dokumentidesse. Arvestada tuleb, et linkandmete puhul on tegemist on tüüpilise innovatsiooniprotsessiga, mille puhul uue tehnoloogia kuluefektiivne rakendamine eeldab eelnevat mahukat investeerimist toimiva taristu loomisesse, linkandmete eelnevasse publitseerimisse, kvalifitseeritud arendajate koolitamisse jne.

## *Olulisemad trendid, Eesti võimekused ja võimalused*

Käesolevas uuringus läbi viidud intervjuudes peetakse riigi infosüsteemide seisukohalt oluliseks peamiselt linkandmete ja semantilise veebi tehnoloogiaid. Neid kahte tehnoloogiat nimetatakse sageli paaris, kuna semantika loob linkandmete kasutuselevõtuks vajalikud eeldused. Samuti on nendes tehnoloogiates Eestil täna esialgne võimekus olemas ning arenguvõimalused väga head, kui arvestada olemasolevat riigi infosüsteemi arhitektuuri, läbiviidud avaandmete tootmise pilootprojekte ning andmete ja teenuste semantilise metainfo esitamise juba olemasolevaid kogemusi. Linkandmete kasutamise eelisena tuuakse välja riigi paindlikumat reageerimist muutustele ning kvaliteetsemat poliitikakujundamist. Informatsiooni kättesaadavuses ja kasutamises ei nähta konkurentsieelist mitte ainult erasektorile, vaid ka riigile. Asjade interneti, suurandmete ja koosloome osas arvamused hajuvad, kuna kogemus ja teadmised nende tehnoloogiate kasutamisel ei ole ekspertidel sarnased. Kokkuvõttes, kõige tähtsamaks linkandmete kasutuselevõttu peetakse andmete avamist ja semantiliste tehnoloogiatele kasutuselevõttu, milleta ei ole linkandmete kasutuselevõtt võimalik. Samuti tundub avaandmete osas olevat mahajäämus suurim võrreldes teiste linkandmetega seotud tehnoloogiatega. Ülejäänud tehnoloogiate järjestuses tekivad eriarvamused ning ei joonistu välja selgeid prioriteete.

## *Aktiivsus rahvusvahelistes initsiatiivides*

Aktiivsus rahvusvahelistes initsiatiivides võib anda uut kogemust, aidata levitada oma tulemusi, parandada riigi mainet ja tuua kaasa lisaressursse. Samas nõuab see omapoolset raha ja kõrge kvalifikatsiooniga inimeste ajaressursside panustamist. Eesti riigi väiksuse ja ressursside piiratuse tõttu tuleks initsiatiive hoolega valida. Ankeetides hinnati Eesti osalemise mõju rahvusvahelistes linkandmete ja avaandmete projektides/algatustes keskmiseks, võrreldes teiste uuritud valdkonna alaste tegevustega.

Eesti peaks olema aktiivne järgmistes linkandmetega seotud initsiatiivides (prioriteetsuse järjekorras): Euroopa Komisjoni initsiatiivid ja programmid (sh strateegiliste dokumentide loomine, ISA programm, Joinup foorum, Horizon raamprogramm, eSENS projekt jt), standardimistegevused, OECD avaliku sektori avaandmete (*Open Government Data*) projekt, Euroopa andmete foorumi konverentside seeria, Euroopa avaliku sektori infotribüün ePSIplatform, avatud valitsemise partnerlusprogramm jne.

Osavõtt Euroopa Komisjoni tegevusest vajab üldist link- ja avaandmete teemalist kompetentsi; osavõtt EL raamprogrammist Horizon 2020 eeldab põhjalikku tehnilist koosvõime, linkandmete ja teadustöö alast kompetentsi; osavõtt standardimisest eeldab väga head kompetentsi standarditavas valdkonnas, teadmisi standardimise protsessist ja üsna palju tööaega. Arvestades Eesti väiksust ning ressursside nappust tuleks linkandmetega seotud tegevus võimalusel ühendada muude tegevustega vastavate initsiatiivide raames.

Eesti erasektori esindajad (näiteks ITL aga ka erafirmad, eriti idufirmad) võiksid astuda ODI toetajaliikmeks (liikmemaks 720 Inglise naela aastas). ODI partnerliikmeks saamine ja olemine on suhteliselt kallis (70000 Inglise naela aastas), mis pole ilmselt jõukohane ja ei vasta ka ootustele.

## Pilootprojektide käivitamine

Linkandmete kasutuselevõtt eeldab järgmisi etappe: andmete avamist (kättesaadavaks tegemist), nende puhastamist ja korrastamist (vajadusel), viimist vähemalt neljanda või viienda järni tasemele, linkimist, taaskasutamist (nt päringute võimaldamist) ja haldamist<sup>1</sup>. Kõigil neil etappidel on vaja lahendada suur hulk uurimisküsimusi ning näiteks EL uus teaduse ja innovatsiooni raamprogramm Horizon 2020 perioodiks 2014-2020. a rahastab aastatel 2014-2015 kokku enam kui poolteise miljardi euro maksumuses ITK alaseid initsiatiive, sealhulgas palju ava- ja linkandmetega seonduvaid teadus- ja arendusprojekte.

Eesti siseriiklikke linkandmete kasutuselevõtu pilootprojekte on soovitatav kavandada nii, et nad arvestaksid Eesti eripära (nt X-tee ja RIHA olemasolu), ei dubleeriks mõjuva vajaduseta rahvusvahelisi suuremahulisi uurimisprojekte, looksid võimaluse kolmandatele osapooltele linkandmete rakenduste loomiseks, tekitaksid lõppkasutajale või üldsusele demonstreeritava kasuliku näitetulemuse („*proof of concept*“), annaksid asjaosalistele uusi kogemusi ja oleksid jätkusuutlikud. Näiteks võiks suhtuda ettevaatlikult linkandmete publitseerimise pilootprojektide algatamisse, mis ei demonstreeri selgelt saadavat kasu ning mis ei kindlusta selle kasu jätkusuutlikkust, sealhulgas loodud linkandmete uuendamist ja haldust.

Üheks pilootprojektide käivitamise eesmärgiks on kogemuste andmine uute tehnoloogiatega töötamisel. Anketeerimise tulemustest selgus, et kogemused avaandmetega puudusid 41% vastajatest, kogemused ülejäänud vaadeldavate tehnoloogiatega – 57-68% vastajatest.

