

# Avatud teadus Eestis ja Euroopas: võimalikkus ja potentsiaal erinevate sidusgruppide vaates

RITA 4: TAI poliitika seire

## Lõpparuanne

Tallinna Ülikool

Kristel Toom, Arko Olesk, Raivo Ruusalepp, Esta Kaal, Sigrid Mandre

Eesti Teaduste Akadeemia

Rein Vaikmäe

Uuringu tellis SA Eesti Teadusagentuur.

Tööd rahastatakse „Valdkondliku teadus- ja arendustegevuse tugevdamise (RITA)“ tegevuse 4 raames Euroopa Regionaalarengu Fondist.

Uuringu tegi Tallinna Ülikool ja Eesti Teaduste Akadeemia.

Uuringu autorid: Kristel Toom  
Arko Olesk  
Raivo Ruusalepp  
Esta Kaal  
Sigrid Mandre  
Rein Vaikmäe

Viitamine: Toom, K., Olesk, A., Ruusalepp, R., Kaal, E., Mandre, S., Vaikmäe, R.(2017). Avatud teadus Eestis ja Euroopas: võimalikkus ja potentsiaal erinevate sidusgruppide vaates. Tallinn: Tallinna Ülikool ja Eesti Teaduste Akadeemia.

RITA on Euroopa Regionaalarengu Fondist toetatav programm, mille eesmärk on suurendada riigi rolli teaduse strateegilisel suunamisel ning teadus- ja arendusasutuste võimekust ühiskondlikult oluliste uuringute läbiviimisel. Programmi kaudu rahastab SA Eesti Teadusagentuur Eesti riigi vajadustest lähtuvaid sotsiaal-majanduslike eesmärkidega rakendusuringuid.

Tegevus 4: Teadus- ja arendustegevuse ning innovatsiooni (TAI) poliitika seire. RITA tegevus 4 eesmärk on jälgida TAI poliitika elluviimist ning anda soovitusi uute poliitikate kujundamiseks. Analüüse ja uuringuid teevad Tallinna Ülikool, Tartu Ülikool, Tallinna Tehnikaülikool, Teaduste Akadeemia ning Eesti Teadusagentuur.

## Sisukord

Sissejuhatus .....	5
Kokkuvõte.....	8
Summary .....	12
1 Dokumendianalüüs.....	17
1.1 Euroopa Liidu regulatsioonid.....	17
1.2 Valdkondlikud uuringud .....	18
1.3 Teiste riikide praktikad .....	24
1.3.1 Suurbritannia .....	25
1.3.2 Holland .....	26
1.3.3 Taani .....	28
1.3.4 Soome.....	29
1.3.5 Sloveenia .....	30
1.3.6 Iirimaa.....	31
1.3.7 Poola.....	33
1.3.8 Ungari .....	33
Kokkuvõte.....	34
2 Veebipõhine küsitlusuuring ja fookusgruppiintervjuud .....	35
2.1 Metoodika .....	35
2.1.1 Veebiküsitlus .....	35
2.1.2 Individuaal- ja fookusgrupp-intervjuud.....	36
2.2 Uuringu tulemused.....	37
2.2.1 Avatud juurdepääs teadusartiklitele .....	38
2.2.2 Teaduse avaandmed.....	43
2.2.3 Avatud teaduse võimalikud mõjud.....	49
2.2.4 Kokkuvõte huvirühmade kaupa.....	52
3 Avatud teadusandmete säilitamise maksumust mõjutavad tegurid .....	54
3.1 Teadusandmete säilitamise taristu maksumust mõjutavad olulisemad komponendid .....	55
3.1.1 Milliseid andmeid avaldada ja arhiveerida?.....	55
3.1.2 Andmete kvaliteet avaldamiseks ja arhiveerimiseks.....	58
3.1.3 Kus andmeid avaldada ja arhiveerida?.....	59
3.1.4 Kuidas tagada teadlastele tugi ja nõustamine?.....	59
3.2 Vabakasutus-kirjastamisega seotud kulud .....	61
3.3 Teadusandmete haldamise ja arhiveerimisega seotud kulud .....	64
4 Teadusandmete säilitamise korraldamine: stsenaariumid .....	69

4.1	Teenuse komponendid.....	69
4.1.1	Teenuse korraldajate ja kasutajate vaade.....	69
4.1.2	Andmete vaade .....	70
4.1.3	Teenuste ja tööprotsesside vaade.....	71
4.1.4	Tehnoloogiate vaade .....	72
4.1.5	Regulatsioonide ja standardite vaade .....	73
4.2	Teenuse organisatsiooni stsenaariumid.....	74
4.2.1	Institutsionaalsete andmehoidlate võrgustik.....	74
4.2.2	Riiklik andmearhiiv .....	78
4.2.3	Rahvusvahelised teenuste platvormid .....	84
4.3	Visioon avatud teaduse taristu teenusest.....	87
5	Teekaart.....	91
	Viidatud kirjandus.....	93

## Sissejuhatus

Avatud teadus on suundumus, mille vajalikkust ja võimalusi teadusele ning ühiskonnale laiemalt tajuvad kõik teemaga seotud osapooled, hoiakud selle suhtes on valdavalt soosivad. Detailide juurde liikudes leiavad erinevad osapooled, et avatud juurdepääsu võimaldamine kõigile teadusartiklitele ja -andmetele kätkeb endas mitmeid praktilisi probleeme ja teaduskorralduslikke küsitavusi. See raport tutvustab Eesti huvirühmade hoiakuid ja rolle avatud teaduse valdkonnas, eesmärgiga aidata kaasa tõhusa ja osapoolte parimaid huviseid arvestava avatud teaduse süsteemi kujundamisel Eestis.

Uuringu raport on osa RITA programmist<sup>1</sup>, mille vahendusel toetab Eesti Teadusagentuur Eesti riigi vajadustest lähtuvate sotsiaal-majanduslike eesmärkidega rakendusuuringute läbiviimist. Uuring keskendub kahele kesksele avatud teaduse eesmärgile: avatud juurdepääs teadusartiklitele ja avatud juurdepääs teadusandmetele. Uuringu üks peamisi lähtekohti oli 2016. aasta juulis Eesti Teadusagentuuri avatud teaduse ekspertkomisjoni koostatud põhimõtted ja soovitused riikliku poliitika kujundamiseks.

Raporti põhieesmärgiks on esitada faktiline alusmaterjal, mille alusel saaks mõtestada ja kirjeldada avatud teaduse suundumuse tähendust, võimalikkust ja potentsiaali Eesti riigi ja teaduse arenguvajaduste kontekstis ning esitada soovitused avatud teaduse kui horisontaalse põhimõtte sisseviimise kohta olemasolevatesse meetmetesse ja tingimustesse.

Nendest lähtuvalt on olulisemateks uurimisülesanneteks:

1. kaardistada olemasolevad vahendid ja viisid, millega avatud teadust toetatakse;
2. identifitseerida avatud juurdepääsu erinevad strateegiad erinevates riikides ning eristada (sh kuluefektiivsuse alusel) Eestile enam sobivad meetmed;
3. analüüsida Eesti võimalusi EL avatud teaduse tegevuskava, sh varasema TIPSi uuringu valguses;
4. kirjeldada T&A asutuste ja teadlaste rolli avatud teaduse põhimõtete kujundamisel, mille koosvaatlus annaks baasi soovitusteks teadlase-asutuse-riigi tasandil vastutuse jagunemise mõistlike proportsioonide kohta.

Eeldatavad tulemused:

1. Milline on avatud teaduse tegelik maksumus Eesti riigile, teadusasutustele ja teadlastele ja millistest komponentidest see koosneb?

Laiemas kontekstis saame anda osalise vastuse ka küsimusele, mis mõnevõrra laiendab, kuid samas kontekstualiseerib täpsemalt lähteülesande sisu:

2. Milliseid institutsionaalseid ja protsessilisi muudatusi, regulatsioone ja uusi lahendeid eeldab avatud teaduse põhimõtete järgimine?

Uuringu tulemuste mõju on mitmetasandiline: a) need on vahetuks sisendiks Eesti teaduspoliitika kujundamisele ja avatud teaduse põhimõtete rakendamisele teadustöös ja teadustegevuse korraldamisel läbivalt teadlane-asutus-riik tasanditel, b) sihtrühmade

---

<sup>1</sup> <http://www.etag.ee/rahastamine/programmid/rita/>

(tudeng, teadlane, raamatukogu, kirjastaja, ettevõtja, poliitikakujundaja jt) spetsiifilised ja nende vajadusi arvestavad ning c) lõppväljundina koostatav teekaart (*roadmap*) loob võimaluse tulevikku suunatult kujundada laia spektri tegutsejate käitumismustreid.

Uuring on läbi viidud Tallinna Ülikooli (TLÜ) ja Eesti Teaduste Akadeemia (ETA) koostöös. Raporti esimene osa kirjeldab erinevaid regulatsioone, uuringuid ning teiste riikide praktikaid, mis on kokku pandud Rein Vaikmäe (ETA), Kristel Toomi (TLÜ) ja Sigrid Mandre (TLÜ) poolt. Teine osa kirjeldab veebipõhise küsitlusuuringu ja fookusgrupiintervjuude tulemusi, mis on peamiselt TLÜ uurimisrühma liikmete Arko Olesk, Esta Kaal ja Kristel Toom kokku pandud. Avatud teaduse maksumuse, stsenaariumite ning teekaardi kirjeldamise osa põhineb eelnevatel osadel ning ülevaate on koostanud Raivo Ruusalepp (TLÜ), mida ilmestab näitena avatud publitseerimise kulude ülevaade Eesti Teaduste Akadeemia Kirjastuse näitel, mille on koostanud Ülo Niine (ETA). Töörühma juhtis Katrin Niglas (TLÜ).

Alljärgnevas tekstis kasutatakse võtme-termineid järgmistes tähendustes:

Termin	Inglisekeelne vaste	Seletus
Andmearhiiv	<i>Data repository, data archive, data library</i>	Teadusandmete haldamiseks ja säilitamiseks kasutatav tarkvaralahendus, koos selle pidamiseks vajaliku taristu ja töötajatega
Andmehaldus	<i>Research data management</i>	Teadusandmete haldamiseks ja säilitamiseks kasutatav tarkvaralahendus, koos selle pidamiseks vajaliku taristu ja töötajatega
Andmehaldusplaan	<i>Data Management Plan</i>	Andmehaldusplaan on osa uuringu või projekti planeerimisel, mis kirjeldab andmete kogumist, hoiustamist, analüüsimist ning pikaajalist säilitamist, sealhulgas avalikustamist ja tundlike andmetega seonduvat
Andmete salvestamine	<i>Data storage, storage</i>	Passiivne andmete hoiustamine andmekandjal, s.h. varundamine
Avaandmed, avatud andmed	<i>Open data</i>	Analüüsis käsitlesime teadusandmed kui andmed, mis on kogutud või loodud teadustöö käigus ning mis on aluseks uute teadustulemuste saamiseks. Avaandmetena käsitleme arvandmeid, tekstilisi andmeid, laboripäevikuid/märkmed, küsimustikke, video-, audio-, pildimaterjale, programme, mudeleid, näidiseid, algoritme, kirjavahetusi jms, mis on avatud ligipääsuks ja kasutamiseks (sh osaliseks) teistele isikutele. Kuna avaandmete haare on laiem, siis käsitleme kitsamalt teaduse avaandmeid.
Avatud juurdepääsuga	<i>Open Access publishing;</i>	Avatud juurdepääsuga avaldamise all peame silmas kirjastusmudelit, kus avaldatud artikkel on

avaldamine	<i>golden open-access</i>	kirjastuse veebilehel viivituse ja lõplikul kujul kõigile tasuta kättesaadav. Kirjastamiskulud võivad olla publikatsiooni autorite kanda (artikli avaldamistasu) või toetab ajakirja väljaandmist mõni asutus või organisatsioon.
Digihoidla, andmehoidla, repositoorium	<i>Institutional repository</i>	Spetsiaalselt digitaalsete andmete hoidmiseks loodud keskkond
Digitaalne säilitamine	<i>digital curation, active digital preservation</i>	Digitaalsete andmete vormingute ja tarkvara seiramine ning vajadusel migreerimine uutesse vormingutesse, et tagada andmete kestev kasutatavus
Ise-arhiveerimine	<i>Self-archiving; green open-access</i>	Ise-arhiveerimise all peame silmas tava, kus autor arhiveerib avaldatud teadusartikli või retsenseeritud käsikirja institutsionaalses (nt ülikooli digiarhiiv), riiklikus (nt ETIS) või rahvusvahelises repositooriumis (nt ArXiv, PubMed Central, sh kommertslahendused nagu ResearchGate), mille kaudu muutub see kõigile kättesaadavaks.
Teadusandmed, uurimisandmed	<i>Research data</i>	Uurimisandmeteks peame kõiki andmeid, mis on loodud või kogutud uurimisteema/projekti raames, olenemata sellest, kas nende põhjal on avaldatud teaduspublikatsioon või mitte. Uurimisandmete näited on statistika, mõõtmiste või katsete tulemused, välitööde märkmed ja tähelepanekud, küsitluste tulemused, intervjuude salvestused ja transkriptsioonid, fotod ja muud kujutised.

## Kokkuvõte

Tulenevalt lähteülesandest on aruandes esitatud faktiline alusmaterjal võimaldamaks mõtestada ja kirjeldada avatud teaduse suundumuse tähendust, võimalikkust ja potentsiaali Eesti riigi ja teaduse arenguvajaduste kontekstis ning on esitatud soovitusi avatud teaduse kui horisontaalse põhimõtte sisseviimise kohta olemasolevatesse meetmetesse ja tingimustesse. Kaardistati olemasolevad vahendid ja viisid, millega avatud teadust toetatakse. Dokumendianalüüsi ja personaalsete kontaktide kaudu identifitseeriti avatud juurdepääsu erinevad strateegiad arenenud riikides ja eristati Eestile enim sobivad meetmed. Analüüsiti Eesti võimalusi EL avatud teaduse tegevuskava ja varasema TIPSi uuringu valguses. On kirjeldatud T&A asutuste ja teadlaste rolli avatud teaduse põhimõtete kujundamisel, andmaks baasi soovitusteks teadlase-asutuse –riigi tasandil vastutuse jagunemise mõistlike proportsioonide kohta. Analüüsiti milline on avatud teaduse ühiskondlik kasu erinevate sihtgruppide vaatepunktist. Uuringus kasutatud metoodika sisaldas dokumendianalüüsi, veebipõhist küsitlusuuringut ja fookusgruppiintervjuusid.

Dokumendianalüüs käsitleb lähtekohana avaandmetega seotud Euroopa Liidu regulatsioone, milledest on lähtunud ka kõigi vaadeldud riikide Hollandi, Iirimaa, Poola, Sloveenia, Soome, Suurbritannia, Taani, ja Ungari praktikad.

Kuna teadusraamatukogud on olnud teaduse avalikustamise protsessis olulisel kohal, siis on ka teadusandmete avalikustamise protsessis üheks oluliseks lüliks raamatukogud.

Euroopa teadusraamatukogud pakuvad või plaanivad pakkuda tervet rida teadusandmete haldamisega seotud teenuseid. Näib, et on välja kujunenud teatud teenused, mida raamatukogud enam pakuvad. Euroopa raamatukogud, nagu ka teadusraamatukogud Põhja-Ameerikas (uurimistulemused 2012.a ja 2015.a), pakuvad pigem meelsamini konsultatsioonitüüpi (andmehaldusplaani, metaandmed, andmete standardid) teenuseid kui praktilisi/tehnoloogilisi andmehalduse teenuseid (nt andmete identifitseerimine raamatukogu repositooriumisse sisestamiseks).

Konsultatsiooniteenused (kuidas leida informatsiooni andmehaldusplaani koostamiseks, metaandmete standardeid või viidata andmetele) haakuvad suures osas raamatukogudes juba kaua pakutud infoteenustega.