Otstarbekas on käivitada orienteeruvalt kuni kuus tehnilist pilootprojekti alternatiivsete linkandmete tehnoloogiate järeleproovimiseks praktikas erinevate seotud tehnoloogiate võtmes valdkonnaüleste infoühiskonna arengukavas väljatoodud oluliste mõõdikute parendamise võtmes. Ühe pilootprojekti maksumuse suurusjärguks võib pakkuda kuni sada tuhat eurot. Pilootprojektid võiksid olla soovitatavalt tervishoiu, georuumilises, hariduse või transpordi, e-äri ja infrastruktuuri valdkondades, kusjuures väga oluline on vastavate organisatsioonide valmisolek projektide algatamiseks. Pilootprojekte on soovitatav kavandada nii, et nad arvestaksid Eesti eripära, ei dubleeriks mõjuva vajaduseta rahvusvahelisi suuremahulisi uurimisprojekte, looksid võimaluse kolmandatele osapooltele linkandmete rakenduste loomiseks, tekitaksid lõppkasutajale või üldsusele demonstreeritava kasuliku näitetulemuse, annaksid asjaosalistele uusi kogemusi ja oleksid jätkusuutlikud.

Järgmine suurem programm oleks enamkasutatavate andmete põhjal linkandmete tuumiku loomine, kuhu ülejäänud andmeid saab vastavalt vajadusele juurde linkida. See programm peaks algama õiguslike aspektide analüüsi sisaldava pilootprojektiga ning sisaldama mitmeid järjestikuseid projekte oluliste andmekogude linkimiseks. Programmi mahtu saab täpsemalt hinnata peale tehniliste pilootprojektide läbiviimist.

Lisaks ülalmainitud tehnilistele pilootprojektidele tuleks kasutada linkandmete platvormi samaliigiliste andmete erinevatest allikatest kokkukogumiseks, näiteks kohalike omavalitsuste andmete ühendamiseks, üldistades kohalike omavalitsuste raamatupidamise avaandmete rakendust. Selliste projektide maksumus jääb tehniliste pilootprojektide maksumuse ja linkandmete tuumiku programmi maksumuse vahepeale.

Asjade Interneti ja tulevikuinterneti arendamiseks oleks esmajoonel vaja pöörata tähelepanu internetiprotokoll 6. versiooni kasutusele võtmisele Eestis. Riigi poolne initsiatiiv võiks seisneda ülemineku käivitamises IPv6-le avalikus sektoris. Tuleb arvestada, et rahalised ja tööjõu kulutused IPv6-le üleminekuks üle tuhande töötajaga ministeeriumi valitsemisalas võivad olla märkimisväärsed ning need tuleks eelarvesse sisse planeerida või rahastada eraldi projektidena.

## Linkandmete alase võimekuse tõstmine

Võimekuse tõstmine on kompleksne teema. Nii näitab UK kogemus, et linkandmete RDF-kujul kättesaadavaks tegemisest ei piisa nende laialdaseks tarbimiseks rakenduste kaudu. Seetõttu loodi Ühendkuningriikides vahevara linkandmete kättesaadavaks tegemiseks ka JSON ja XML andmekirjelduskeeltes. Lisaks loodi valdkonnaspetsiifilised API-d vastavate andmete kasutamiseks. Samuti

<sup>1</sup> <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6237454>

toob aruanne välja linkandmete tootjate vajaduse konkreetsete eeskirjade ja muustrite järele linkandmete tootmiseks ning vajaduse juhendamise järele linkandmete tootmisel ja haldamisel. Eestis näitavad esialgsed Ehitisregistri linkandmete kogemused RDF-i kasutamise osas nii õppetöös kui erasektori tarbijate poolt, et piisava huvi olemasolu korral ei saa linkandmete tehnilised detailid tarkvaraarendajatele rakenduste loomisel takistuseks. Pigem on küsimus teadlikkuse tõstmises ja ajendi tekitamises linkandmete loomise ja kasutamise motiveerimiseks vahendite piiratust arvestades.

Linkandmete alase võimekuse tõstmiseks tuleks eelkõige paralleelselt käivitada linkandmete teemalised kursused ja koolitused, algatada pilootprojektid ning luua tehnilise infrastruktuuri lahendused. Lisaks tuleks osaleda rahvusvahelistes initsiatiivides, maandada riske ning korrastada õigusruumi.

## *Linkandmete alase teadlikkuse tõstmine*

Tavaelanikud võivad tänu linkandmete kasutamisele saada paremaid teenuseid, kuid nad pole eeldatavalt linkandmete nn toorel kujul kasutajate peamine sihtgrupp. Seda arvestades peaks peamine teavitustöö linkandmete olemuse osas toimuma sihtgruppides, kes loovad, haldavad või kasutavad otseselt linkandmeid (sh andmete omanikud ja rakenduste loojad).

Anketeerimine näitas, et teadlikkuse tõstmine on Eestis oluline – 77-st laekunud täidetud ankeedist vastati kahekümne ühes (27%), et vastajal ei olnud teadmisi ühestki küsitletud tehnoloogiast. Nende seas, kes olid mõnest tehnoloogiast teadlikud, oli jaotus tehnoloogiatega lõikes küllaltki ühtlane – 9-25% vastajatest ei teadnud vastavast tehnoloogiast midagi, 38-53% vastajatest olid üldisel tasemel teadmised, 20-32% vastajaid hindas oma teadmisi keskmiseks.

Linkandmete alase teadlikkuse tõstmiseks tuleks alustada infopoliitika ja selle rakendusplaanide uuendamisest, koosvõime raamistiku täiendamisest ja juhendite koostamisest. Tuleks käivitada esmased pilootprojektid, mis demonstreerivad linkandmete kasulikkust ja mida saab koolitustel näidetena kasutada. Koolituste esimene etapp peaks keskenduma üldistele (orienteeruva mahuga üks päev) ja süvendatud (tehniline lisakoolitus orienteeruva mahuga kaks päeva) koolitustele. Teavitustöö linkandmetega seotud võimalustest ja ohtudest võiks jälgida häid praktikaid, mis on tekkinud ISKE turvameetmete süsteemi või riigi IT koosvõime raamistiku alase teadlikkuse tõstmise kampaaniate käigus.

Lisaks tuleks välja valida andmete linkimiseks suurema potentsiaaliga andmekogud ning aidata ja motiveerida nende omanikke andmete avamisel ja linkandmete ettevalmistamisel. Samuti on suurema mõjuga selliste andmete avamine või linkimine, mis on ka teistes riikides avatud või lingitud.

## *Linkandmetega kaasnevad võimalikud ohud ja riskid*

Ankeetides hinnati kõige suuremateks andmete vale interpreteerimise, privaatsuse vähenemise ja keerukate sõltuvuste ning andmete muutumise ohtusid. Muuhulgas, linkandmete puhul võib piisata isikuandmetest ühe lingi küljes selleks, et tuvastada ülejäänud andmete seotus vastava isikuga; näiliselt ebaolulised isikulised detailid võivad kombineeritult anda isikust põhjaliku pildi; andmete linkimine võib laiendada võimalusi statistilise üksuse kaudseks tuvastamiseks; isikustatud andmed kombineeritult isikustamata andmetega võivad kaasa tuua isikuandmete kaitse nõuete riive; linkandmed võimaldavad seostada isikuandmeid avaliku sektori andmete ja muude allikate (meedia, sotsiaalmeedia jm) kombineerimisel; oht väärarusaamade, põhjendamatu hinnangute, inimesele suunatud vaenulike kampaaniate jms tekkeks on suurem. Seetõttu tuleb isikuandmete kaitseks tagada, et ükski avalik lingitud andmete komplekt ei sisaldaks isikuandmeid.

Täpsustada sellised isikuandmete kaitse aspektist küsitavad olukorrad, kus ühe õigusakti alusel saab isikukoodi või muid isikuandmeid avalikustada, AvTS alusel saab seda teavet taaskasutada ning isikuandmete kaitse seaduse alusel ei kohaldata tulemusele isikuandmete seadusega võimaldatavat kaitset. Realiseerida andmekogusid sätestatavates õigusaktides täielikumalt IKS § 11 ja 12 pakutavad andmesubjekti

poolse isikuandmete avalikustamise võimalused. Täpsustada, millal rakendub isikuandmete kaitse seaduse § 17 (automaatsed otsused).

Lisaks õigusruumi täiendamisele nähtub eelnevalt linkandmeid rakendanud riikide kogemusest, et linkandmete loojad vajavad kvaliteetsete andmete loomiseks ja rämpsandmete osakaalu vähendamiseks konkreetseid eeskirjasid ja juhendamist<sup>2</sup>. Nt Ühendkuningriikide ja Belgia kogemus (üleriigilise ja regionaalsete aadressiregistrite andmete ühendamine), FP7 (LOD) ja teised rahvusvahelised linkandmete projektid annavad hea sisendi parimate praktikate näol kvaliteetsete linkandmete loomise, haldamise (sh versioneerimise) jm eeskirjade jaoks.

Rämpsandmete vähendamiseks on vaja kehtestada andmekogude põhimäärustes ja juhendmaterjalides avalikustatavate andmete kvaliteedi nõuded ja määratleda vastutus andmete uuendamise eest. Publitseerida nende nõuete ja nende vastavuse taseme hinnangud vastavas andmekogus või teha nad kättesaadavaks RIHA kaudu. Lisada avaliku teabe seadusesse sätte, mis selgelt määratleb vastutuse väliste osapoolte töödeldud andmete ja nende osutatud teenuste kvaliteedi eest.

## Vajalikud muudatused õigusruumis

Linkandmete kasutuselevõtt hõlmab endas kokkuleppeid koosvõime neljal tasemel - tehnilisel, semantilisel, organisatoorsel ja õiguslikul tasemel. Muudatused õiguslikul tasemel peavad toetama ka ülejäänud tasemeid ja vastupidi. Näiteks uuring „*Open data: an international comparison of strategies*“<sup>3</sup> toob välja nn suletud valitsemiskultuuri kui ühe barjääridest andmete avalikuks tegemisel – nimelt ametnikul organisatsioonis on lihtsam andmeid mitte avada kui riskida sellest tulenevate sanktsioonidega. Taoliste barjääride mahavõtmine eeldab muutusi nii organisatsioonides kui õigusruumis. Vajalikud muudatused õigusruumis, sealhulgas piirangud ja reguleerimist vajavad teemad, on kokku võetud allpool.

- Viia AvTS ning andmekogude aluseks olevad regulatsioonid vastavusse taaskasutuse printsiipidega ja direktiividega 2003/98/EL ning 2013/37/EL. Kehtestada direktiivi 2013/37/EL järgimiseks vajalikud õigus- ja haldusnormid, muuhulgas täiendades RIHA määrust avaandmete korrastamisega seotud sätetega.
- Vaadata läbi ja vajadusel uuendada riigi andmekogude kohta käivad õigusaktid, kus sätestatakse andmete väljastamise tasumäärad, eriti piirkulusid ületava tasu võtmise osas.
- Sätestada avalikustatava teabe litsentseerimise kord ja litsentsitingimused või litsentsitüübid avaliku teabe seaduses, võimalusel soovitades seejuures kasutada sobivat litsentsi *creative commons* litsentsiperest. Kehtestada andmekogude põhimäärustes ja juhendmaterjalides avalikustatavate andmete kvaliteedi nõuded ning publitseerida nende nõuete ja nende vastavuse taseme hinnangud vastavas andmekogus või teha nad kättesaadavaks RIHA kaudu.
- Sätestada konkreetsete andmekogude kohta käivates õigusaktides, võimalusel ka avaliku teabe seaduses (nt AvTS § 31, § 4 või § 8), millised on õigused päringute tulemusena saadavate andmete publitseerimiseks. Andmete (linkandmetena) publitseerimisel tuleks arvesse võtta riske (sh isikutele, omandile, seadmetele, kriitilisele taristule ja riigi julgeolekule), mis tekivad nende andmete kombineerimisel juba olemasolevate publitseeritud andmetega. Võimalik andmete publitseerimise kitsendamine selliste riskide tõttu peaks siiski olema pigem osaline (enamikes andmekogudes on andmeid, mida saab suurema riskita avalikustada), põhjalikult läbi analüüsitud ning detailselt põhjendatud.

<sup>2</sup> [http://events.linkedata.org/ldow2010/papers/ldow2010\\_paper14.pdf](http://events.linkedata.org/ldow2010/papers/ldow2010_paper14.pdf)

<sup>3</sup> [http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012\\_1.pdf](http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012_1.pdf)



## Ettepanekute elluviimise korraldamine

Kuna linkandmete puhul on tegemist on tüüpilise innovatsiooniprotsessiga, mille puhul uue tehnoloogia kuluefektiivne rakendamine eeldab eelnevat mahukat investeerimist, siis tuleks leida poliitiline tugi linkandmete ja avaandmete tehnoloogiate soodustamiseks. Sellise toe motiveerimiseks võib tugineda rahvusvahelisele praktikale (nt USA-s president Obama poliitiline tugi avaandmetele ja eelmise Ühendkuningriikide peaministri Gordon Browni osalemine andmete avamise poliitikate loomisel) ning avaliku sektori andmete avalikustamisest saadava võimaliku kasu suuruse kaudsetele hinnangutele.

Eraldi struktuuri loomist Eesti valitsusasutuste juurde vaid selleks, et ellu viia Lingitud Eesti suunas viivaid ettepanekuid, ei pea käesolev uuring otstarbekaks, välistamata sellise keskuse võimalikkust laiemate ülesannete täitmiseks. Samas tuleks tähtsustada suunatud strateegiat ja konsensuslikku tegevuskava, viia tegevuskava ametnikeni ning kontrollida selle täitmist, luua valitsuse juurde ava- ja linkandmete töögrupp ning teha rahvusvahelist koostööd ODI-ga.