Euroopa raamatukogudes pakutavate teenuste hulka kuulub peamiselt andmete haldamise teenuste käsitlemine ning poliitikate planeerimine ja väljatöötamine. Seni on andmete haldamise poliitika vähem kui pooltel raamatukogudel, mis näitab, et nad on alles oma teenuste arendamise suhteliselt varajases staadiumis. Euroopa raamatukogud pakuvad andmete haldamise teenuseid mitmetes valdkondades, kuid kõige enam sotsiaal- humanitaar- ja inseneriteaduste alal. Uurijatele tuli üllatusena, et uuringus osalenud raamatukogudes pakutakse väga vähe teenuseid meditsiiniteaduste valdkonnas. Sellele tuuakse kaks võimalikku põhjendust: 1. meditsiiniteaduste uurimisrühmades on oma spetsialistid, kes tegelevad teadusandmete haldamisega või 2. uuringus ei olnud esindatud piisavalt meditsiini raamatukogusid. Andmete haldamise teenuse väljatöötamine eeldab raamatukogutöötajaid, kellel on teadmisi ning võimalus omandada uusi oskusi. Paljud raamatukogud pakuvad oma töötajatele võimalust õppida teadusandmete haldamist, mõned raamatukogud võtavad sellele tööle uusi inimesi. Eeltoodust lähtudes tekib paratamatult küsimus, mis hetkel on vastuseta - kas raamatukogud on mõelnud ka sellele, kas ja millised raamatukoguteenused ära jätta, et andmete haldamise teenust paremini osutada. Läbiviijad peavad seda edasistes uuringutes käsitlemist nõudvaks temaks. Värskel avaldatud uuring COAR (Confederation of Open Access Repositories) poolt



juhhib tähelepanu avaandmete avalikustamise juures eelkõige meta-andmete olulisusele ja korrektsusele, eriti sotsiaalteadustes. Uuring toob välja ka institutsionaalse võimekuse tõstmise vajaduse, kuna uuringu tulemusena joonistub välja, et teadusasutustel puudub valdavalt avatud andmete põhimõtteid reguleeriv kord (sh andmete pikaajalise säilitamise kord). Samuti on ühe nõrkusena uuringus välja toodud tervet protsessi hõlmava tugisüsteemi puudumine. Uuring juhhib tähelepanu ka, et andmete standardiseerimise põhimõtted ja suuremahuliste andmemassiivide haldamine vajab vastava ekspertiisi välja arendamist.

Dokumendianalüüsis vaadeldud riikide lõikes võib tõdeda, et suundumused on samad ning neid kujundab eelkõige EL avatud teaduse suunaline poliitika ning rahastusmeetmete nõuded. Riigid ja ülikoolid, kes sõltuvad rohkem EL rahastusest on hakanud vastavaid meetmeid teaduse avaandmete toetamiseks ja reguleerimiseks varasemalt rakendama ning protsessid on paremini välja arendatud.

Veebipõhise küsitlusuuringu põhilised tulemused näitasid, et Avatud teadus on suundumus, mille vajalikkust ja võimalusi teadusele ning ühiskonnale laiemalt tajuvad kõik teemaga seotud osapooled, hoiakud selle suhtes on valdavalt soosivad. Ent detailide juurde liikudes leiavad paljud, et avatud juurdepääsu võimaldamine kõigile artiklitele ja teadusandmetele kätkeb endas mitmeid praktilisi probleeme ja teaduskorralduslikke küsitavusi. Analüüs tutvustab Eesti huvirühmade hoiakuid ja tegevusi avatud teaduse valdkonnas, eesmärgiga aidata kaasa tõhusa ja osapoolte parimaid huviseid arvestava avatud teaduse süsteemi kujundamisel Eestis.

Teadlaste seas läbi viidud küsitlusest nähtub, et avatud juurdepääsus artiklitele nähakse teaduse jaoks kasu: väite „avatud juurdepääs kõigile artiklitele aitab tõsta teaduse ja hariduse kvaliteeti“ keskmine toetus viiepallisüsteemis on 3,78. Toetavad ollakse ka idee suhtes, et kõik artiklid võiksid olla vabalt kättesaadavad (keskmine hinne 3,73).

Kõige kõrgema toetusega (3,97) sai aga väide „ilmumiskulude katmise kohustus autorite poolt soodustab ebavõrdsust publitseerimisel“. Uuring tõi välja teadlaste seas levinud mõtteseose, mille kohaselt avatud juurdepääsu tagamine seostub eelkõige avaldamisega *open-access* ajakirjades, eelkõige sellistes, kus avaldamise kulud tuleb katta autoritel. Nendega seoses tuuakse enamasti esile kaks probleemi: avatud juurdepääsuga ajakirjade kvaliteet ning neis avaldamise rahastusmudel.

Kvaliteedi hindamise osas ilmnevad erisused teadusvaldkondade vahel. Mõne vastanu sõnul ongi nende valdkonna parim ajakiri *open-access* (ajakirja mainet tuuakse küsitluse vastanute seas esile peamise põhjusena, miks nad avaldavad avatud juurdepääsuga ajakirjades, viiendik vastanuist avaldab seal põhimõtte pärast). Ja kvaliteedi puudumist („need on meie erialal siiski rämp“) toodi küsitluses avatud vastusena välja neis mitteavaldamise põhjusena. Paludes hinnangut väitele „avatud juurdepääsuga teadusajakirjade kvaliteet on sama kõrge kui tavalistel teadusajakirjadel“ jagunesid hinnangud viiepallisüsteemi hinnete vahel pea võrdselt (keskmine 2,72). See võib ühelt poolt peegeldada teatud vastajarühmade jäikaid mõttemustreid, kuid näitab kindlasti ka avatud juurdepääsuga ajakirjade ebaühtlast arengut ja teadusvaldkondade erinevat valmisolekut võtta omaks avatud teaduse põhimõtteid.

Keskne aruteluteema *open-access* avaldamise puhul oli siiski raha: teadlaste tavapärase esimene reaktsioon on küsida, kes avaldamistasud nende jaoks kinni maksab. On selge ootus, et kui rahastaja nõuab, et tulemused oleksid avaldatud avatud juurdepääsuga ajakirjades, siis eraldatakse selle jaoks ka täiendavaid ressursse. Samas on sõnum otsustajate poolt – lisaraha ei tule ning avaldamise kulud tuleks katta praeguste projektide raames. Täpselt seda hirmu, et uued nõuded võivad kaasa tuua reaalse teadusraha kahanemise, väljendasid teadlased intervjuudes korduvalt.

Intervjuudes osalenute (sh raamatukogude esindajate) hinnangul pole realistlik ka see, et loobuda senistest teadusajakirjade tellimustest ja suunata need vahendid avaldamiskulude katteks, kuna see tooks siiski kaasa senisest oluliselt halvema juurdepääsu teadusajakirjadele.

Fookusgrupiintervjuudest tuli siiski välja ka see, et kui teadlase motivatsioon *open-access* avaldamiseks on piisav, muutub avaldamiskulude küsimus teisejärguliseks.

Kuigi üldised hoiakud on artiklite vaba juurdepääsu soosivad, juhib teadlaste avaldamisstrateegiaid peamiselt põhimõtte, et nende artikkel oleks hõlpsalt leitav ja (oma valdkonna inimeste poolt) võimalikult laialt loetav. Sõltuvalt valdkonnast võib see tähendada avaldamist valdkonna tippajakirjades (sõltumata selle litsentsitingimustest), teadlikke valikuid *open-access* ajakirjade kasuks, laialdast isearhiveerimist, alternatiivsete publitseerimisvormide eelistamist (nt monograafiad, konverentsiannaalid) või hoopis tulemuste kiiret levitamist *working paper*'ite või *pre-print*-artiklite vormis. Mõni valdkond on artiklite vaba juurdepääsu põhimõtted laialdaselt omaks võtnud, teine näeb potentsiaalses nõudes levitada artikleid eelistatud *open-access*-formaadis vastuolu nende valdkonnas tunnustatud parimate publitseerimistavadega ja seeläbi ka ohtu enda töö nähtavusele.

Artiklitele vaba juurdepääsu kujundavaid tingimusi mõjutavad oluliselt globaalsed aspektid, mis ei ole kohalikul Eesti tasandil mõjutatavad – eelkõige valdkondlikud trendid eelistatud publitseerimistavade osas (mis võivad olla nii *open-access*'i soosivad kui taunivad), kirjastajate käitumine (enamasti piiranguid seadev) kui ka suurte teadusrahastajate seatud tingimused (enamasti *open-access*'i eelistavad).

Kohaliku tasandi tegevused (koolitused, teavitustegevus, rahastajapoolsed nõuded) saavad toetada globaalset suundumust avaliku juurdepääsu suunas ning kujundada siinsete teadlaste ja teiste huvirühmade hoiakuid ja käitumist.

#### **Soovitused:**

- Kasutada teadlaste hoiakute ja käitumise suunamiseks sama valdkonna esindajaid, kes on avatud teaduse eeskõnelejad;
- Kaardistada valdkondlikud võimalused;
- Isearhiveerimiskeskondade kasutustingimuste selgitamine;
- Arvestada valdkondlikke erinevusi.

Nagu selgus dokumendianalüüsil on ka teiste riikide puhul avaandmete temaatika märksa komplitseeritum, kui OA publitseerimise temaatika.

Meie uuringus tunnistasid rohkem kui pooled küsitlusele vastanutest, et hoiavad enda uurimisandmeid isiklikus arvutis või andmekandjal. Valdkondlikke repositooriume kasutab oma andmete hoiustamiseks vaid mõni protsent vastanuist. Kuigi vastanute enamik ütles, et on enda valitud hoiuviisiga rahul, tunnistasid paljud siiski, et isiklikus seadmes säilitamisega kaasnevad andmekaotuse risk, ruuminappus ning raskused andmete jagamisel. Sellega seoses väljendasid teadlased huvi institutsionaalsete või riiklike andmehoidlate vastu. Sellest nähtub, et andmete asjakohase hoiustamise vastu oli neil suurem huvi kui avalikult kättesaadavaks tegemise vastu. Et aga esimene on vajalik samm teise võimaldamiseks, tasub panustada teadlikkuse tõstmisse ja säilitusvõimaluste tutvustamisse. Need on raamatukogude hinnangul juba praegu piisavad, katmaks eri teadusvaldkondade vajadusi.

Nii saab andmete hoiustamiseks (ja avaldamiseks) valdkonnast sõltuvalt valida Eestis loodud valdkondlikke repositooriume, rahvusvahelisi valdkondlikke repositooriume ja raamatukogupõhiseid repositooriume. Mõned vastanud väljendasid huvi, et võiks olla ka üks riiklik keskne Eesti

repositoorium siinsete andmete jaoks, teised leidsid, et see oleks liigne ressursside kulutamine ja olemasolevad võimalused on piisavad (kuid peaks olema koht, nt ETIS, kuhu on koondatud lingid andmekogudele või pakutud muu moodus neid hõlpsaks leidmiseks).

Ühe olulise teemana kerkis üles teaduse avaandmete suhestumine riigi avaandmetega. Teadlaste hinnangul annaks parem juurdepääs riigi andmetele meie teadusele palju enam kui teadusandmete avamine, seega oleks riigi eeskuju oluline ka teadlaste motiveerimisel. Ka riigi ja selle allasutuste tellitud uuringute andmete puhul puudub praegu ühtne poliitika nende kättesaadavuse osas. Lisaks haakub teaduse avaandmete teema ka kogude digiteerimisega, mida eelkõige mainisid raamatukogud.

Võrreldes kirjastamisega (ning teatud määral ka isearhiveerimisega) on avaandmete juures praegu vähem ärihuvisid (mis on näiteks kirjastamise juures üks teadlastes tõrkeid tekitavaid aspekte). Kui need ärihuvid tekivad (andmekaeve jms näol) – mida mõned eestkõnelejad ka eesmärgina näevad – võib see samuti kaasa tuua potentsiaalse vastuseisu oma andmete jagamisele.

Tervikuna selgus et juhtudel, kus on näha valdkondlik loogika ja potentsiaalne kasu ning kasutajad, on teadlased rohkem motiveeritud algandmeid kättesaadavaks tegema ning on ka juurutanud sellekohased praktikad. Kuid selline selge positiivne mõju on praegu vaid teatud tüüpi (valdavalt kvantitatiivsete loodusteaduslike) andmete korral, ülejäänud valdkondades on palju enam kahtlusi.

Põhimõttena toetavad avatust pea kõik vastanud ja huvirühmad. Vastava käitumise vähesust saab ühelt poolt selgitada napi teadlikkusega, teisalt aga motivatsiooniküsimustega. Motivatsiooni oluline komponent on tegevuse eesmärgipärasus, kasude ja tulemuste tajumine. Just avatud teaduse konkreetsemate kasude või mõjude sõnastamisega ei saanud huvirühmad kuigi edukalt hakkama. On paar teema eestkõnelejat, kes suudavad välja tuua erinevaid mõjusid, kuid pigem abstraktselt ja tulevikkuvaatavalt, mitte juba toimivate konkreetsete näidetena (või on need üsna spetsiifilised, puudutades näiteks seireandmete kasutamist).

Analüüsi tulemused näitavad, et teadlased mõtleavad avatud teaduse kasudest peamiselt teadusekeskselt (toob parema koostöö, tagab tulemuste laiema nähtavuse, parandab kvaliteeti jne). Seda peamiselt seoses avatud juurdepääsuga artiklitele, vähem aga teaduse avaandmete puhul, sest teaduskogukonnas on andmete omavaheline jagamine (omavahelise suhtluse vahendusel) niigi toimunud.

Avatud teaduses nähakse küll potentsiaalset kasu ka teistele, teadusvälistele huvirühmadele (nt otsustajad, ajakirjanikud, õpilased), ent lähemalt analüüsidest ollakse siiski skeptilised nende potentsiaalsete kasutajate suutlikkuses artikleid või andmeid mõista ja töödelda. Ka need huvirühmad ise tunnevad põhimõttelist huvi avatud juurdepääsu vastu ja ütlevad, et nende tegevus võiks olla teaduspõhisem, kuid tajuvad olulisemana teisi barjääre peale juurdepääsu puudumise (eelkõige keerukus).

Meie töö üks keskseid soovitusi on sõnastada avatud teaduse positiivseid mõjusid eri ühiskonna tasanditel, koos näidetega, ning kasutada neid teadlikkuse tõstmisel ning avatud teaduse suunal liikumise vajalikkuse põhjendamisel. Raamatukogud on motiveeritud teemaga tegelema ning näevad enda rolli eelkõige harija ja teavitajana, mõnel juhul ka taristu (repositooriumite) pakkujana. Samas kurdavad ressursside (eelkõige inimeste ja aja) puuduse üle. Teadlased tunnevad hirmu, et avatud teaduse nõuetest saab järjekordne lisakoorem, mis pidurdab nende võimalusi pühenduda teadustegevusele – nii raha kui ka muude ressursside osas. On väliste nõuete osas pigem skeptilised ja eelistaksid otsustusvabadust nii artiklite kui ka andmete kättesaadavaks tegemise tingimuste üle. Samas on mõnes valdkonnas need põhimõtted juba omaks võetud ja laialdaselt rakenduses.

## Summary

Ensuing from the initial assignment, the report has produced the factual material to enable the interpretation and describe the meaning, possibility and potential of the trends of Open Science in the context of the developmental needs of the Estonian sciences and the State. Suggestions have been made for the adoption of Open Science as a horizontal principle into the existing measures and conditions. The mapping of the existing means and methods by which open science is supported was carried out. Through document analysis and personal contacts the different strategies for open access were identified in developed countries and the measures best suited for Estonia were identified. The possibilities of Estonia within the context of the European Union Open Science policy and earlier Research and the Innovation Policy Monitoring Programme (TIPS Programme) were analysed. The role of the Research and Development Institutions and researchers in shaping the principles of Open Science has been described to create a basis for recommendations for the sensible and proportionate division of responsibilities between the researchers, institutions and the State. Analysis was conducted on what is the societal profit of Open Science from the viewpoint of different target groups. The methods used for the research were document analysis, web based survey research and focus group interviews.

The document analysis used as its starting point the European Union regulations related to open access data, which has been the ground principle for the practices of all the investigated countries: Netherlands, Ireland, Poland, Slovenia, Finland, Great Britain, Denmark and Hungary.

Since the research libraries have held a significant position in the process of providing free access to science, the libraries are an important link in the process of generating free access to research data.

The European research libraries provide or plan to provide several services related to the management of research data, and there appear to be certain services that the libraries offer more often. Both European libraries and the research libraries in North America (research data of 2012 and 2015) offer more willingly consultation type of services (data management plans, meta-data, standards of data) than practical and technological data management services (for example: identifying data for the purpose of entering it into a library repository).

The consultation services (how to find information for compiling a data management plan, standards of meta-data, or how to refer to the data) are mostly related to the information services that the libraries have been offering for a long time.

Among the services provided by the European libraries there are mainly services related to data management and planning and developing related policies. So far less than half of the libraries have data management policies which shows that they are still in the early stage of development of their services. The European libraries offer data management services in several fields, most widely in Social Sciences, Humanities, and Engineering. It came as a surprise to the research team that the libraries that participated in the research offered very few services in the field of Medicine. Two possible explanations are given: 1. Research teams in medical sciences have their own specialists who deal with management of research data; 2. The research had an insufficient representation of medical libraries. As a prerequisite for the

developing of data management services there should be librarians who have suitable knowledge and opportunities to acquire new skills. Several libraries offer for their employees the opportunity to learn management of scientific data, some libraries recruit new employees. Based on this a question arises that remains at the moment unanswered: have the libraries thought whether they should drop any library services and which ones in order to provide better data management services? The conductors of the present research consider this a topic that should be dealt with in future research. The recently published research by COAR (Confederation of Open Access Repositories) points out that when making available free access data the meta-data is significant and has to be correct, especially in the Social Sciences. The report points out the need for increasing institutional capacity since, as a result of the report it becomes evident that the research institutions commonly do not have any regulations concerning the principles of free access data (including the rules for long term conservation of data). The report points out as one of the weaknesses the non-existence of a support system that would cover the whole process. The report also points out that for adopting principles of standardising data and managing big data sets the development of appropriate expertise is required.

The document analysis on the countries included confirms that the trends are the same and these are influenced mainly by the European Union policies of Open Science and the requirements of financing measures. The countries and universities who depend more on European Union financing have started to adopt measures for supporting and regulating free access scientific data sooner than others and consequently their related processes are better developed.

The main results of the web based survey research indicated that the importance and opportunities that Open Science provided for research and for the society more widely are recognised by all the parties related to the topic and the attitudes towards it are mainly positive. Nevertheless, when tackling the details many believe that free access to all articles and research data holds several practical problems and doubts related to scientific organisation. The analysis introduces the attitudes and activities of Estonian stakeholders in the field of Open Science and aims to foster the development of an efficient Open Science system in Estonia that would also take the best interests of the parties involved into consideration.

From the survey conducted among the scientists it becomes apparent that the open access to articles is seen as profitable for science. On a five point scale the average of 3.78 was given to the claim that “open access to all articles helps to enhance the quality of research and education”. The idea that all articles could be freely accessible is also supported (average 3.73).

The highest support (3.97) was given to the claim that “compulsion of the authors to cover the publishing costs favours inequality of publishing”. The research confirmed the common assumption among scientists that for guaranteeing open access for articles relates to publishing in open-access journals, mainly in those where the publishing costs must be covered by the author. In relation to these publications two obvious problems are pointed out: the quality of the open access journals, and the financing scheme for publishing in them.

When assessing quality, differences between areas of science become apparent. According to some respondents the best journal of their field is the open-access publication (the reputation of a publication is pointed out by the respondents as the main cause why they publish in open access journals, one fifth of the respondents publishes there out of principle). The lack of quality (“in our field they are rubbish”) was pointed out in the open answer section as a reason why one is not publishing in some journals. When assessing the claim that “the quality of open access scientific journals is as high as the quality of ordinary scientific journals” the answers were divided almost equally (2.72 on the five point scale). This could reflect the rigid thinking patterns of some respondent groups but shows undoubtedly the uneven development of open access journals and the differences in readiness of various science areas to accept the principles of Open Science.

The core topic of discussion in relation to open access publishing is nevertheless money: the common reaction of scientists is to ask who will pay their publishing costs for them. There is clear expectation that if the financier demands the research results to be published in an open access journal, additional funding is provided for it. At the same time there is a message from the decision makers: additional funding will not be provided and the publishing costs should be covered within the existing project. The same fear that the new requirements could cause actual diminishing of research funding was expressed by the scientists during their interviews on several occasions.

The participants of the interviews (including the representatives of the libraries) did not consider it realistic that the orders for scientific journals should be discontinued and the resources reallocated for covering the publishing costs since this would result in diminishing access to scientific journals when compared to the present situation. The focus group interviews revealed, nevertheless, that if a scientist has sufficient motivation for publishing in an open access journal the issue of publishing costs is no longer a primary concern.

Although the general attitudes are supportive of free access to articles the publishing strategies of scientists are mainly directed by the principle that their article should be easy to find and read as widely as possible (by the specialists of the same field). Depending on the field it may mean publishing in the top-ranking journals of the specific field (regardless of the licensing terms), conscious choices for open access journals, big scale self-archiving, preference for alternative publication forms (for example monographs, conference annals) or quick dissemination of the results as working papers or pre-print articles. Some fields have widely accepted the principles of free access to articles, some others see in the potential demand to disseminate the articles preferably in the open access format, a contradiction to the best publishing practices acknowledged in their field and as a result, a threat to the visibility of their own work.

The conditions shaping the free access to articles are significantly influenced by global aspects which cannot be influenced at local, Estonian level; mainly the field related trends of preferred customs of publication (which can be both either favouring or disapproving open access), the behaviour patterns of the publishers (mainly limit setting), and the conditions set by the large scale science financiers (mainly preferring open access).

The activities at local level (training courses, information dissemination, requirements of the financier) may support the global trends towards open access, and shape the attitudes and behaviour of the local scientists and other stakeholders.

### **Recommendations:**

- For guiding the attitudes and behaviour of the scientists, representatives of the same field who are also promoters of open science should be used;
- Mapping the field related possibilities;
- Explaining the user conditions of self-archiving environments:
- Taking into account field related differences.

As the document analysis has shown, the topic of free access data is more complex in case of other countries than just the topic of open access publishing.

During our research more than half of the respondents admitted that they keep their research data in their private computer or data carrier. The field related repositories are used for storing data only by a few percent of the respondents. Although the majority of the respondents said that they are content with their chosen method of data storing, many admitted that when storing data on a private carrier there is the risk of data loss, lack of space and difficulties with sharing data. In relation to this the scientists expressed interest towards institutional or national data repositories. This indicates that their interest towards suitable storing of data is greater than towards making the data freely accessible. While the first is a necessary step for making the second possible, more attention should be given to raising the awareness and introducing the options for data storage. The storage facilities are already sufficient, according to the estimates of the libraries, to cover the needs of different scientific fields.

For storing (and making available) data, depending on the field, Estonian based file related repositories, international field related repositories and library based repositories can be chosen. Some respondents expressed interest that there could be one Estonian central and national repository for the local data, but others found that it would be too draining for the resources and the existing options are sufficient (although, there should be a place, for example ETIS, that stores the links to the data repositories, or some other means and method for making it easy to find them).

One important topic came up concerning the relation of scientific free access data to the free access data of the State. According to the estimates of the scientists, better access to the State data would be more profitable for our science than allowing open access to the research data; therefore, the example of the State would be an important factor for motivating the scientists. At present there is no unifying policy regulating the accessibility to the data of such researches as have been conducted upon the order of the State or its various establishments. Furthermore, the topic of free access scientific data is related to the topic of digitations of the collections and this was mainly pointed out by the libraries.

Compared to publishing (and to some extent to self-archiving) there is less business interests related at the moment to free access data (business interests being one of the aspects causing reluctance in scientists in relation to publishing). Should these business interests arise (in the shape of data dig or something else), which some of the spokespersons see as an aim, it can also result in the potential resistance to sharing one's data. It became clear that on such occasions when the field related logic is visible and there is potential gain and users, the scientists are more motivated to make their primary data available and there are related practices in place. This kind of positive influence, nevertheless, is at the moment related to only certain types (mainly quantitative data of natural sciences) of data, other fields have significantly more doubts.

As a principle, almost all respondents and stakeholders support openness. The lack of corresponding behaviour can be explained not only by the lack of awareness but also by the issues of motivation. A significant component of motivation is the purposefulness of the activities, the cognisance of profits and results. But the specific profits or influences of Open Science were those the stakeholders could not very successfully define. There are few spokespersons who are able to bring out different influences, although in an abstract and looking into the future manner, not as already functioning concrete examples (or, the examples are rather specific and mean, for example, the use of monitoring data).

The analysis shows that scientists think of the profits of Open Science mainly in a professionally oriented way (brings better cooperation, provides wider visibility of the results, enhances quality etc). The main context for it is open access articles, less in the context of free access scientific data, since sharing data with co-researchers within the scientific circles has functioned anyway as part of mutual communication.

Although the potential gains of Open Science for other stakeholders outside sciences (for example, decision makers, journalists, pupils) are seen, on closer analysis there is still scepticism about the abilities of these potential users to understand and process these articles and/or data.

The stakeholders themselves are in principle interested in free access and say that their activities could have better scientific grounding, although they sense as significant other barriers beside the lack of access (mainly the complexity).

One of the central recommendations of our research is to define the positive influences of Open Science on different levels of society, provide examples, and use these to raise the level of awareness and for giving impetus for the necessity of moving towards Open Science. The libraries are motivated to deal with the topic and see their role mainly as educators and information disseminators, in some cases as infrastructure (repositories) providers. At the same time they complain about the lack of resources (mainly staff and time). The scientists are afraid that the demands of Open Science will become an additional burden that prohibits the opportunities to dedicate themselves to science, both financially and in relation of other resources. They are sceptical about the outside demands and would prefer the freedom of choice upon the conditions of making available both articles and research data. In some fields, though, these principles have already been adopted and in widespread use.



# 1 Dokumendianalüüs

Lähteülesanne: Eesti, OECD ja ELi dokumendid, programmid, regulatsioonid, ELi tegevuskava, temaatilised teadusartiklid, uuringud, analüüsid – kaardistamaks Eesti riigi, T&A asutuste ja teadlaste senist praktikat; b) teiste riikide praktika kaardistamine. Vaadeldakse 2 tüüpi riike: eeskujuriigid: Inglismaa, Holland, Taani, Iirimaa ja Eestiga võrreldavas positsioonis olevad riigid nt Ungari, Poola, Sloveenia. Praktikaid võrreldakse kahel tasandil: riik ja T&A asutused.

Dokumendianalüüs on jagatud kolme teemablokki ning võtab vaatluse alla (1) Euroopa Liidu regulatsioonid, mis avatud teadusandmetega seonduvat raamivad; (2) vaatleb erinevaid avatud andmete valdkonnas läbi viidud uuringuid; ning (3) kirjeldab erinevate riikide praktikaid nii riiklikul kui ka institutsionaalsel tasemel.

## 1.1 Euroopa Liidu regulatsioonid

Eesti teadusüsteemi, sealhulgas Eesti teadusasutusi mõjutavad oluliselt Euroopa OA (Open Access) eeskirjad ja regulatsioonid. Euroopa tasandil on vastu võetud mitmed õigusaktid, mis raamivad ja kujundavad EL (Euroopa Liidu) liikmesriikide vastavaid poliitikaid.

Euroopa Komisjon (EK) on oma soovitusel (*Commission's recommendation from 17 July 2012 on access to and preservation of scientific information (2012/417/UE)*)<sup>2</sup> liikmesriikidele pannud kohustuseks täpsustada riiklike OA strateegiaid ning üksikasjalikult ette kirjutanud selle, mida strateegiad peavad sisaldama. Dokumendis välja toodud soovitusi kohaldatakse mitte ainult teaduslike publikatsioonide avaldamisele, vaid ka teadusandmete avaldamisele. Soovituste koostamisel on silmas peetud EK plaani rakendada avatud teaduse põhimõtteid avalikult rahastatavatele projektidele. Euroopa Parlamendi ja Euroopa Nõukogu poolt kehtestatud Horisont2020 raamdirektiivi alusel muudetakse kohustuslikus OA põhimõtete rakendamine teaduslikele publikatsioonidele, mis on loodud Horisont 2020 programmi raames. Sama määrus sätestab ka avatud teadusandmete pilootprogrammi rakendamise seotud raamistiku<sup>3</sup>.

Horisont2020 avaandmete pilootprogramm hõlmab valitud teadusvaldkondi ning annab taotlejale võimaluse põhjendatud juhtudel pilootprogrammist loobuda. Pilootprogrammi täpsemad eesmärgid ja reeglid on kirjeldatud Horisont2020 programmi määruses<sup>4</sup>. Pilootprogrammiga liitudes on oluline teada nii õiguslikku raamistikku kui ka tehnilisi nõudmisi teadusandmete kogumisel, säilitamisel ning avalikustamisel. Kuna teadusandmete puhul on erinevalt OA publikatsioonidest mitmeid erisusi (nt isikuandmed), siis on nõuded teadusandmete avalikustamisele komplekssemad. Avatud teadusandmete aluseks on andmehaldusplaan (*Data Management Plan, DMP*), mis kirjeldab üksikasjalikult andmete haldamise elutsüklit – andmete kogumist, töötlemist, säilitamist ja avalikuks tegemist. Andmehaldusplaan peab sisaldama järgmist teavet:

<sup>2</sup><http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:EN:PDF>

<sup>3</sup>[http://ec.europa.eu/research/participants/portal/desktop/en/funding/reference\\_docs.html#h2020-mga](http://ec.europa.eu/research/participants/portal/desktop/en/funding/reference_docs.html#h2020-mga)

<sup>4</sup>[http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

- Andmete käitlemise põhimõtteid uuringute ajal ja pärast projekti lõppu;
- Milliseid andmeid kogutakse, töödeldakse ja/või luuakse;
- Millist metoodikat ja standardeid rakendatakse;
- Millisel moel andmed jagatakse ja kuidas tagatakse avatud juurdepääs;
- Kuidas andmeid säilitatakse (sh pärast projekti lõppu).

Horisont2020 2016. aastal avaldatud vaheraporti põhjal selgus, et suur osa pilootprogrammi põhivaldkondade projekte osalesid avatud andmete pilootprogrammis (65,4% 3699-st allkirjastatud toetuslepingust). Lisaks, 11,9% vabatahtliku osalemisega projektidest otsustas pilootprogrammiga liituda.<sup>5</sup>

EL avatud andmete poliitika eesmärk on üheltpoolt vähendada juba läbi viidud uuringute korduvat finantseerimist ning juba kogutud andmete kasutamist laiemalt:

- teadusuuringute käigus kogutud andmeid saab hiljem kasutada muudes teadusuuringutes;
- teadusuuringute andmeid saab kasutada muudel mitte-teaduslikel eesmärkidel teiste sihtrühmade poolt – ettevõtjad ja avaliku sektori asutused.

Järjest rohkem on EL (kui ka liikmesriikide) tasemel hakatud soosima juba eksisteerivate avaandmete põhjal tehtava teadustöö finantseerimist eraldi rahastusmeetme raames, mis muuhulgas kujundab tava andmete taaskasutamiseks.

Horisont2020 avaandmete põhimõtete järgimise lihtsustamiseks tuleb teadusasutustel tegeleda kahe peamise regulatsiooni loomise ja rakendamisega:

1. Teadusandmete haldamist reguleeriv kord, mis sisaldab andmehaldusplaani (*Data Management Plan*) koostamise põhimõtteid ning asutustepoolset tugisüsteemi teadlastele;
2. Andmete pikaajalise säilitamise tagamise strateegia, sh andmekogumite säilitamiseks ettevalmistamise standardid (metaandmete struktuur, isikuandmete hoidmise põhimõtted jms).