Andmete avamisel ja lingitaval kujul esitamisel tuleks eristada kolme erinevat tüüpi tegevusi. Spetsiifilised tegevused ava- ja linkandmete loomiseks tuleks läbi viia nende asutuste organisatsiooni ja vahenditega, kes andmeid toodavad ja haldavad, vastavalt õigusaktidele. Eelnevaid tegevusi toetavad ettevõtmised, näiteks koolitused, tuleks finantseerida ja tellida keskselt, nt RIA, RISO või Riigikantselei poolt. Konkreetseid arendusprojekte võib läbi viia nii asutuste vahenditega kui ka keskselt; neid tuleks võimalusel teha laiemate projektide kontekstis, mis viivad asutuste või kodanike jaoks nähtavate kasulike tulemusteni.

Era- ja vabasektori kaasamiseks tuleks kehtestada kõikidele tellitavatele infosüsteemidele koosvõimelisuse, avaandmete ja linkandmete osa kohustuslikkuse ning semantiliste klassifikaatorite ja URI-de olemasolu nõuded, propageerida edukaid linkandmete projekte nii era- kui ka avalikus sektoris ning arendada tarkvara linkandmete esitamiseks ja päringute tegemiseks ning pakkuda Eesti vabasektori ja kodanikkonna kaasamiseks linkandmete alast teavitamist ja koolitamist, huvipakkuvaid rakendusi, suhteliselt lihtsaid tööriistu ja kättesaadavaid linkandmeid. Nende jaotiste, samuti intervjuude ja ankeetide põhjal ei saa prognoosida olulist era- ja vabasektori rahastuse kaasamist Lingitud Eesti suunas viivate ettepanekute elluviimisesse.

## Koosvõime arhitektuuri arendamine

### Linkandmetega seotud avatud standardid

Uuring „Standardiseerimise majanduslik mõju“<sup>4</sup> toob välja, et ligi 25% iga-aastasest Prantsusmaa SKP kasvust tuleneb standardimisest. Uuringus viidatakse ka analoogilistele tulemustele teiste Euroopa tööstusriikide Saksamaa ning Ühendkuningriikide kohta. Lisaks otsestele tuludele suurendab standardimine ettevõtte väärtust, võimaldab kiiremat innovatsiooni, parandab vastavust regulatsioonidega, muudab tulemused rahvusvaheliselt kasutatavateks ning soodustab toodete ja teenuste kvaliteeti. (Rahvusvahelise) koostöö võimaldamine on standardimise üks olulisemaid tulemusi. Näiteks on Eestis kokku 24003 kehtivat standardit, neist vaid 275 ehk 1,1 % on algupärased Eesti standardid (8.01.2014 seisuga). Ülaltoodud põhjustel on investeringud standardite kasutuselevõttu suhteliselt suure tulususega.

Linkandmete ja semantilise veebi standardite kasutuselevõtmine ei tähendaks olulisi muutusi riigi infosüsteemi arhitektuuris – standardid RDF, RDFa, microdata, OWL, SA-WSDL, URI, URL, HTTP jt on riigi infosüsteemi koosvõime raamistik<sup>5</sup> juba olemas ning piisavalt küpsed kasutuselevõtuks. Pigem on neist osade puhul probleemiks nende rakendamine praktikas, nagu ka linkandmete puhul üldiselt. Kuna linkandmete standardite rakendamine ei ole eesmärk omaette, tuleks seda korraldada seoses linkandmetega seotud projektide, koolituste ja muude algatustega. Ankeetide vastustes märgitakse, et avatud ja lingitud

<sup>4</sup> <http://www.iso.org/iso/fr/home/standards/benefitsofstandards/benefits-detail.htm?emid=7>

<sup>5</sup> <http://www.riso.ee/et/koosvoime/raamistik>

andmed ei ole väärtuslikud „asjana iseeneses“ ning mingi standardi eelistamine ei tähenda ilmtingimata rahalist säästu.

## Tehnilise infrastruktuuri lahendused

Linkandmete tootmiseks on mõistlik võimalusel ära kasutada olemasoleva riigi infosüsteemi arhitektuuri poolt pakutavaid võimalusi. Riikliku Ehitisregistri linkandmete pilootprojektis kasutati näiteks X-tee andmeteenuid linkandmete komplekti ehitamiseks. Kuna X-tee kaudu on riigi infosüsteemide andmed kättesaadavad, on X-tee kasutamine linkandmete tootmiseks igati otstarbekas. Selline kasutamine võimaldab paremini korraldada ka linkandmete uuendamist. Samas on X-tee andmeteenuid kasutamiseks linkandmete loomisel vaja koostada andmeteenuid disainimise metoodika ja teha selle jälgimine kohustuslikuks, läbi Riigi infosüsteemi haldussüsteemi määruse, et uute andmeteenuid loomisel tekiks linkandmete loomiseks vajalikud andmeteenuid ning samal ajal pareneks andmeteenuid taaskasutus. Lisaks on vaja luua avatud linkandmete tootmiseks infrastruktuur, kus saaks avaandmeid pakkuvaid X-tee jaoks loodud andmeteenuid rakendada, sh linkandmete tootmiseks, ilma turvaservereid kasutamata.

Tuleks edasi arendada Eesti avaandmete portaali [opendata.riik.ee](https://opendata.riik.ee) ning avatud linkandmete portaali, seostades neid teiste andmeallikatega, nagu Eesti geoportaal, Maa-ameti geoportaal, Keskkonnaagentuuri portaal ning Statistikaameti andmebaas. Kuna paljud avaliku sektori asutuste andmed on portaalis registreerimata, tuleks käivitada projekt Eesti avaliku sektori avaandmete loetelu loomiseks ning publitseerimiseks.

Suurema sünergia saavutamiseks nii linkandmete vallas kui seotud tehnoloogiates on mõistlik tellida ideekonkursi kaudu avatud tehnilise pilveplatvormi (PaaS) loomine ja haldamine linkandmete *kasutamise lihtsustamiseks* tehnoloogilistes pilootides (nt linkandmete loomine ja haldamine koosloome kaudu, linkandmete reaalaajanalüütika infoühiskonna arengukava meetrikate pidevaks mõõtmiseks, valdkonnaühene linkandmete kasutamine e-äri lahendustes, jm) ja mis uuendaks linkandmeid linkandmete portaalist ning oleks avatud linkandmetele, funktsioonidele ja linkandmete rakendustele. Platvorm peab toetama valdkondlike REST teenuste loomist andmete publitseerijate, firmade või kogukonna poolt nii kommertseesmärkidel kui tasuta kasutamiseks. Avatud tehnilise platvormi jätkusuutlikkuse tagamiseks on vaja eralda eelarvelised vahendid selle organisatsiooni ülalpidamiseks.