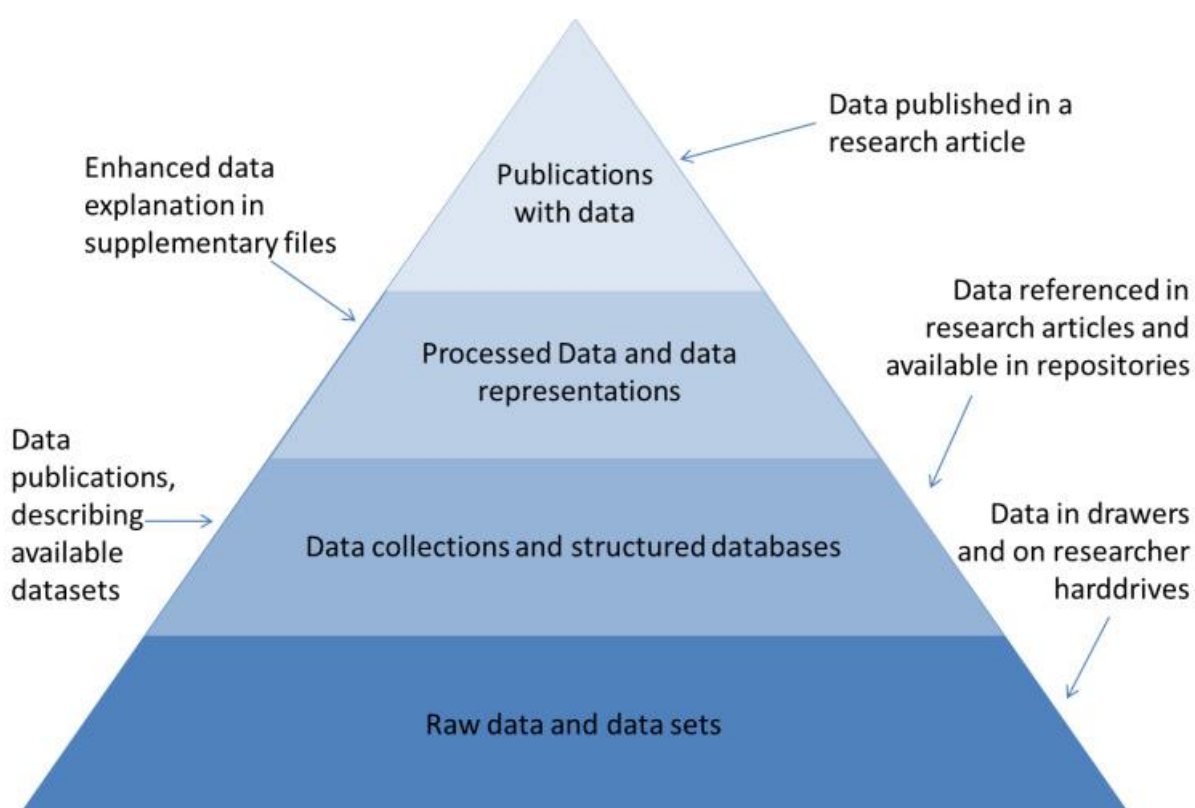
## 1.2 Valdkondlikud uuringud

LERU (*League of European Research Universities*) on 2012. aastal avaldanud oma seisukoha avatud andmete osas.<sup>6</sup> Avatud andmete kontseptsioon on küll juba varasemalt teada, kuid on viimastel aastatel selgepiirilisemalt formuleeritud ning sellest tulenevalt on järjest olulisem tähelepanu pöörata andmetega seotud õiguslikele aspektidele ja põhimõtete rakendamisega seotud protsessidele (sh tehnilistele nõudmistele). Dokument toob välja, et oluline on kirjeldada täpselt nii rahastajate poolset raamistikku kui ka ülikoolide, raamatukogude ning teadlaste vastutust protsessis.

<sup>5</sup><http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

<sup>6</sup><http://www.leru.org/files/general/Open%20Access%20to%20Research%20Data-FINAL.docx.pdf>

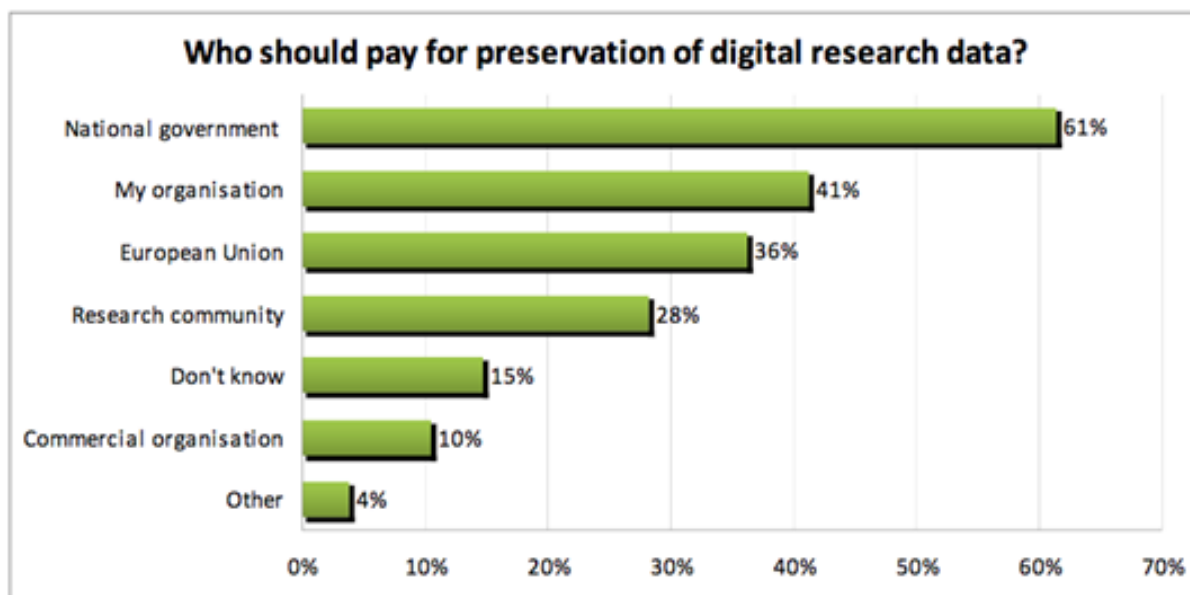
Raamatukogude rollist teadusandmete säilitamise ja kättesaadavuse protsessist (joonis 1) hakati LIBER (*Association of European Research Libraries*) eestvedamisel rääkima juba mitukümmend aastat tagasi. Projekti *Opportunities for Data Exchange* (rahastatud RP7 raames) 2011. aastal publitseeritud raport<sup>7</sup> toob välja, et peamiselt kasutatakse teadusandmeteni jõudmiseks omavahelist mitteformaalset andmevahetust (58% uuringus osalenutest). Peamise puudusena tuuakse välja, et kui teadlastel on väga spetsiifilised vajadused teadusandmete osas ning nende võrgustikus puudub nende andmetele ligipääs või teave nende kohta, siis on väga keeruline leida vajalikku infot. Sama uuring toob välja teadusandmetele juurdepääsu võimaldamisel kaks peamist takistust: õiguslikud küsimused ning andmete väärkasutamine, samas on olulise takistusena märgitud ka ebasobiv andmete formaat, tehnilise võimekuse puudumine ja piiratud rahalised võimalused.



**Joonis 1.** Andmete avalikustamise püramiid (*Opportunities for Data Exchange (ODE) Report on Integration of data and publications, 2011*)

Andmetele avatud juurdepääsu võimaldamine ja andmete säilitamise tagamine takerdub erinevate uuringute põhjal peamiselt rahaliste vahendite taha. Eeltoodud projekti raport toob välja, et peamiselt oodatakse andmete säilitamise osas vastuse võtmist riigi, institutsiooni või Euroopa Liidu poolt (Joonis 2).

<sup>7</sup> [http://rep4-vm.awi.de/31397/1/ODE-ReportOnIntegrationOfDataAndPublications-1\\_1.pdf](http://rep4-vm.awi.de/31397/1/ODE-ReportOnIntegrationOfDataAndPublications-1_1.pdf)



**Joonis 2.** Teadusandmete säilitamise osas ootavad teadlased tuge eelkõige riigilt (PARSE report on First Insights into Digital Preservation of Research Output in Europe, 2009)<sup>8</sup>

LIBER president P. Ayris on juhtinud mitmel korral tähelepanu, et tehnoloogia areng võimaldab teadus- ja avalike andmete ulatuslikumat ja mitmekülgsemat kasutamist, mistõttu on järjest olulisem andmetele ligipääs. Kuna teadusraamatukogud on olnud teaduse avalikustamise protsessis olulisel kohal, siis on ka teadusandmete avalikustamise protsessis ühes oluliseks lüliks raamatukogud.

Aastatel 2015-2016 viidi Euroopa teadusraamatukogudes läbi ulatuslik uuring „*Research Data Services in European Academic Research Libraries*“. Uuringu eestvedaja, Carol Tenopir oli viinud varem läbi sarnased uuringud Põhja-Ameerikas: 2012. aastal teadusandmete haldamise teenuste kaardistamine teadusraamatukogudes, ning 2015. aastal jätkuuring teadusandmete haldamise teenuste arengutest. Nimetatud uuringute põhjal koostati mõningate muudatuste ja kohendustega metoodika Euroopas läbi viidud uuringu jaoks.

Kutse uuringule saadeti 333 Euroopa teadus/ülikooliraamatukogule, kes kuulusid Rahvusvahelise Raamatukogude Assotsiatsioonide Föderatsiooni (IFLA) koosseisu. Vastused saadi 119 raamatukogust, 22 riigist. Vastanute hulgast jäid välja Horvaatia, Küprose, Prantsusmaa, Ungari, Luksemburgi, Malta, Portugali, Rumeenia, Serbia, Slovakkia ja Türgi teadusraamatukogud.

Uuringu tulemustes toodi välja, et Euroopa teadusraamatukogud pakuvad või plaanivad pakkuda tervet rida teadusandmete haldamisega seotud teenuseid. Näib, et on välja kujunenud teatud teenused, mida kõik raamatukogud pakuvad. Euroopa raamatukogud, nagu ka teadusraamatukogud Põhja-Ameerikas (uurimistulemused 2012.a ja 2015.a), pakuvad pigem konsultatsioonitüüpi (andmehaldusplaanid, metaandmed, andmete standardid) teenuseid kui praktilisi/tehnoloogilisi andmehalduse teenuseid (nt andmete identifitseerimine raamatukogu repositooriumisse sisestamiseks).

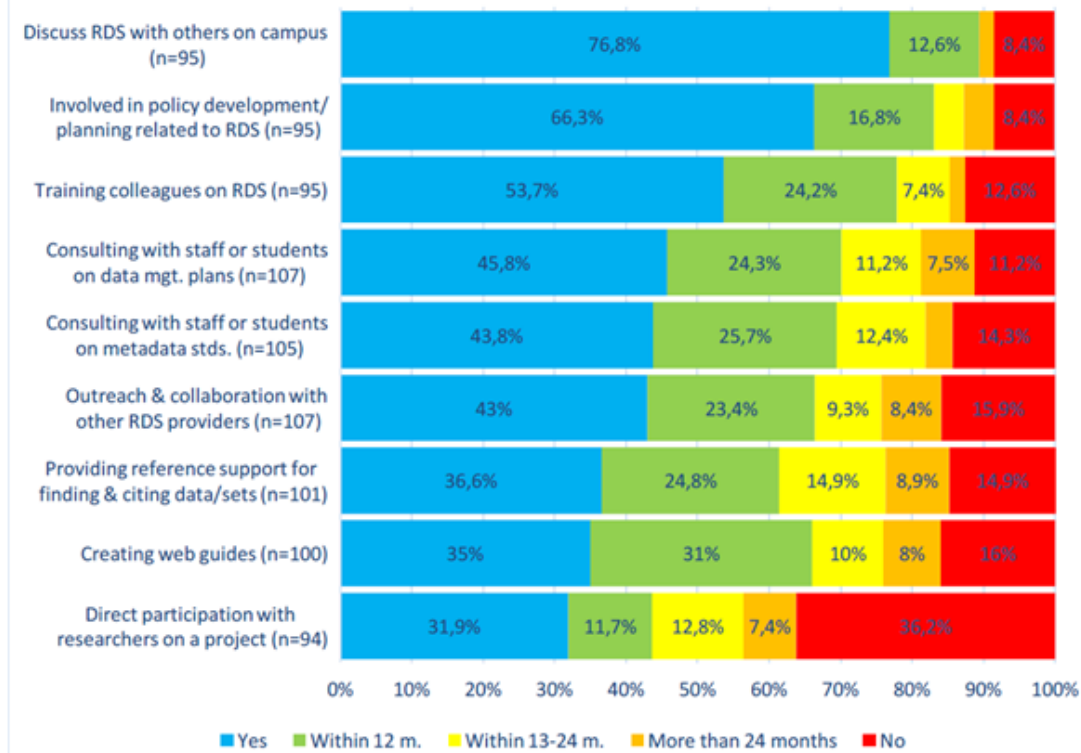
<sup>8</sup> [http://libereurope.eu/wp-content/uploads/PARSE-Insight\\_D3-5\\_InterimInsightReport\\_final.pdf](http://libereurope.eu/wp-content/uploads/PARSE-Insight_D3-5_InterimInsightReport_final.pdf)

Konsultatsiooniteenused (kuidas leida informatsiooni andmehaldusplaani koostamiseks, metaandmete standardeid, või viidata andmetele) haakuvad suures osas raamatukogudes juba varasemalt pakutud infoteenustega.

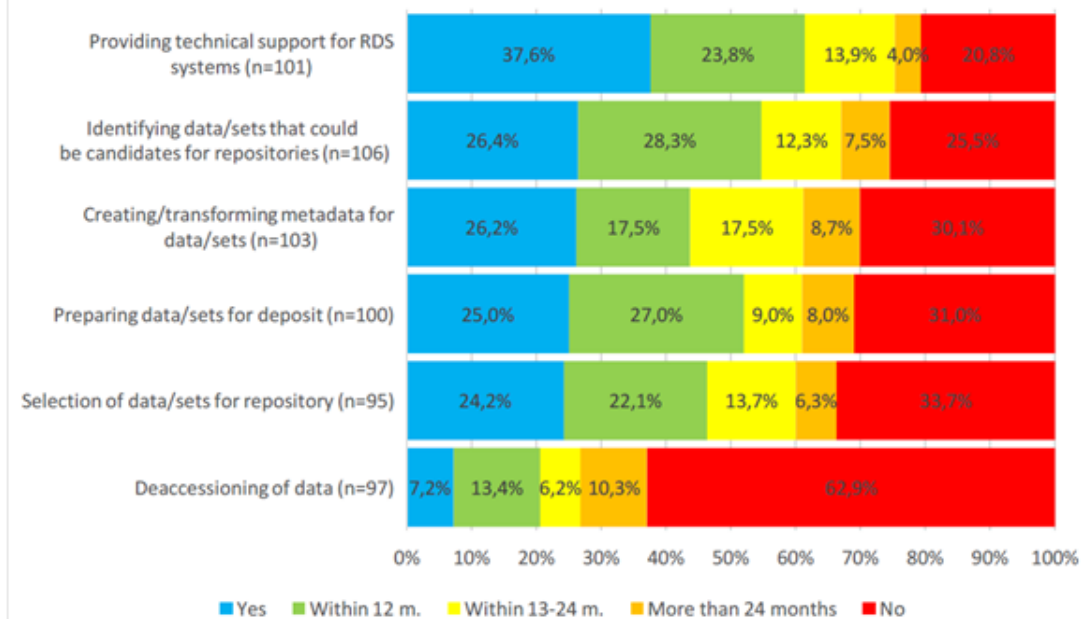
Euroopa raamatukogudes pakutavate teenuste hulka kuulub peamiselt andmete haldamise teenuste käsitlemine ning poliitike planeerimine ja väljatöötamine. Seni on andmete haldamise poliitika vähem kui pooltel raamatukogudel, mis näitab, et ollakse oma teenuste arendamisega suhteliselt varajases staadiumis. Andmete eemaldamist repositooriumist on planeerinud vaid mõni üksik raamatukogu, kuid arvestades arengu varast staadiumi, ei ole jõutud koguda piisavalt andmeid.

Joonis 3 näitab konsultatsioonitüüpi teenuste ja tehniliste teenuste olemasolu Euroopa raamatukogudes.

### Consultative Research Data Services (RDS)



### Technical Research Data Services (RDS)

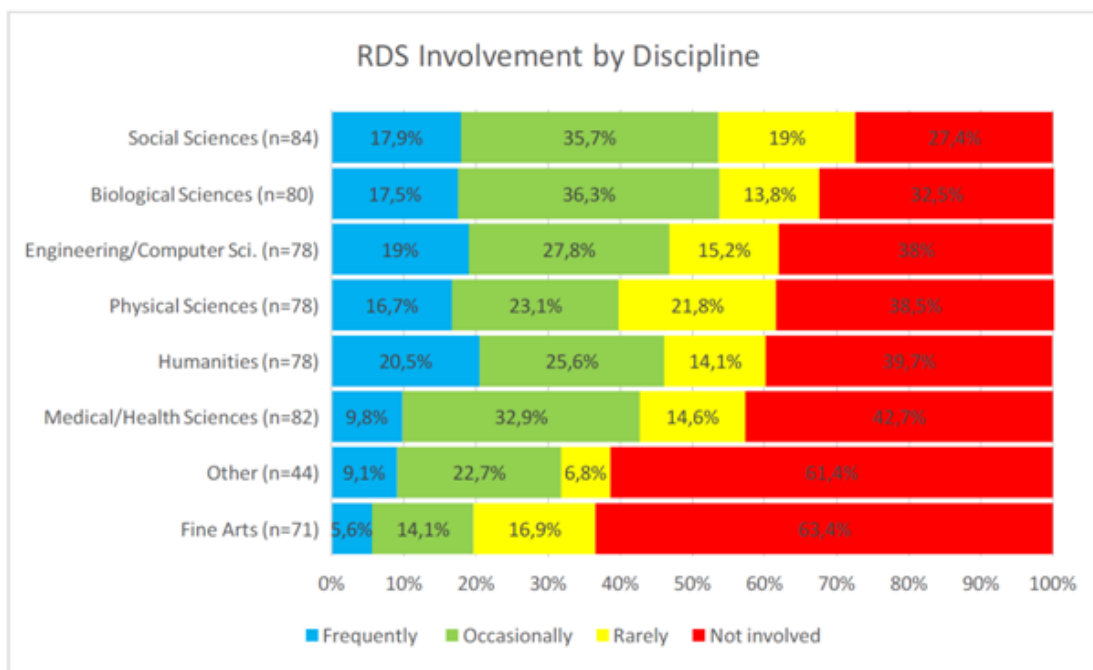


Labels not shown for values of less than five percent.

**Joonis 3.** Teadusandmete teenused. Allikas: *Data Services in European Academic Research Libraries. Carol Tenopir jt (2016)*

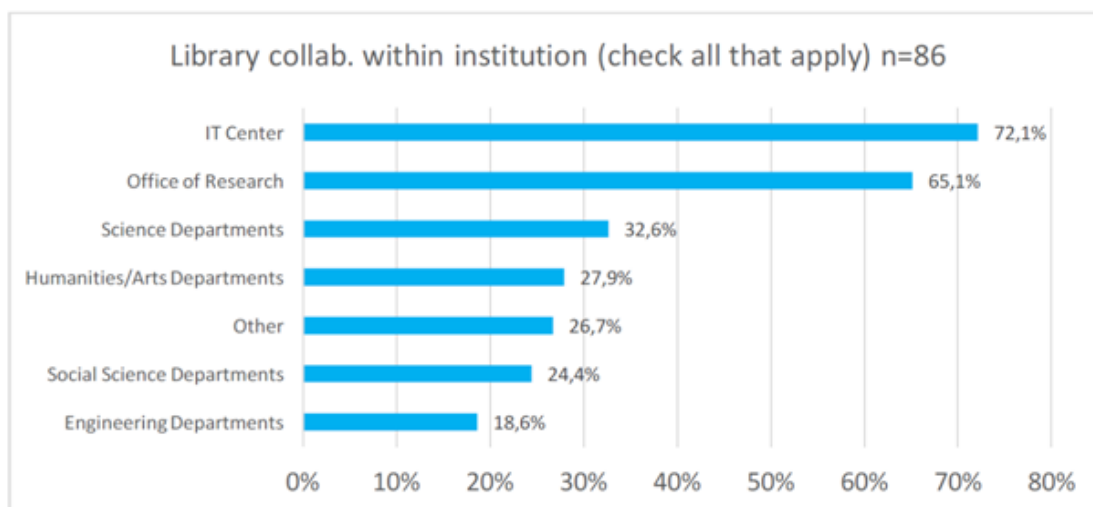
Euroopa raamatukogud pakuvad andmete haldamise teenuseid mitmetes valdkondades, kuid kõige enam sotsiaal- humanitaar- ja inseneriteaduste alal (joonis 4). Uurijatele tuli üllatusena, et uuringus osalenud raamatukogudes pakutakse väga vähe teenuseid meditsiiniteaduste valdkonnas. Sellele tuuakse kaks võimalikku põhjendust: (1) meditsiiniteaduste

uurimisrühmades on oma spetsialistid, kes tegelevad teadusandmete haldamisega või (2) uuringus ei olnud esindatud piisavalt meditsiiniraamatukogusid.

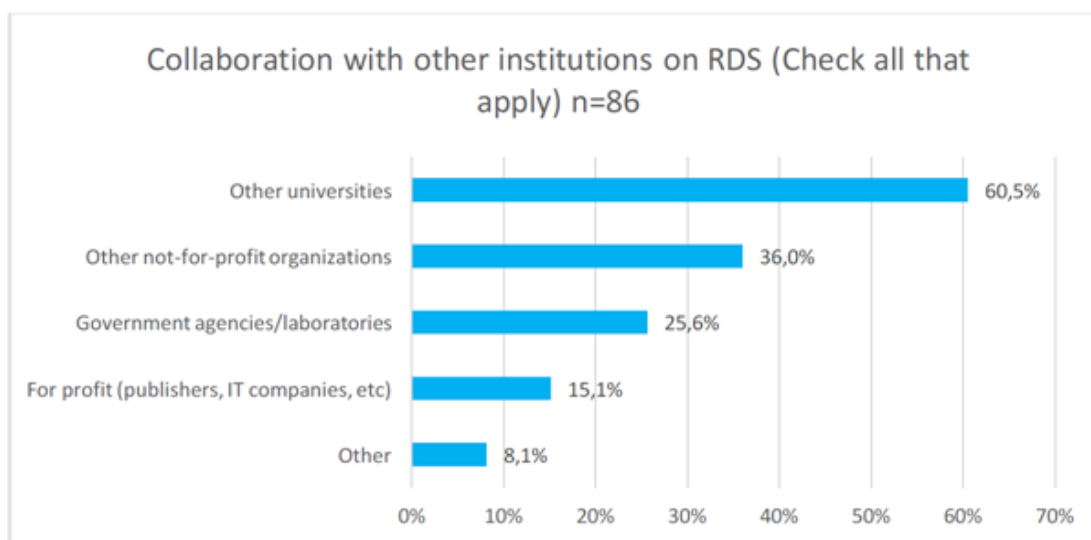


**Joonis 4.** Teadusandmete tugiteenused valdkondade lõikes. Allikas: Data Services in European Academic Research Libraries. Carol Tenopir jt (2016)

Üheks oluliseks valdkonnaks teadusandmete haldamisel on koostöö. Kuna teadusraamatukogude direktorid mõistavad, et vaid oma jõududega ei ole võimalik kõiki keerulisi teadusandmetega seotud probleeme lahendada, siis teevad raamatukogud koostööd mitmete institutsiooni siseste ja väliste partneritega (Joonistel 5 ja 6 on näha, millised on raamatukogude peamised koostööpartnerid). Koostöö ülikoolide ja teiste organisatsioonidega on väga oluline, kuna paljud Euroopa raamatukogud töötavad välja arengustrateegiat ning arutlevad võimaluste üle pakkuda paremaid teadusandmete haldamise teenuseid.



**Joonis 5.** Raamatukogude koostöö teiste üksustega. Allikas: Data Services in European Academic Research Libraries. Carol Tenopir jt (2016)



**Joonis 6.** Raamatukogude koostöö teiste asutustega. Allikas: Data Services in European Academic Research Libraries. Carol Tenopir jt (2016)

Andmete haldamise teenuste väljatöötamine eeldab töötajaid, kellel on valdkondlikke teadmisi ning võimalus omandada uusi oskusi. Paljud raamatukogud pakuvad oma töötajatele võimalust õppida teadusandmete haldamist, samas on raamatukogusid, kes võtavad sellele tööle juba vastava kvalifikatsiooni omandanud inimesi. Eeltoodust lähtudes tekib paratamatult küsimus, mis hetkel on vastuseta – kas raamatukogud on valmis loobuma mõningatest täna pakutavatest teenustest, et osutada kvaliteetsemat andmete haldamise teenust. Uuringu läbiviijad peavad seda oluliseks küsimuseks mida uurida.

Värskelt avaldatud COAR (*Confederation of Open Access Repositories*) raport „*Next Generation Repositories. Behaviours and Technical Recommendations of the COAR Next Generation Repositories Working Group*“<sup>9</sup> juhib tähelepanu avaandmete avalikustamise juures eelkõige meta-andmete olulisusele ja korrektsusele, eriti sotsiaalteadustes. Raportis tuuakse välja institutsionaalse võimekuse tõstmise vajadus, kuna läbi viidud uuringu tulemusena joonistub välja, et teadusasutustel valdavalt puudub avatud andmete põhimõtteid reguleeriv kord (sh andmete pikaajalise säilitamise kord). Samuti on ühe nõrkusena uuringus välja toodud tervet protsessi hõlmava tugisüsteemi puudumine. Uuring juhib tähelepanu ka sellele, et andmete standardiseerimise põhimõtted ja suuremahuliste andmemassiivide haldamine vajab vastava ekspertiisi välja arendamist.

Raportis tuuakse muuhulgas välja standardite ja mõtteviisi muutmise olulisus, mis lisaks andmearhiivide struktuuri ja tugiteenuste arendamisele innustab rohkem mõtlema ka andmete võimalike kasutajate kirjeldamisele ning andmete kasutamise meetrika väljatöötamisele.

### 1.3 Teiste riikide praktikad

Teiste Euroopa riikide avatud teaduse maastiku ning heade praktikate kirjeldamine on hea lähtekoht Eesti riikliku poliitika kujundamisel. Uuringu plaanis oli välja toodud Suurbritannia,

<sup>9</sup> <https://www.coar-repositories.org/files/NGR-Final-Formatted-Report-cc.pdf>



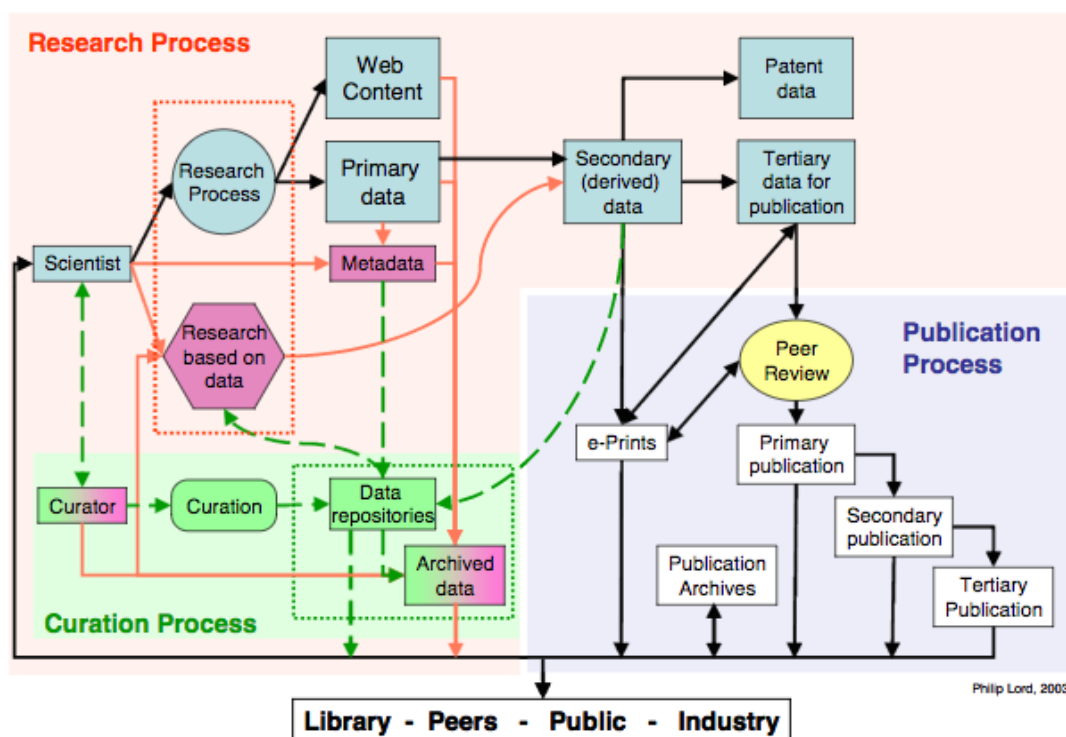
Hollandi, Taani, Poola ja Ungari praktikate lähem kirjeldamine. Dokumentide ja regulatsioonide läbitöötamise käigus lisandusid veel Soome, Iirimaa ja Sloveenia. Eelkõige on kirjeldamisel lähtunud riiklikust raamistikust, mida võimalusel on ilmestamas näide institutsionaalsest teaduse avaandmete protsessi praktikast.

### 1.3.1 Suurbritannia

Riiklikul tasemel on Suurbritannia vastu võtnud memorandumiga avatud teadusandmete haldamise sisuliste põhimõtetega (*Concordat on Open Research Data, 2016*)<sup>10</sup>, mille eesmärgiks on kindlustada, et kogutud ja loodud teadusandmed on avatud kõigile soovijatele. Teadusandmete avatusele on kehtestatud vastavad õiguslikud, eetilised kui ka valdkonnast tulenevad normid.

EPSRC (Engineering and Physical Sciences Research Council) loodud andmearhiiv on juba aastakümneid säilitanud sotsiaal- ja humanitaarteaduste andmeid. See on andnud riiklikul tasemel avatud andmete poliitika väljakujundamisel suure eelise. Täna on kõik Suurbritannia riiklikud rahastajad aktsepteerinud avatud teadusandmete poliitikat ning ülikoolid on vastavalt sellele kujundanud välja oma sisemised protsessid ning regulatsioonid.

Suuremates Suurbritannia ülikoolides on välja on arendatud nii tehniline võimekus kui ka tugisüsteem teadusandmete säilitamiseks ning avatud juurde pääsu tagamiseks (Joonis 7).



**Joonis 7.** Suurbritannia ülikoolide andmehalduse näidismudel (rakendatakse näiteks University of Essex, Kings College, UCL)<sup>11</sup>

<sup>10</sup> <http://www.rcuk.ac.uk/documents/documents/concordatonopenresearchdata-pdf/>

<sup>11</sup> <http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/150.pdf>

Leedsi Ülikoolis läbi viidud uuring näitas, et avaandmete temaatikat peetakse väga oluliseks nii teadlaste kui ka juhtkonna poolt. Uuring toob välja, et teadlased on avaandmeid kasutanud ning oma andmeid avaldanud erinevates repositooriumites (3/4 vastanutest). Avaandmete autorluse küsimuses pidas ligi 70% teadlastest andmetele viitamist sama oluliseks kui publikatsioonile viitamist ning 10% vastanutest pidas andmetele viitamist isegi olulisemaks kui artiklile viitamist.

### 1.3.2 Holland

Holland oli üks esimesi EL liikmesriike, kus algatati riiklikul tasemel teadusandmete repositooriumi loomine. Selle initsieeris peamiselt see, et ülikoolid või mõnel juhul ülikoolide koostööna oli hakatud looma oma repositooriumie, mis oli aga suhteliselt ressursimahukas. Samuti erines mitmel juhul andmekogude ja metaandmete struktuur, mis ei võimaldanud andmeid lihtsasti vahetada.