## Koosloomel põhinevate lahenduste loomine

Riigi seisukohast on koosloome oluline instrument kodanikuühiskonna kujundamisel, kuna see lihtsustab indiviidide panustamist suurte probleemide lahendamisse. Koosloome elluviimiseks on kriitilise tähtsusega mitmeotstarbelise koosloome platvormi (sh liitunud indiviidid) loomine, mis lihtsustab isikute kaasamist erinevatesse initsiatiividesse. Sellised platvormid eeldavad ühiste mõistete ja probleemihalduse kasutamist, mida saab korraldada nii linkandmete kasutamiseks kui ka ilma selleta.

Riik saab läbi viia pilootprojektid koosloome platvormi ja selle näidisrakenduste loomiseks kas siis üldiste konkurssidena või konkreetsete teenuste moderniseerimise eesmärgil (nt seotult julgeoleku või keskkonna seirega). Selliste pilootprojektide raames tuleks üheks eesmärgiks seada sellise platvormi loomine, mida saaks kasutada ka väljaspool avaliku sektori teenuste pakkumist. See lihtsustaks tunduvalt uute koosloome lahenduste loomist.

Selliste konkursside algatamisel tuleks tugineda vähemalt ühele sisukale rakendusele, kus linkandmete kasutamine annaks koosloome protsessile lisaväärtust. Kuna koosloome eelduseks olevate ühiste mõistete ja probleemihalduse kasutamist saab paljudel juhtudel realiseerida ka ilma linkandmeteta, pole sellise rakenduse leidmine triviaalne. Eesti valitsus võiks koondada ava- ja linkandmete kasutajate grupid, kes kasutades koosloome platvormi annaksid riigile nõu, milliseid andmeid avada ja linkida.

## Vajalikud semantikaressursid

Linkandmete kasutamiseks tuleks olemasolevaid semantilisi varasid täiendada, ühtlustada ja/või kooskõlastada, eriti valdkonnaüleste rakenduste puhul. Lisaks olemasolevate semantiliste varade täiendamisele vastavalt kehtivale semantilise koosvõime raamistikule tuleb kaasajastada semantilise koosvõime raamistiku järgmist versiooni linkandmete ja seotud tehnoloogiate vaatevinklist.

Esmajoones tuleks keskenduda juriidilisi ja füüsilisi isikuid, aadresse, kinnistuid ning ruumiandmeid kirjeldavate ontoloogiate loomisele, sest need andmed on esindatud enim riigi infosüsteemides, neil on suurim ühisosa riigi infosüsteemide lõikes ja nad on eelduseks andmete linkimisele. Seoses linkandmete kasutamisega peaks nende üleilmse lingitavuse parandamiseks võtma kasutusele nii palju standardontoloogiaid kui võimalik. Ka Eesti riigi infosüsteemi ontoloogiaid tuleks linkandmete loomise aspektist üle vaadata, eriti just avalike linkandmete loomise nõuetest lähtuvalt. Lisaks tuleks luua rohkem valdkonnaüleste ontoloogiaid.

Riigi infosüsteemide semantilise koosvõime raamistikus on viimaste esitamiseks ette nähtud valdkonnasõnastikud, klassifikaatorid, ontoloogiaid, nendega seotud veebiteenuste semantilised kirjeldused ja muud semantilised varad, mis tuleb publitseerida RIHA-s. Kuigi RIHA määruses on öeldud, et kõik uued infosüsteemid tuleb varustada semantiliste kirjeldustega, pole selge, kas see on nii läinud ja kas toimub vastav järelevalve. Seda küsimust tuleks eraldi uurida näiteks auditi raames.

## Linkandmed avaandmete publitseerimisel

Kuna linkandmete üks põhilisi ideid on võimaluse andmine kolmandatele osapooltele uute teenuste ja rakenduste loomiseks, siis suur osa linkandmete tehnoloogiate kasutuselevõtu raskusest langeb kolmandatele osapooltele. Ankeetides hinnati, et **avaandmete rakenduste loomine kuulub projektide rühma, millel on Eestile kõige suurem mõju.**

Registrites, infosüsteemides ja andmekogudes (sh andmete avaandmetena publitseerimisel) tuleks linkandmete tehnoloogiate kasutusele võtmiseks jälgida W3C linkandmete publitseerimise parimaid praktikaid: modelleerida andmed, nimetada asju püsivate URI-dega, taaskasutada sõnastikke niipalju kui võimalik, avaldada avaandmete kohta nii inim- kui masinloetavad kirjeldused (metaandmed), teisendada andmed RDF-i, määrata sobiv litsents, avalikustada avaandmed linkandmetena ja kuulutada nad välja. Mitteavalike andmete puhul on mõistlik eelkõige lingitavana hoida see osa andmetest, mille juures on näha vajadust kombineerida neid andmetega välistest allikatest.

Andmed tuleks publitseerida vastavalt kehtivale õigusaktidele, masinloetaval kujul ja avatud litsentsiga, sätestades vastutused andmete kvaliteedi ning nende haldamise osas, järgides riigi infosüsteemi koosvõime raamistikus toodud standardeid andmete ja metaandmete publitseerimiseks, kasutades valitud linkandmete kasutuselevõttu toetavaid tehnilise infrastruktuuri lahendusi ning soovituslikult pakkudes andmete kasutajale andmeid harjumuslikus andmekirjelduskeeles üle kasutaja pöördumise jaoks suhteliselt lihtsa vahekihi, kasutades sobivaid semantikaressursse andmete metamudelite kirjeldamiseks vastavalt semantilise koosvõime raamistikule, vajadusel olemasolevaid kirjeldusi täiendades ja korrastades.

## Sisutöötlus veebisaitides ja -liidestest

Linkandmete kasutuselevõtt veebisaitides ja -liidestest peaks algama uute vajaduste selgitamisest, mida linkandmete kasutuselevõtt rahuldab, eelistest, mida linkandmed seejuures pakuvad ning linkandmete kasutuselevõtu tulu-kulu analüüsist. Eriti tuleks olemasolevate veebisaitide puhul hinnata, kas nende ümberkorraldamise ja linkandmete kasutuselevõtu positiivsed efektid ületavad tehtavaid kulutusi.