Peamised Hollandi deponitooriumid (sisaldavad ainult Hollandi TA asutuste teadusandmeid):

- NARCIS (<http://www.narcis.nl/>) – Hollandi teadusandmete portaal. Sisaldab kümneid tuhandeid teaduslikke andmestikke Hollandi ülikoolidelt ja uurimisasutustelt. Lisaks teadusandmetele sisaldab ka teaduspublikatsioone.
- DANS (*Data Archiving and Networked Services*) – ligipääs sotsiaal- ja humanitaarteaduste andmetele läbi EASY platvormi (<https://easy.dans.knaw.nl/>).
- 4TU.Centre for Research Data (4TU.ResearchData) (<http://researchdata.4tu.nl/home/>) – koostööarendus TU Delfti, TU Eindhoveni ja Twente Ülikooli vahel, mis võimaldab teadusandmeid säilitada ja tagada juurdepääsu (peamiselt tehnikavaldkonna uuringute andmed).
- DataverseNL (<https://dataverse.nl/>) - 2013. aastal loodud Hollandi ülikoolide ühine teadusandmete konsortsium teaduse avaandmete arhiveerimiseks ja avatud juurdepääsu võimaldamiseks.

2013. aasta veebruaris algatas *Data Archiving and Networked Services* (DANS) projekti CLIO-DAP, mille eesmärk oli parendada juhtivate majandus- ja sotsiaaljaloo valdkonna ajakirjades avaldatud publikatsioonidega seotud teadusandmete kättesaadavust. Projekti raames töötati välja standardid ja kirjeldati parimaid praktikaid teadusandmete avatud juurdepääsu võimaldamises.<sup>12</sup> See aitas suuresti kaasa ühtsete põhimõtete väljakujunemisele riiklikul tasandil.

Tänaseks rakendub Hollandis riikliku rahastuse toel läbiviidud teadustöö raames kogutud ja loodud andmetele kohustus hoida ja hallata teadusandmeid avalikult.<sup>13</sup> Erinevate rahastajate/programmide puhul on täpsemad nõuded erinevad:

---

<sup>12</sup><https://www.clariah.nl/en/about>

<sup>13</sup>[http://www.library.tudelft.nl/fileadmin/redacteur\\_upload/hulp-tips/publiceren/openaccess/FunderRequirementsonOpenScience.pdf](http://www.library.tudelft.nl/fileadmin/redacteur_upload/hulp-tips/publiceren/openaccess/FunderRequirementsonOpenScience.pdf)

NWO - the Netherlands Organisation for Scientific Research	<p>Pilootprogramm alates 1. jaanuar 2015.</p> <p>Teadlastel on kohustus lisada taotlusele andmehaldust kirjeldav osa ning positiivse rahastusotsuse puhul tuleb rahastajale esitada andmehaldusplaan.</p> <p>Pilootprogramm rakendub ainult osadele programmidele.</p>
ZonMw, the division for health research of NWO	<p>Andmehaldust kirjeldav osa lisatakse taotlusele ning positiivse rahastusotsuse puhul tuleb esitada andmehaldusplaan, mis sisaldab muuhulgas kuidas tagatakse andmetele juurdepääs ning andmete avatuks tegemise aega. Andmed, mis on kogutud projekti raames peavad olema dokumenteeritud ja säilitatud viisil, mis võimaldab nende avatuks tegemist.</p>
STW, the division for the Technical Sciences of NWO	<p>Tõstmaks teadustulemuste laialdasemat kasutamist eelkõige kolmandate osapoolte poolt on rahastaja liitunud andmehalduse pilootprogrammiga (vt NWO)</p>

Kõikide lähemalt vaadeldud ülikoolide (Leiden, Maastricht, Rotterdam, Delft) süsteem ja regulatiivne raamistik tugineb peamiselt rahastajate (riiklike ja Horisont 2020) regulatsioonidel. Ülikoolide veebilehtedel või ülikooli vastava valdkonna kordades on lisaks protsesside reguleerimisele väga selgelt kirjeldatud ülikooli tugisüsteemid ja erinevate osapoolte rollid. Väga suurt osa andmete haldamise tugisüsteemis mängivad raamatukogud (tugi andmete arhiveerimiseks ettevalmistamisel, andmehaldusplaani koostamisel jms). Teine oluline tugiroll on seotud projektide ettevalmistamise tuge pakkuva üksusega, kelle ülesandeks on projektitaotlustega seotud andmehaldusprotsessi kirjeldamine ning rahastaja valdkondlike reeglite vahendamine. Pikaajaline säilitamine ning andmete avalikustamine toimub riikliku andmerepositooriumi vahendusel. Teadlikkuse tõstmiseks on mitmel ülikoolil avaandmete temaatilised ajaveebid. Omavahelises vestluses on ülikoolide esindajad välja toonud kommunikatsiooni ja info kättesaadavuse olulisust.

Kiiret arengut avatud teadusandmete valdkonnas näitab pilguheit 2011. aastal (6 aastat tagas) Hollandi teadlaste seas läbi viidud uuringu tulemustele, mis näitasid, et Hollandi teadlased hoiustavad enda andmeid enamasti lokaalselt, juurdepääsu mitte võimaldaval moel. Süvaintervjuudel (n=32) ja ankeetküsitlusel (n=556) põhinev uuring<sup>14</sup> tõi välja, et andmete jagamine on üsna tavaline, eriti loodusteadlaste seas. Samas on sealgi teadlasi, kes näevad andmete avalikuks tegemises ohte, kaalutlusteks isikuandmete kaitse, omanikutunne andmete suhtes, soov ise publitseerida või hirm väärtõlgenduste ees. Toodi välja, et andmete jagamine ei paku jagajale (teadlase või institutsioonina) erilist kasu. Suur hulk, 70% küsitletud teadlastest ütles, et on kasutanud mujalt pärit andmeid, 61 protsendi sõnul on teised teadlased kasutanud nende andmeid.

<sup>14</sup> [https://pure.knaw.nl/portal/files/478516/The\\_Dutch\\_Datalandscape\\_DEF.pdf](https://pure.knaw.nl/portal/files/478516/The_Dutch_Datalandscape_DEF.pdf)

### 1.3.3 Taani<sup>15</sup>

2014. aasta veebruaris kutsusid Taani Rektorete Nõukogu, DeIC (*Danish e-Infrastructure Cooperation*) ja DEFF (*Denmark's Electronic Research Library*) kokku riikliku avaandmete töörühma, kuhu kuulusid lisaks Taani ülikoolide esindajatele ka Taani Kuningliku Raamatukogu, riikliku ja ülikoolide raamatukogu ja Taani Rahvusarhiivi esindajad. Töörühma eesmärgiks oli välja töötada riiklik teaduse avaandmete strateegia, mis omakorda suurendaks Taani teadlaste poolt loodavate ja kogutavate teadusandmete (ja seeläbi teaduse) paremat kättesaadavust ja levikut. Strateegia avaldati 2015. aastal.<sup>16</sup>

Strateegia kirjeldab ühtseid põhimõtteid ning tegevuskava teaduse avaandmete põhimõtete levitamiseks läbi teadlikkuse tõstmise ning erinevate uuringute läbiviimist. Strateegia määratles riiklikult rahastatud teadustöö käigus kogutud andmetele avatuse soovitus, mis viis selleni, et 2016. aastal alustati riiklikult rahastatud teadusprojektide avaandmete pilootprojektiga. Taani riiklik teadusrahastus (*Danske offentlige forskningsråd og fonde*) eeldab, et kogutud (ja ka ostetud/tellitud) kvantitatiivseid või kvalitatiivseid andmeid esitatakse dokumenteeritud kujul Taani Andmearhiivi.

Taani Andmearhiiv (<https://www.sa.dk/en/services/danish-data-archive>) kuulub Taani Rahvusarhiivi alla ning selle ülesandeks on uuringuandmete kättesaadavaks tegemine teadlastele ja üliõpilastele. Oluline on välja tuua, et teadusandmeid on lubatud taaskasutada ainult teaduslikel, statistilistel või ajaloolistel eesmärkidel. Andmekaitse eeskirjade kohaselt tuleb isikuandmed hävitada või anonümiseerida ning täpsemat andmete arhiveerimist (sh metaandmete struktuuri) reguleerib arhiveerimise eeskiri (*Statens Arkiver*).

Riiklikud avaandmete põhimõtted ning teadusandmete arhiveerimisvõimalused Taani Rahvusarhiivis on kättesaadavad Taani teadusportaalis.<sup>17</sup> Rahvusarhiiv säilitab teadusandmeid, mis on kogutud või loodud nii riikliku rahastuse kui ka erarahastuse toetusel.

Riiklikul tasemel kasvas 2015. aastal loodud töörühmast välja teadusandmete haldamise foorum (DeiC), kuhu kuuluvad esindajad kõikidest ülikoolidest ja teadusraamatukogudes ning andmearhiividest.<sup>18</sup>

Aastal 2017 avaldati Taani Teadusfondi läbi viidud uuringu<sup>19</sup> tulemused, mille käigus uuriti avatud teaduse põhimõtete rakendamist, sh. põhjuseid, takistusi ja motivaatoreid. Peamise põhjusena andmete avatuks tegemisel toodi välja, et andmetele ligipääs tagab parema kvaliteedi, nähtavuse ja usaldusvääruse, aga on ka lihtsalt hea teaduspraktika osa ning annab aluse koostööle. Samas tõid teadlased välja, et andmete avatuks tegemine on ajamahukas, võib põhjustada väärkasutamist või väärtõlgendamist ning teadlane/teadusgrupp kaotab publitseerimisvõimaluse. Motiveerivate teguritena toodi välja, andmete avatuks tegemisega peaks kaasnema täiendav rahastus, andmete kasutamisel publikatsioonide ühisautorlus või andmete tsiteerimise põhimõtete täiendamine, avaandmete tunnustamist akadeemilise

<sup>15</sup><https://openscience.fi/documents/14273/0/OSR-Denmark/4fcc4732-eacc-4c20-9060-c79c9b7c7116>

<sup>16</sup><https://www.deic.dk/sites/default/files/uploads/PDF/National%20Strategi%20for%20Forskingsdata%20Management%202015-2018.pdf>

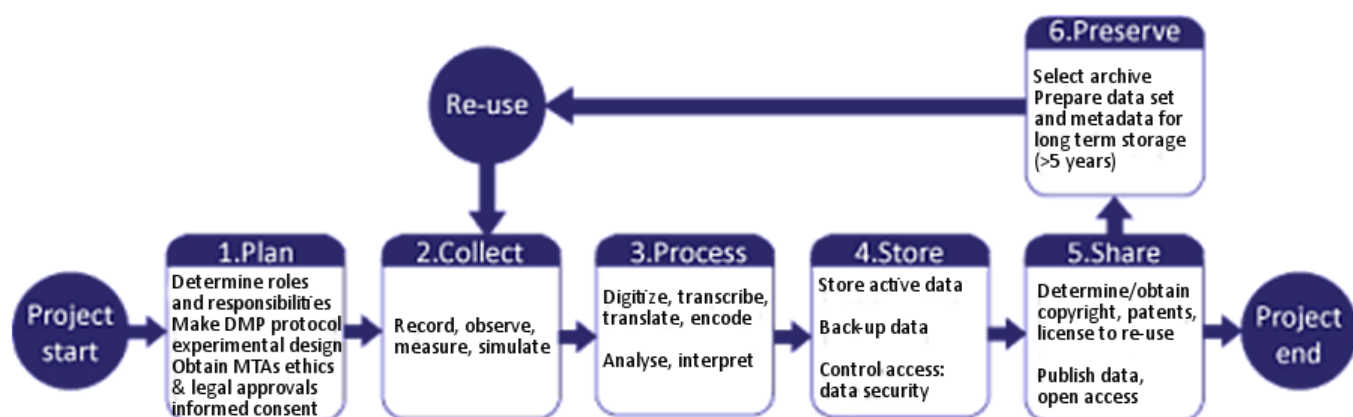
<sup>17</sup><http://forskerportalen.dk/en/research-data-management/>

<sup>18</sup>[https://www.deic.dk/datamanagement/DM\\_forum](https://www.deic.dk/datamanagement/DM_forum)

<sup>19</sup><http://dg.dk/wp-content/uploads/2017/11/Open-Access-to-Data-%E2%80%93-It%E2%80%99s-not-that-Simple.compressed.pdf>

profiili hindamisel. Uuringu raport toob välja ka, et avatud andmete põhimõtete rakendamine erineb valdkonniti olulisel määral.

Taani ülikoolidest toome näitena Kopenhaageni Ülikooli, kus teadusandmete haldamise kord reguleerib andmete kogumist, säilitamist, avalikuks tegemist kui ka tundlike andmete haldamist ning andmete kuuluvust. Reguleerimine võimaldab valdkonniti/teaduskonniti täpsemad korrad kehtestada, et valdkondlike erisusi paremini reguleerida. Tervise- ja meditsiiniteaduskonna teadusandmete säilitamise ja juurdepääsu võimaldamise korras on välja toodud täpsem andmehalduse skeem, mis on abiks teadustöö planeerimisel ja läbiviimisel (Joonis 8).



**Joonis 8.** Kopenhaageni Ülikooli tervise- ja meditsiiniteaduskonna andmehaldusprotsessi skeem<sup>20</sup>

Ülikooli tasandil on loodud andmehalduse nõukogu, mis teostab järelvalvet andmehalduse protsesside elluviimisele ning teeb ettepanekuid protsessi täiendamiseks.

#### 1.3.4 Soome

Soome Haridus- ja kultuuriministerium on käivitanud avatud teaduse initsiatiivi (ATT)<sup>21</sup> edendamaks uuringuandmete kättesaadavuse ja avatud teaduse platvormi väljaarendamist aastatel 2014-2017. Selle käigus kehtestatakse riiklikult rahastatud teadusuuringute raames kogutud andmetele avatud juurdepääsu nõue. Avatud teadusandmete algatuse peamiseks eesmärgiks on Soome teadusuuringute konkurentsivõime ja kvaliteedi tõstmine ning läbipaistvuse ja innovatsiooni propageerimine.

Valitud näitena on Turu Ülikoolis välja arendatud avaandmete kord ja protsess.<sup>22</sup> Turu Ülikooli avaandmete kord kirjeldab ülikooli põhimõtteid ja poliitikat, mis on seotud andmete kogumise, kasutamise ja säilitamisega. Avaandmete kord puudutab digitaalseid andmeid, mille hulka võivad kuuluda mõõtmistulemused, statistika, digitaalsed pildid, transkribeeritud intervjuud, uuringute andmed või märkmed, mis põhinevad välitöödel.

<sup>20</sup> <http://healthsciences.ku.dk/>

<sup>21</sup> <http://www.aka.fi/en/funding/responsible-research/open-science/>

<sup>22</sup> <https://www.utu.fi/en/news/Documents/datapolitiikka-en-2.pdf>

Turu Ülikooli avaandmete projekti juht Mari Riipinen on öelnud omavahelises vestluses, et ülikool on läbi viinud erinevaid küsitlusi ning intervjuusid nii teadlaste kui ka doktorantidega ning kaardistanud ära kõige suuremad vajadused, sh infrastruktuuri vajadused andmete pikaajaliseks säilitamiseks ning koolitusvajaduse avaandmetepõhise andmehaldusplaani tegemiseks. Koolitusstrateegia peamised sihtrühmad on kraadiõppurid ja teadlased, kuid ka tugitöötajad, kelle roll on toetada teadustöö läbiviimist spetsiifilistes avaandmetega seotud küsimustes. Avaandmete teemal on planeeritud välja arendada kohustuslik ainekursus doktorantidele ning spetsiifiline koolitusprogramm uutele teadustöötajatele.

Avaandmete põhimõtete kujundamise ja rakendamise protsess oli väga kaasav ning Mari Riipinen on välja toonud olulised tähelepanekud:

Avaandmetega seotud protsesside väljakujundamisse kaasati ülikooli töötajaid erinevatest valdkondadest ja erinevatelt tasanditelt, mis tagas läbipaistvuse ning kõikide osapoolte kaasatuse;

Väga oluline roll regulatsiooni kujundamises oli andmetega seotud õiguslikud suhted, sh andmete looja ja ülikooli vahel;

Infrastruktuuri väljaarendamise ning kasutamise protsessi kirjeldamise üks osa oli teadlaste ja doktorantide koolituskava;

Võtmetegevusena on protsessi väljaarendajad märkinud kommunikatsiooni nii ülikooli sees kui ka väljapoole ülikooli – partneritele, avalikkusele, rahastajatele.

### 1.3.5 Sloveenia

Juba Sloveenia Teadus-ja innovatsioonistrateegia 2011-2020 sätestas, et avaliku sektori rahastamise toel kogutud teadusandmed peavad saama vabalt kättesaadavateks ja Sloveenia teaduse infrastruktuuri arengukavas 2011-2020 sätestati nii teaduspublikatsioone kui teadusandmeid sisaldava rahvusliku avaandmete infrastruktuuri loomine seotuna Sloveenia teaduse informatsioonisüsteemiga SICRIS (Eesti ETIS analoog) kooskõlas ESFRI teekaardiga.

Nende põhimõtete realiseerimiseks moodustas Sloveenia Hariduse, teaduse ja spordiministeerium 2014. a erinevate osapoolte esindajatest töörühma rahvusliku avatud juurdepääsu strateegia "*National Strategy of open Access to Scientific Publications and Research Data in Slovenia 2015-2020*"<sup>23</sup> väljatöötamiseks ja kasutas selleks tööks EL PASTEUR 4OA projekti (<http://pasteur4oa.eu>) vahendeid. Strateegia esimene mustand valmis 2015. a maikuu, millele järgnesid avalikud arutelud, dokumendi kooskõlastamine erinevates ministeeriumides ja kinnitamine Sloveenia valitsuses 2015. a septembris. Sellele järgnes avaandmete pilootprogramm, mille tulemuste põhjal koostatud avaandmete strateegia tegevuskava kinnitati 2017. a maikuu. Avaandmete strateegia tegevuskava, on momendil kättesaadav vaid sloveeniakeelsena.

Sloveenia avatud teaduse strateegia põhiseisukohad on järgmised:

---

<sup>23</sup>[http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National\\_strategy\\_for\\_open\\_access.pdf](http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Zakonodaja/Strategije/National_strategy_for_open_access.pdf)

1. Avaliku sektori rahastuse toel ilmunud eelretsenseeritavad teaduspublikatsioonid peavad olema avatud juurdepääsuga ja tungivalt soovitatav on tagada vaba juurdepääs ka kõigile muudele teaduspublikatsioonidele.
2. Eesmärk on, et 80% 2017.a riikliku rahastuse toel valminud teaduspublikatsioonidest oleksid avatud ligipääsuga 2018. aastaks ning et 100% 2020. aasta analoogilistest publikatsioonidest oleksid avatud ligipääsuga 2021. aastaks.
3. Sloveenias baseeruvate ja oma tegevuseks rahvuslikest vahenditest toetust saavate kirjastuste publitseeritavad teadusajakirjad peavad olema nn kuldse avatud juurdepääsuga.

Väärib märkimist, et vaatamata juba 2010. aastal alanud avatud juurdepääsu poliitika väljakujundamisele Sloveenias ja viidatud strateegiadokumentide olemasolule ei olnud 2017. aasta maikuuks kõik Sloveenia teadusasutused ja -organisatsioonid Avatud teaduse poliitikat veel ametlikult aktsepteerinud.

Seejuures võib tõdeda, et nagu enamuses teisteski EL riikides on ka Sloveenia edukam olnud teaduspublikatsioonidele avatud ligipääsu tagamisel, aga keerukam on tänaseni olukord teadusandmetele avatud juurdepääsu tagamisega.

### 1.3.6 Iirimaa<sup>24</sup>

Iiri Teadusorganisatsioonide Komitee avalikustas 2012. a dokumendi "Rahvuslikud avatud juurdepääsu poliitika põhimõtted" (*National principles for Open Access policy statement*)<sup>25</sup>, mis hakkasid kehtima 2013. aastast.

Üldosa sätestab, et see poliitika

- kinnitab teadlaste vabadust publitseerida oma teadustöö tulemusi kohtades, kus nad peavad seda sobivaks;
- on mõeldud selleks, et suurendada Iiri Vabariigi poolt rahastatud teadustöö tulemuste nähtavust ja juurdepääsu teadustööle pärast selle avaldamist;
- on kujundatud toetamaks info vaba liikumist rahvuslike ja rahvusvaheliste teaduskogukondade vahel;
- on mõeldud toetama teaduspõhise õpetamise ja õppimise põhimõtteid ning avatud haridusressursside (*Open Educational Resources, OER*) loomist;
- täiendab ja kindlustab (tugevdab) Iiri teadusrahastuse varasemat avatud juurdepääsu poliitikat ja praktikat, Iiri teadusasutusi, raamatukogusid ja repositooriume.

Põhimõtted ise on järgmised:

<sup>24</sup> <http://openaccess.thehealthwell.info/>

<sup>25</sup> <http://www.dri.ie/sites/default/files/files/National%20Principles%20on%20Open%20Access%20Policy%20Statement%20%28FINAL%2023%20Oct%202012%20%29.pdf>

1. Eelretsenseeritud ajakirjade artiklid ja muud teadustulemused, mis on saanud kas täieliku või osalise riikliku rahastuse, peavad olema deponeeritud avaliku juurdepääsuga repositooriumis ning tehtud avalikult leitavaks, juurdepääsetavaks ja taaskasutatavaks.
2. Repositooriumid peavad avaldama metaandmed kohe peale deponeerimist. Avatud juurdepääs artikli täistekstile tuleb anda kohe peale deponeerimist või kohe peale artikli ilmumist.
3. Teadlasi innustatakse publitseerima avatud juurdepääsuga ajakirjades, kuid see ei asenda artikli deponeerimist repositooriumisse. Ka avatud juurdepääsuga ajakirjades avaldatud artiklid tuleb deponeerida repositooriumis.
4. Repositoorium loetakse sobivaks juhul, kui see võimaldab vaba juurdepääsu sisule, toetab koostööd teiste repositooriumite või teadusinfo süsteemidega, on ühendatud rahvuslike ja rahvusvaheliste portaalidega ning toetab pikaajalist säilitamist.
5. Teadusandmed peavad olema deponeeritud alati, kui see on võimalik, ning lingitud nendega seotud artiklitele, kui see on sobiv/soovitav. Teadusandmete puhul tuleb arvesse võtta nii Euroopa kui ka riiklike andmekaitse reegleid. Minimaalselt tuleb deponeerida teadusandmete metaandmed, mis kirjeldavad andmeid, nende asukohta ja juurdepääsuvõimalusi.

2017. aasta juulis avaldas liri valitsuse kapitalimahutuste ja reformide osakond Avatud andmete strateegia aastateks 2017-2020 (*Open Data Strategy 2017 - 2022*)<sup>26</sup>. Dokumentis on sõnastatud avatud andmete strateegia mis on raamistikuks ka teaduse avaandmetele. Dokument sedastab, et "Teaduskogukonnal on võtmeroll avatud andmete kasutamise majanduslike, sotsiaalsete ja demokraatlike võimaluste väljaselgitamisel, avatud andmete kultuuri edendamisel ning asjakohase andmete haldamise praktika väljatöötamisel eestvedaja ja toetajana."

Tegevused, mida teaduskogukonnalt oodatakse on sõnastatud järgmiselt:

- kindlustavad avatud juurdepääsu avalikult rahastatud teadusuuringute toorandmetele;
- kasutavad avatud andmeid ja aitavad edendada avatud andmete kasutuselevõttu;
- selgitavad välja avatud andmete rahvuslikud/rahvusvahelised trendid;
- arendavad välja mõõdikud tulemuste saavutamiseks/hindamiseks;
- selgitavad välja vajalikud uued oskused, et tagada üliõpilaste edu andmepõhises majanduses;
- selgitavad välja kuidas kodanikud saaksid avatud andmetele ligi pääseda ja neid kasutada;

---

<sup>26</sup> [https://data.gov.ie/sites/default/files/files/Final%20Strategy%20online%20version\(1\).pdf](https://data.gov.ie/sites/default/files/files/Final%20Strategy%20online%20version(1).pdf)



- selgitavad välja avatud andmete abil läbipaistvuse ja vastutuse suurendamise võimalused;
- tutvustavad avatud andmete arhiveerimise, turvalise hoiustamise ja säilitamise, juurdepääsu ja jagamisega seotud parimaid praktikaid.

### 1.3.7 Poola

Aastal 2015 avaldatud raport *“Open Science in Poland 2014. A Diagnosis.”*<sup>27</sup> avab põhjalikult regulatiivse raami, mille järgi Poola riiklikku avatud teaduse poliitikat on kujundama hakatud. Eelkõige lähtub riiklik plaan Horisont2020 rahastusmehhanismiga kaasnevatest regulatsioonidest ning sellest tulenevast muudatuste vajadusest.

Poola teadus- ja kõrgharidusministeerium algatas teadlaskonna teavitamise avatud teaduse põhimõtetest juba 2004. Aastal, kohe pärast OECD *Declaration on Access to Research Data from Public Funding* allkirjastamist Poola valitsuse poolt. Avatud teaduse ja järjest rohkem selle raames ka avatud teadusandmete põhimõtete peamiseks eesmärgiks on läbipaistvus, vastutus, professionaalsus, kvaliteet, efektiivsus, kuid ka intellektuaalse omandi raamistik ja andmete selgem kaitse.

Teisalt aga seisab 2015. aastal avaldatud raportis, et riiklik teadusrahastus ei ole toetanud avatud teaduse rakendamist, kuigi sellele on juba 2013. aastal tähelepanu juhtinud Poola Rektorate Nõukogu koos Poola Teaduste Akadeemiaga nende ühises pöördumises: *“OA as the primary model of sharing research publications; in order to succeed, OA needs to be included in assessments and metrics; recommending depositing in OA repositories + granting non-exclusive licenses to publishers.”* (Bolikowski 2014)

2014. aasta aprillis on Poolas loodud 23 institutsionaalset teadusandmete repositooriumi, lisaks kasutatakse DSpace, OpenAIRE ja Figshare platvorme. Repositooriumite haldajate hulgas viidi läbi küsitlus, et kirjeldada arhiveeritavate andmete mahtu, iseloomu ning avatust (täpsemad tulemused raportis *“Open Science in Poland 2014. A Diagnosis.”* lk 49-52, kus on loetletud kõikide kasutatavate repositooriumite info). Uuringu tulemused näitavad, et teadusandmete säilitamine on killustunud ning ühtse riikliku depositeerimise puudumine on viinud selleni, et teadlaskond eelistab rahvusvahelisi valdkondlikult välja kujunenud platvorme.

Mitmetel ülikoolidel on loodud oma valdkondlikud depositeerimised. Näitena võib tuua Warssawi ülikooli, millel on mitmes valdkonnas teadlaste enda poolt hallatavad valdkondlikud andmete repositooriumid.

Raport toob välja ka avatud teaduse ühe tööriistana teaduse ajaveebi, mis on kogunud Poolas laialdaselt populaarsust ning mis on avaandmete põhimõtete kajastamisel olulist rolli mänginud.

### 1.3.8 Ungari

On tähelepanuväärne, et Avatud teaduse põhimõtete ajaloost rääkides meenutatakse valdavalt nn. Budapesti Avatud Teaduse Initsiatiivi (*The Budapest Open Access Initiative*

<sup>27</sup> [http://pon.edu.pl/images/plon\\_publications/files/16\\_Open%20Science%20in%20Poland%202014.pdf](http://pon.edu.pl/images/plon_publications/files/16_Open%20Science%20in%20Poland%202014.pdf)

BOAI). 2001. a detsembris korraldas Budapesti Avatud Ühiskonna Instituut esimese avatud teaduse probleemistikule pühendatud konverentsi, mille tulemusena formuleeriti Avatud Teaduse põhimõtted, mis saidki tuntuks kui Budapest *Open Access Initiative*. Selle tulemusena on välja arenenud tänapäevane Avatud Teaduse poliitika Euroopas ja ka laiemalt. Selle taustal on tähelepanuväärne, et Euroopa Liidus Avatud Teaduse poliitika edendamiseks loodud PASTEUR 4OA projekti käivitamise raames 2014. a läbiviidud uuringu “*Hungary Open Access Case Study*” kokkuvõtte alguses tõdetakse, et Ungaris on vähe valdkonda reguleerivaid poliitikadokumente, ja kuigi infrastruktuuri osas on avatud juurdepääsu olukord üsna hea, siis teadusandmete avatuse osas hinnatakse olukorda väljendiga “dire”(kohutav!). Märgitakse, et Ungari kui väikese riigi puhul oleks EL tasemel koordineeriv tegevus Avatud Teaduse poliitika kujundamisel hädavajalik.

### Kokkuvõte

Vaadeldud riikide lõikes võib öelda, et suundumused on samad ning seda kujundab eelkõige Euroopa Liidu avatud teaduse suunaline poliitika ning rahastusmeetmete nõuded. Riigid ja ülikoolid, kes sõltuvad rohkem EL rahastusest on hakanud vastavaid meetmeid teaduse avaandmete toetamiseks ja reguleerimiseks varasemalt rakendama ning protsessid on paremini välja arenenud (Tabel 1).

**Tabel 1. Vaadeldud riikide võrdlus osapoolte rollide lõikes**

	Euroopa Liit	UK	Holland	Taani	Soome	Iirimaa	Poola	Ungari
<b>Raamatu-kogude roll</b>	Valdavalt oluline	Oluline	Oluline (jagatud)	Ebamäärane	Ebamäärane	Ebamäärane	Ebamäärane	Ebamäärane
<b>Riiklikud regulatsioonid</b>	Varieeruv	Olemas	Olemas	Olemas	Olemas	Olemas	Protsessis	Varieeruv
<b>Institutsionaalne valmisolek</b>	Varieeruv	Olemas	Olemas (protsessis)	Olemas (protsessis)	Varieeruv	Varieeruv	Varieerub	Varieeruv
<b>Seadusandlus</b>	Varieeruv	Olemas	Olemas	Olemas	Olemas (protsessis)	Varieeruv	Osaliselt	Varieeruv
<b>Rahastaja</b>	Varieeruv	Toetab	Toetab (piloot)	Toetab (piloot)	Toetab	Toetab	Ebamäärane (osaliselt)	Varieeruv

## 2 Veebipõhine küsitlusuuring ja fookusgruupiintervjuud

*Lähteülesanne: Avatud teaduse kasutajaskonna praeguste praktikate ja vajaduste kaardistamine. Viime läbi 10 fookusgruupiintervjuud sihtrühmade lõikes: a) institutsioonid, ülikoolid, b) teadlaskond (4 põhivaldkonna lõikes), c) ettevõtlussektor – aluseks nutika spetsialiseerumise jaotus ning arvestades ettevõtete ja organisatsioonide tüüpe (sh suur-, keskmised, väikeettevõtted, innovaatilised, innovatsiooni toetavad organisatsioonid jt); d) avaliku sektori esindajad, sh rahastajad ja poliitikakujundajad, e) kolmas sektor ja vabavahendused, f) teenusepakkujad. Osad fookusgruupiintervjuud on sihtrühmapõhised, osadesse on kaasatud erinevate sihtrühmade esindajad. Küsitluse ja intervjuude tulemused on sisendiks TÜ rühmale, et analüüsida eri praktikate õiguslikke ja majanduslikke aspekte.*

### 2.1 Metoodika

Uuring koosnes kahest peamisest osast: küsitlus Eesti teadlaste seas ning fookusgruupi- ja individuaalintervjuud teadlaste ja teiste huvirühmade esindajatega.