Linkandmete kasutuselevõtt veebisaitides ja -liidestest peaks algama RDFa, mikroandmete või *microformats* märgendite lisamisest HTML dokumentidele. Konkreetne märgendite keel ja sõnavara valitakse pragmaatilistest kaalutlustest lähtuvalt. Avaliku sektori veebides on esmajärjekorras oluline märgendada

organisatsiooni töötajate kontaktid ja sündmused. Suuremate organisatsioonide puhul on mõistlik selleks laiendada olemasoleva sisuhaldussüsteemi mooduleid, mis töötajate kontakte ja uudiseid veebi väljastavad, selliselt, et väljastatud andmetele lisatakse märgendid automaatselt. KOVide puhul annaks suurima mõju KOVTP lahendusele analoogse toe lisamine. Otstarbekas on kasutada esitusi, mille puhul mitmete valdkondade jaoks on loodud vastavad metamärgendite komplektid, nt mikroandmete jaoks loodud <http://schema.org>, <http://purl.org/goodrelations> jt. Linkandmete kasutuselevõtul tuleks joondada sisutöötlus LOD2 Stack poolt pakutud linkandmete elutsükliga ning järgida veebide koosvõime raamistiku soovitusi ja W3C linkandmete publitseerimise parimaid praktikaid. Üheks linkandmete kasutamise rakenduseks sisutöötluse vaatevinklist oleks KOV-ide ja asutuste veebidest sisu automaatne kokkukogumine ja esitamine suuremates portaalides, näiteks eesti.ee.

Kirjeldatud nõuete ja tehnoloogiate rakendamiseks tuleks olemasolevate veebisaitide puhul käivitada linkandmete kasutuselevõttu toetavad projektid, mis võivad olla iseseisvad või seotud linkandmete kasutuselevõtu pilootprojektidega. Uute veebisaitide projektide hangetesse tuleks lülitada linkandmete kasutuselevõttu toetavad sisutöötluse nõuded, muuhulgas seoses metaandmete lisamise, semantilise kirjeldamise ja andmete agregeerimisega.

## Organisatsioonilise koosvõime arendamine

### Linkandmete tehnoloogiate monitoorimine, õpetamine ja propageerimine

Linkandmete, semantilise veebi jt käesolevas uuringus käsitletud tehnoloogiate temaatikat on seni Eesti ülikoolides käsitletud väheste kursuste ja tudengiprojektide raames õppejõudude initsiatiivil. Ülikoolid peaksid seda initsiatiivi võimendama läbi temaatiliste õppe-, koolitus- ja teadusprojektide, mille raames kaasatakse nii eksperte välismaalt kui kohalikke tudengeid. Ka on vaja ülikoolide õppekavasid uute trendide valguses regulaarselt kaasajastada.

Koolituste esimene etapp peaks keskenduma üldistele (orienteeruva mahuga üks päev, koolitatavate suurusjärg viissada inimest ehk orienteeruvalt 20 koolituspäeva) ja süvendatud (tehniline konkreetse probleemi linkandmetega lahendamise lisakoolitus orienteeruva mahuga kaks päeva - suuremate andmekogude haldajatele, kes avaldavad oma andmeid ava- ja linkandmetena, orienteeruvalt 200 inimest ehk kokku 20 süvendatud koolituse päeva) koolitustele.

Erinevad uuringud on üksmeelel selle kohta, et ava- ja linkandmetel on positiivsed mõjud – näiteks, võimaldades pakkuda paremaid teenuseid, luues paremat investeerimiskliimat ja vähendades korrupsiooni. Riik peaks selliseid mõjusid toetama, isegi kui täpsed kulude/tulude uuringud ei ole kättesaadavad. Kindlasti õigustavad pakutud hinnangud ülaltoodud koolituste läbiviimist – nende orienteeruv esimese etapi kogumaht jääb suurusjärku 50 000 eurot aastas.

### Linkandmed Eesti IKT sektori lahendustes

Intervjuudest nähtub, et Eesti IKT sektori võimalusena nähakse eelkõige teenuste loomist. Innovatsiooni väga ei usaldada ning kõige olulisemate tegevustena tuuakse välja riigi poolset eestvedamist ning vastutuse võtmist. Erasektori kaasamiseks tuleks riigihangetele lisada ava- ja linkandmeid käsitlevad reeglid, mis on kohustuslikud iga infosüsteemi tellimisel. Ka ankeedivastustes hinnati avaandmete rakenduste mõju (näiteks teenustes) kõrgelt.

Samal ajal näitavad globaalsed trendid kasvavat turgu (suur)andmete analüüsimise lahendustele. Globaalsed trendid viitavad märkimisväärsele ja kasvavale puudusele andmete töötlemise ja analüüsimise oskustega tööjõu osas lähima 5 aasta perspektiivis. Sellest tulenevalt on suurenenud (suur)andmeid kasutavate analüütiliste tööriistade ja SaaS lahenduste tootmine nii globaalses mastaabis. Lingitud ja puhastatud

andmete kasutamine lihtsustab taoliste lahenduste loomist. Veelgi enam, kasutades globaalselt aktsepteeritud metamudeleid linkandmetes muutub lihtsamaks nende lahenduste eksportimine teistele turgudele. Eesti IKT sektor saab tõsta oma ekspordivõimekust linkandmetel põhinevat tarkvara tootes ning avalik sektor saab aidata sellele kaasa nõudes enda poolt tellitud või toetatud toodetes järjepidevalt linkandmete kasutamist.

Linkandmete kasutuselevõtu suurendamisele aitavad kaasa temaatilised teadusarendus-, tootearendus- ja innovatsioonitoetused ja Garage48-tüüpi ettevõtmised. Need võivad aidata kaasa olukorra muutumisele Eesti IKT sektoris, kus on suhteliselt vähe sektorisisest tootearendust ning riskide maandamiseks tehakse üldiselt tööd klientide tellimisel. Temaatilised rakendusuuringud ja tootearendusprogrammid võiksid tagada tootearendusele suunatud projektide jätkusuutlikkuse. Samal ajal saab avalik sektor tellijana suunata IKT sektorit linkandmetesse rohkem panustama, nõudes uutes või jätkuprojektides mõistlikkuse piires linkandmete kasutamist. Erasektor saab tellijana suunata IKT sektorit linkandmeid kasutama, kui avalikult sektorilt vajaminevad andmed on linkandmetena kättesaadavad.

Lisaks tuleks võtta kasutusele meetmed avatud, valdkondlike linkandmete PaaS platvormide loomise ja tööhoidmise või valdkondlike/valdkonnaüleste avatud teenuste olemasolevale avatud linkandmete pilveplatvormile loomise motiveerimiseks (sealhulgas e-äri jaoks nt CRMid, finantsanalüüsi, turundusanalüütika ja toote planeerimise lahendused, mis kasutavad linkandmeid).

Vastavalt avaliku teabe seadusele tuleb avalikustamisele kuuluv teave avalikustada ajakohasena ning viisil ja vormis, mis võimaldab selle allalaadimist masinloetaval kujul. Seepärast ei too täiendav linkandmetena publitseerimise nõue, kui see lisatakse kohe projekti algatamise faasis, kaasa väga suuri lisakulutusi. Innovatsioonitoetuste ja Garage48-tüüpi meetmete maht sõltub pakutavatest toetust vajavatest projektidest ning seda saab hoida valitud piires, näiteks samas suurusjärgus koolituste ja teavitamise mahuga. Arvestades eelmises jaotises toodud linkandmete kasude hinnanguid võib ülalpakutud meetmeid seega hinnata kuluefektiivseiks.

## *Avaliku sektori ülesanded infosüsteemide tellijana*

Intervjuudes rõhutati, et avalik sektor peaks kehtestama kõikidele tellitavatele infosüsteemidele koosvõimelisuse, avaandmete ja linkandmete osa kohustuslikkuse ning semantiliste klassifikaatorite ja URId e olemasolu nõuded. Nõuete järelevalve peaks struktuurifondidest rahastatavate projektide puhul olema seotud rahade väljamaksmisega. Muude projektide üle võiks järelevalvet korraldada RISO, kes kooskõlastaks projekti rahade väljamaksmist sõltuvalt nõuete täidetusest.

Eriti oluline on ontoloogiade loomine ja uute linkandmete tekitamine, sest linkandmete väärtus kasvab hüppeliselt võrreldes lingitud andmete mahu kasvuga. Seetõttu võiks iga uus projekt, kus andmeid kasutatakse ja/või uusi andmeid luuakse, lisada uusi linkandmeid läbi tekkinud andmete linkimise. Ontoloogiade loomisele on kaasa aidanud tellijapoolsed nõuded täitjatele luua uute infosüsteemide jaoks ka sobivad ontoloogiad. Lisaks on veel teisigi näiteid, nt ehitussektoris, kus tellija nõuete kaudu saab suunata valdkonna arengut.

Avalik sektor peaks suurendama analüütiliste rakenduste osakaalu, kus kasutatakse andmeid ka väljastpoolt nende endi haldusala ja kasu linkandmetest on otseselt näha. Taoliste lahenduste tekkimise soodustamiseks on mõistlik infosüsteemide lähteülesannete püstitustes eraldi välja tuua väljaspool haldusala paiknevate andmete võimalik mõju rakendusele valdkondlike andmekomplektide ja avaandmete kasutamise lõikes. Sel juhul saab infosüsteemide loomise finantseerimise otsuseid tehes arvestada nende valdkondadeülest mõju.

## *Soovitusi erasektorile infosüsteemide arendamisel*

Intervjuude ja ankeetide põhjal ei ole erasektori esindajad oma rolli hindamisel eriti optimistlikud. Ei usuta, et uued tehnoloogiad toovad suurt edu – innovatsiooni on tehtud, kuid see on enne pooleli jäetud, kui tulemus tekib. Kõige olulisemate tegevustena tuuakse välja riigi poolset eestvedamist ning vastutuse võtmist.

Samuti nimetatakse edulugude vajadust, mis võimaldaks näidata innovatiivset probleemilahendust ning mida Euroopasse skaleerida. Erasektori rolli nähakse eelkõige teenuste loomises, taandades avalik sektor teenuseosutaja rollist. Ka ankeetides on turunduse ja müügi, erasektorile tüüpiliste avaandmete kasutusvõimaluste mõju hinnatud keskmisest madalamana.

Samas on avatud standardite kasutamine ning standardiseerimises osalemine arenenud tööstusriikides andnud olulise osa SKP kasvust. Erasektorile võiks soovitada linkandmete käsitlemist avatud standardina ning nende positiivse mõju arvestamist sarnaselt valdkondlikele metamudelitele, nagu nt SID, HR-XML jne. Tuleks olla kursis linkandmete tehnoloogia arenguga, osaleda koolitustel, jälgida linkandmete edulugusid ja kasutada avaliku sektori poolt publitseeritavaid linkandmeid. Eestis on näiteid linkandmete kasutuselevõttust, näiteks on Riikliku Ehitisregistri lingitud kujul andmed leidnud kasutust kinnisvaraga seotud portaalides.

Kuna erasektori esindajad ei ole oma rolli hindamisel linkandmete kasutuselevõttus eriti optimistlikud, tuleks soovitada riigi poolset eestvedamist ning vastutuse võtmist ning edulugude esiletoomist. Erasektorile võiks soovitada linkandmete käsitlemist avatud standardina ning nende positiivse mõju arvestamist. Tuleks olla kursis linkandmete tehnoloogia arenguga, osaleda koolitustel, jälgida linkandmete edulugusid ja kasutada avaliku sektori poolt publitseeritavaid linkandmeid. Erasektor saab linkandmete tehnoloogiaid oma infosüsteemide tellimisel ja arendamisel kasutada, propageerides edukaid linkandmete projekte nii era- kui ka avalikus sektoris, arendades tarkvara linkandmete esitamiseks ja päringute tegemiseks, pakkudes avaliku sektori andmete publitseerimise ja linkimise teenust ning näidislahendusi, pakkudes ekspordipotentsiaaliga linkandmetel põhinevaid andmeanalüüsi ja pilveteenuseid, kasutades ava- ja linkandmeid teenuste pakkumiseks, rakendades linkandmete tehnoloogiat oma ettevõtte andmete haldamisel (linkandmed ei pruugi olla avaandmed), rakendades linkandmete tehnoloogiat firmade vaheliseks andmete vahetamiseks ning kasutades võimalusi, mida pakuvad innovatsioonitoetused ja Garage48-tüüpi ettevõtmised.

## Vabasektor ja kodanikkonna kaasamine

Vabasektor ja kodanikkond on üle maailma olnud linkandmete alal üsna aktiivne. Näiteks on DBpedia Live koosloome initsiatiiv, mille eesmärk on teha Wikipedia andmed linkandmetena kättesaadavateks ja kasutatavateks. Eestis on koosloome läbi vastavate rakenduste leidnud kasutuse isikute poolt linkandmete loomisesse. Näiteks lisatakse veebilehtedele juba praegu märgendeid (nt *microdata* kujul), mida otsimootorid kasutavad linkimiseks, tehes samas saadud andmed otsitulemustes kättesaadavaks.

Üldiselt saab vabasektori ja eraisikute huvi (link)andmete vastu tulla nende seesmistest ajenditest, näiteks soovist teha midagi paremaks, isiklikust huvist, soovist saada tuntuks vms. Linkandmete tehnoloogia on suhteliselt keeruline ja kompetentsi puudumine avaldab siinkohal negatiivset mõju. Inimesed, kes seda tehnoloogiat hästi valdavad, on suure tõenäosusega juba mingis organisatsioonis hõivatud ja tegelevad sellega oma tööülesannete raames. Lisaks on vabasektori ja kodanikkonna jaoks olemas ka palju muid huvitavaid teemasid.

Eelnevast ning intervjuudest nähtub, et Eesti vabasektori ja kodanikkonna kaasamiseks oleks vaja linkandmete alast teavitamist ja koolitamist, huvipakkuvaid rakendusi, suhteliselt lihtsaid tööriistu ja kättesaadavaid linkandmeid. Teavitamist, koolitust ja linkandmete loomist tuleks teha avaliku sektori tööülesannete täitmise ning eelmistes jaotistes toodud algatuste raames, näiteks eraldades koolitustele teatud protsendi kohti vabasektori ja kodanikkonna jaoks. Sellise eraldamise mõistlik maht selgub koolituste käigus ning sellest lähtudes võib arvestada vajalikke lisaressursse.

Lisaks on vaja disainida meetmed selliste valdkondlike ja valdkonnaüleste avatud linkandmete ökosüsteemide loomise või kohandamise võimaldamiseks, kus on toetatud nii avatud kui privaatsete/konfidentsiaalsete linkandmete kooskasutamine ilma andmete privaatsust/konfidentsiaalsust ohustamata. Vastavad ökosüsteemid koondaks erinevaid linkandmete komplekte valdkondlike või valdkonnaüleste probleemide lahendamiseks.

## Tööjaotus eri huvipoolte vahel

Kirjanduse analüüs, nt TNO avaandmete strateegiate uuring<sup>6</sup> ning käesolevas uuringus läbi viidud intervjuud toovad välja avaliku sektori juhtiva rolli ava- ja linkandmete kasutuselevõtus. See on ka arusaadav, sest erasektor näeb andmeid kui oma strateegilisi varasid ning on väga ettevaatlik nende avalikustamisega, vabasektor ja kodanikkond omavad aga suhteliselt vähe andmeid. Samuti on näha tugeva poliitilise juhtfiguuri positiivne mõju. Ka see on mõistetav, sest avaliku sektori töötajad ei pruugi iseenesest olla motiveeritud andmete avalikustamiseks.

Toodud üldine rollijaotus kehtib ka Eestis. Avalik sektor on enamasti initsiaator, sätestab õigusakte, soovib või kehtestab standardeid, publitseerib andmeid, arendab esmaseid teenuseid, tellib pilootprojekte, koolitab ja teavitab teisi osapooli. Erasektor, vabasektor ning kodanikkond on pigem järgijate, teenusepakujate ja andmete kasutajate rollis. Toodud üldisest rollijaotusest saab tuletada ka järgneva detailsema tööjaotuse eri huvipoolte vahel Eestis.

Täpsemalt, RISO haldab ja arendab riigi infosüsteemi semantilise koosvõimega seotud dokumentatsiooni, semantika töörühma materjale, semantilise koosvõime raamistikku ning muid riigi infosüsteemi strateegilisi dokumente<sup>7</sup>. Riigi Infosüsteemi Amet koordineerib riigi infosüsteemi arendamist ja haldamist, teostab järelevalvet riigi infosüsteemi haldamist reguleerivatest õigusaktidest tulenevate nõuete täitmise üle ja peab riigi infosüsteemi haldussüsteemi. RISO ja RIA koostöös koordineeritakse metaandmete, sealhulgas valdkondade vaheliste metaandmete loomist, kasutades selleks valdkondade vahelisi töörühmi ning vajadusel hankides uuringuid või pilootprojekte. MKM infoühiskonna teenuste arendamise osakond (ITAO) koordineerib linkandmete kasutamise aspekte, mis on seotud avalike teenuste ning avaliku sektori asjaajamise ja elektroonilise dokumendihalduse arendamisega.

Avaliku sektori (täpsemalt, riigi, kohaliku omavalitsuse või muu avalik-õigusliku juriidilise isiku või avalikke ülesandeid täitva eraõigusliku isiku) andmekogude vastutavad või volitatud töötajad registreerivad andmekogud ning loovad ja registreerivad vastavad semantilised varad riigi infosüsteemi haldussüsteemis. Samuti vastutavad avaliku sektori asutused oma infosüsteemide jaoks vajalike semantikavarade väljatöötamise ning andmete publitseerimise eest vastavalt avaliku teabe seadusele, Vabariigi Valitsuse määrusele „Riigi infosüsteemi haldussüsteem“ ja muudele õigusaktidele.

Erasektor, vabasektor ning kodanikkond hoiab ennast ava- ja linkandmete valdkonnaga kursis, leiab kasulikke, tulusaid ja/või huvitavaid rakendusi, pakub teenuseid ja kasutab andmeid. Avaliku ja erasektori koostööd tuleb igati soodustada, kuid riigihangete seadusele vastavaid PPP projekte linkandmete valdkonnas võib lugeda Eestis vähetõenäolisteks ning suure mahu ja kestvuse tõttu peaksid nendele eelnema koolitused, pilootprojektid ja muud eelmistes jaotistes pakutud tegevused.

Toodud loomulik tööjaotus taaskasutab Eesti juba olemasolevaid organisatsiooni ja protsesse ning ei too kaasa täiendavaid kulutusi lisaks juba eelpool mainitutele.

## Andmetöötlusprotsessid ja andmete täieliku elutsükli toetamine

Üldjuhul on linkandmete loomine mitte põhieesmärk, vaid vahend põhieesmärkide saavutamiseks. Andmeid kogutakse ja hallatakse rahuldamiseks mingit vajadust ning saamaks mingit sisulist tulemust. Vajadusel ja võimalusel publitseeritakse need andmed ava- või linkandmetena. Seega peaks kasutatavate andmetöötlusprotsesside valik eelkõige lähtuma andmekogu loomise eesmärkidest. Valitud protsesse täiendatakse linkandmete loomiseks ja kasutamiseks vajalike tegevustega.

Linkandmete andmetöötlusprotsesside eripäraks andmete elutsükli mõttes on andmete semantiline hajusus. Seetõttu on avaliku sektori linkandmete jaoks pakutud erinevaid elutsükli mudelid. Need pakuvad erinevat elutsükli ulatust (nt kas mudel käsitleb vaid andmeid või ka äriprotsesse), etappide valikut ning etappide

<sup>6</sup> [http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012\\_1.pdf](http://www.epractice.eu/files/European%20Journal%20epractice%20Volume%2012_1.pdf)

<sup>7</sup> <http://www.riso.ee/et/koosvoime/semantika>

organisatsiooni (nt järjestikune või iteratiivne). Ühte kõikide poolt tunnustatud linkandmete elutsükli mudelit ei ole. Olukord on analoogiline tarkvara arendusega, kus kasutatakse sõltuvalt vajadusest mitmesuguseid elutsükli mudeleid.

Käesoleva töö autorid soovivad Eesti riigi infosüsteemis kasutada linkandmete tootmiseks ja publitseerimiseks iteratiivset elutsükli mudelit, mis hõlmab järgmisi etappe: spetsifitseerimine, modelleerimine, genereerimine, publitseerimine ja kasutamine. See on heas kooskõlas nii TNO avaandmete elutsükli mudeliga kui ka W3C soovitustega.