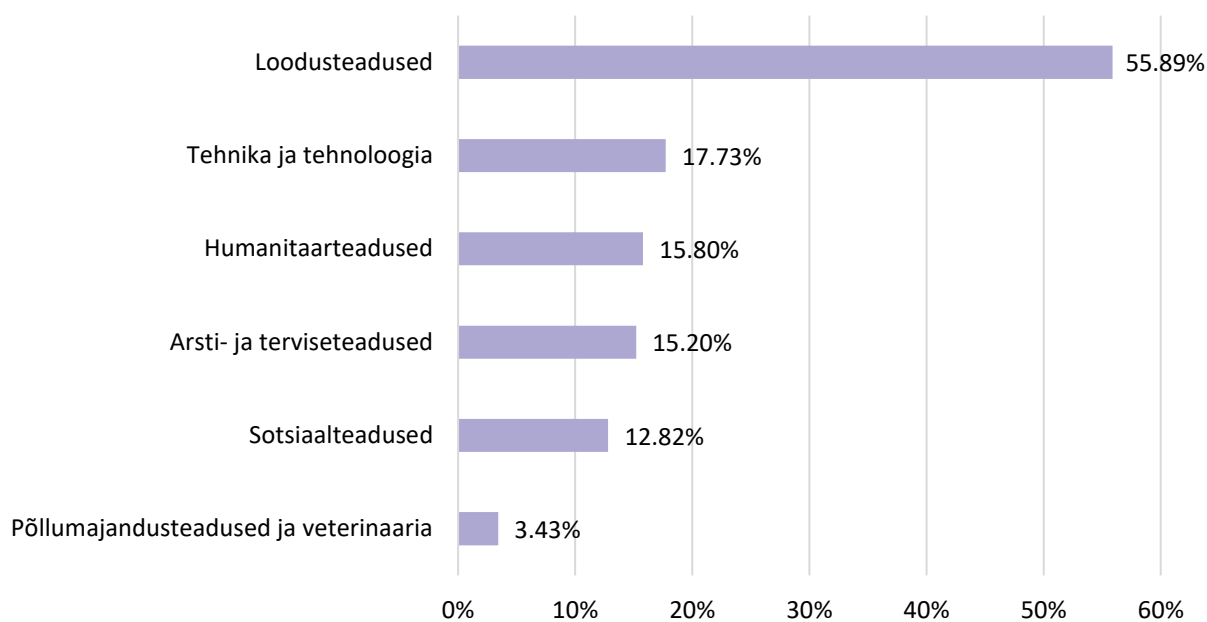
#### 2.1.1 Veebiküsitlus

Veebiküsitluse eesmärgiks oli selgitada avatud teaduse alaseid praktikaid ja hoiakuid Eesti teadlaste seas. Üldkogumi moodustasid ETISE andmebaasis olevad inimesed, kes on viimase viie aasta jooksul osalenud täitjana mõnes projektis (nii Eesti kui rahvusvahelised projektid), seisuga märts 2017.

Küsitlustöö toimus ajavahemikul 21.03-3.04-2017. Kõigile üldkogumi liikmetele saadeti välja 4233 e-kirja, millest veateatega tuli tagasi 200. Kokku jõudis uuringukutse 4033 teadlaseni, kellest vastas 671 teadlast erinevatest teadusvaldkondadest (Tabel 2 ja Joonis 9). Uuringu üldine vastamismäär oli seega 15,8%. Küsimustik sisaldas 33 kinnist ja 5 avatud küsimust. Tulemuste analüüsiks kasutati SPSS statistikapaketti. Üldkogumi suurust ja vastajate arvu arvestades ei ületa tulemuste hinnanguviga kõigi vastajate lõikes  $\pm 3,5\%$  piiri (Tabel 3).

**Tabel 2. Vastajate arv valdkonniti**

Valdkond	Vastanute arv
Loodusteadused	375
Tehnika ja tehnoloogia	119
Arsti- ja terviseteadused	102
Põllumajandusteadused ja veterinaaria	23
Sotsiaalteadused	86
Humanitaarteadused	106
Vastajad kokku	671
Üldkogum (N)	4233



**Joonis 9.** Küsimustikule vastajate osakaalud valdkondade lõikes

**Tabel 3.** Teadlaste veebiküsitluse tulemuste veapiirid.

Üldkogum	ETIS andmebaasis olevadteadlased, kes on viimase viie aasta jooksul osalenud täitjana mõnes projektis (nii Eesti kui rahvusvahelised projektid), seisuga märts 2017.											
Hinnangu jaotus	50%	45%	40%	35%	30%	25%	20%	15%	10%	5%	3%	2%
Vastajate arv	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	97%	98%
25	20,54%	20,44%	20,12%	19,59%	18,82%	17,79%	16,43%	14,67%	12,32%	8,95%	7,01%	5,75%
40	15,89%	15,81%	15,57%	15,16%	14,57%	13,76%	12,72%	11,35%	9,54%	6,93%	5,42%	4,45%
80	11,07%	11,02%	10,85%	10,56%	10,15%	9,59%	8,86%	7,91%	6,64%	4,83%	3,78%	3,10%
100	9,88%	9,83%	9,68%	9,43%	9,06%	8,56%	7,90%	7,06%	5,93%	4,31%	3,37%	2,77%
120	8,91%	8,86%	8,73%	8,50%	8,16%	7,71%	7,13%	6,36%	5,35%	3,88%	3,04%	2,49%
150	7,86%	7,82%	7,70%	7,50%	7,20%	6,81%	6,29%	5,61%	4,72%	3,43%	2,68%	2,20%
200	6,76%	6,73%	6,63%	6,45%	6,20%	5,86%	5,41%	4,83%	4,06%	2,95%	2,31%	1,89%
250	6,01%	5,98%	5,89%	5,74%	5,51%	5,21%	4,81%	4,29%	3,61%	2,62%	2,05%	1,68%
350	5,02%	4,99%	4,92%	4,79%	4,60%	4,34%	4,01%	3,58%	3,01%	2,19%	1,71%	1,40%
400	4,66%	4,64%	4,57%	4,45%	4,27%	4,04%	3,73%	3,33%	2,80%	2,03%	1,59%	1,31%
500	4,12%	4,10%	4,03%	3,93%	3,77%	3,56%	3,29%	2,94%	2,47%	1,79%	1,40%	1,15%
671	3,47%	3,45%	3,40%	3,31%	3,18%	3,01%	2,78%	2,48%	2,08%	1,51%	1,18%	0,97%

### 2.1.2 Individuaal- ja fookusgrupp-intervjuud.

Uuringu teises osas viisime juunist detsembrini 2017 läbi 12 fookusgrupi- ja individuaalintervjuud, milles osales kokku 38 inimest. Valim moodustati eesmärgipärasena. Eesmärgiks seati avatud teaduse elluviimise seisukohalt oluliste siht- ja sidusrühmade esindajate kaasamine, kelle on otsene või potentsiaalselt mõjukas puutumus avatud teaduse kontseptsiooni ellu rakendamisse. Valimise kuuluvus määrati projektigrupi ekspertide arutelu tulemusel. Teadlaste sihtrühma intervjuudesse kutsumisel jälgiti, et kaetud oleks erinevate

teadusvaldkondade mõjukad teadlased (vähemalt vanemteaduri tasemel), kelle kogemus või potentsiaal avaandmete loomiseks. Valimisse kaasati sobivusel ka veebiküsitluses eelneva nõusoleku andnud teadlasi.

Intervjuudes on esindatud eri valdkondade teadlased (valdavalt vanemteaduri ja professori tasandil), ülikoolide raamatukogude esindajad, ülikoolide prorektorid ja teadusosakondade töötajad, rahastajate esindajad (Eesti Teadusagentuur) ning valdkonna otsustajad (haridus- ja teadusministeeriumi esindajad).

Fookusgruppintervjuude kestvus oli 1,5 kuni 2 tundi ning paaris- ja individuaalintervjuudel 0,5 kuni 1 tund. Kõigi sihtrühmade intervjuudes käsitleti järgmisi teemasid: 1) refleksioon avatud teaduse idee ja võimalikule mõjule; 2) senine praktika open-access artiklite ja avaandmetega nii avaldaja, kasutaja kui haldaja perspektiivist; 3) OA ja AA kättesaadavus (respositooriumid täna ja perspektiivis); 4) ajendid ja barjäärid kasutamisel. Täiendavalt sihtrühma spetsiifilised küsimused, nt kommentaarid ETAG-i juures oleva ekspertkomisjoni poolt väljatöötatud soovitudele nende elluviimise perspektiivist, autoriõigused, litsentsitasud, teadlaste karjäärimudel, soovitud poliitikakujundajatele, ülikoolide koostöö ettevõtetega, riiklikud avaandmed.

Intervjuud salvestati vastajate teadlikul nõusolekul. Vastavalt kokkuleppele vastajatega pöörasid uurijad suurt tähelepanu vastajate anonüümsuse tagamisele kogu tööprotsessi vältel, sh uuringu tulemuste avaldamisel. Teadlaste, raamatukogude ning ülikoolide esindajate intervjuude salvestustest tehti täielikud mahakirjutused. Kogutud materjalide süstematiseerimisel kasutati temaatilist kodeerimist.

Paraku osutus keerukaks teadusväliste kasutajarühmade kaasamine fookusgrupp uuringusse, saime arvukalt huvi- ja ajapuudusele viitavaid äraütlemisi. Selle tõttu kompasime 2017. aasta oktoobris nende rühmade hoiakuid kahe eraldi veebiküsitlusega, millest üks oli suunatud ministeeriumite teadusnõunikele ja teine vabaühendustele. Kaheksast teadusnõunikust vastas küsitlusele seitse. Vabaühendustele suunatud küsimustikku levitasime Eesti Mittetulundusühingute ja Sihtasutuste Liidu infokanalite kaudu ja saime samuti seitse vastust.

Küsitlused ja intervjuud viis läbi Tallinna Ülikooli töörühm koosseisus Esta Kaal, Arko Olesk ja Kristel Toom.

## 2.2 Uuringu tulemused

Avatud teadus on suundumus, mille vajalikkust ja võimalusi teadusele ning ühiskonnale laiemalt tajuvad kõik teemaga seotud osapooled, hoiakud selle suhtes on valdavalt soosivad. Ent detailide juurde liikudes leiavad paljud, et avatud juurdepääsu võimaldamine kõigile artiklitele ja teadusandmetele kätkeb endas mitmeid praktilisi probleeme ja teaduskorralduslikke küsitavusi. Analüüs tutvustab Eesti huvirühmade hoiakuid ja tegevusi avatud teaduse valdkonnas, eesmärgiga aidata kaasa tõhusa ja osapoolte parimaid huviseid arvestava avatud teaduse süsteemi kujundamisel Eestis.

Analüüs tugineb huvirühmade esindajate seas läbi viidud küsitlustel ja intervjuudel (täpsemalt meetodika peatükis 2.1). Uuring keskendub kahele keskele avatud teaduse eesmärgile: avatud juurdepääs teadusartiklitele ja avatud juurdepääs teadusandmetele. Uuringu üks

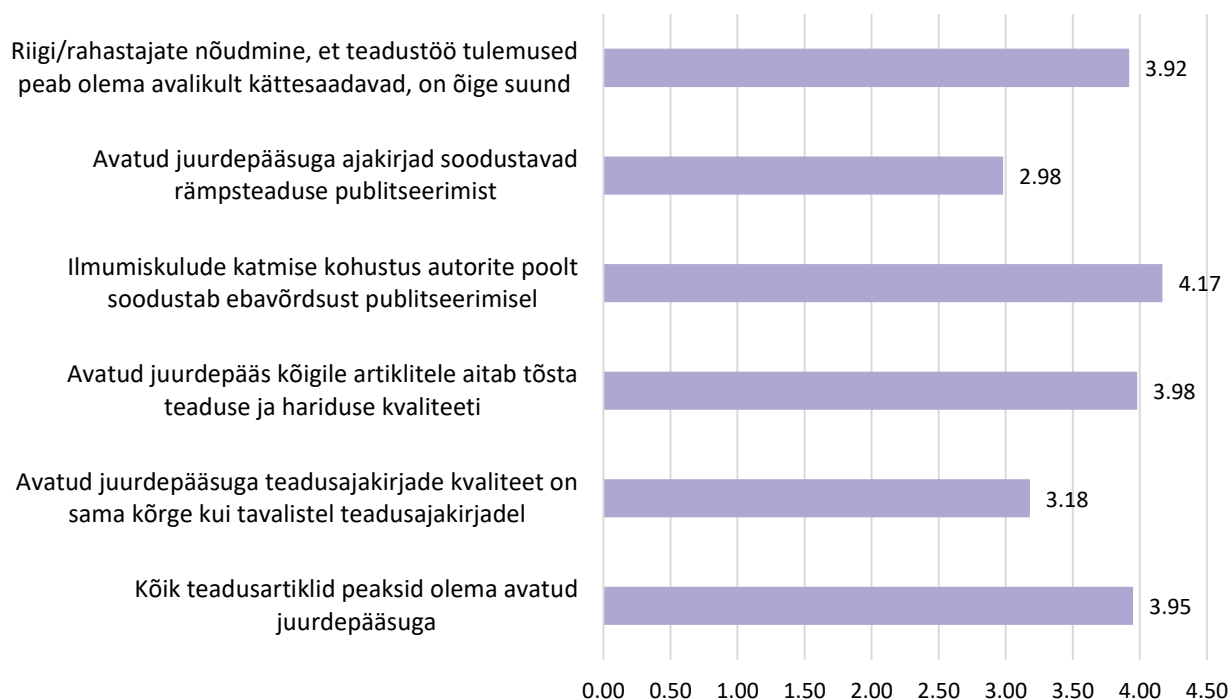
peamisi lähtekohti oli 2016. aasta juulis Eesti Teadusagentuuri avatud teaduse ekspertkomisjoni koostatud põhimõtted ja soovitused riikliku poliitika kujundamiseks<sup>28</sup>.

### 2.2.1 Avatud juurdepääs teadusartiklitele

Eesti Teadusagentuuri avatud teaduse ekspertkomisjoni sõnastatud Eesti strateegiline eesmärk aastaks 2020 on: „Avaliku sektori poolse rahastamise toel avaldatud teadusartiklid on kõigile vabalt kättesaadavad hiljemalt 1 a peale nende esmailumist, vähemalt pooled neist viivitamatult ja lõplikul kujul ning tagatud on nende pikaajaline säilitamine.“

Avatud juurdepääsu võimaldamise all saab mõista nii artikli avaldamist *open-access* formaadis, milles ajakirjad (kohe või viivitusega) annavad tekstile vaba juurdepääsu, kuid ka nn isearhiveerimist ehk artikli või käsikirja üles laadimist vaba juurdepääsuga repositooriumitesse (mille tuntumad näited on ArXiv, PubMed Central ja ResearchGate).

Teadlaste seas läbi viidud küsitlusest nähtub, et avatud juurdepääsus artiklitele nähakse teaduse jaoks kasu (Joonis 10): väite „avatud juurdepääs kõigile artiklitele aitab tõsta teaduse ja hariduse kvaliteeti“ keskmine toetus viiepallisüsteemis on 3,98. Toetavad ollakse idee suhtes, et kõik artiklid võiksid olla vabalt kättesaadavad (keskmine hinne 3,95). Kõrge hinde (3,92) saab ka väide: „riigi/rahastajate nõudmine, et teadustöö tulemused peab olema avalikult kättesaadavad, on õige suund“.



**Joonis 10.** Küsimustikule vastanud hindasid avatud juurdepääsu teadusartiklitele kõikide väidete osas oluliseks.

Ise eelistab *open-access* ajakirjades publitseerimist umbes veerand vastanutest (on enamuse või kõik enda viimase viie aasta jooksul ilmunud artiklitest avaldanud neis väljaannetes).

<sup>28</sup> [http://www.etag.ee/wp-content/uploads/2016/07/Avatud\\_Teadus\\_Eestis\\_1.0.pdf](http://www.etag.ee/wp-content/uploads/2016/07/Avatud_Teadus_Eestis_1.0.pdf)

Viiendiku jaoks on avatud juurdepääsuga ajakirjad on ka oluline infoallikas, nad saavad sealt enamiku oma teadusinfost. Samas toodi välja, et kuna ülikoolide kaudu on juurdepääs ajakirjadele üsna ulatuslik, ei märgatagi tihti, kas kasutatav ajakiri on avatud juurdepääsuga või mitte.

*“Praegu lihtsalt ei ole vahet, kas [artikkel] on open access või mitte, me saame kätte kõik.”* (loodusteadused)

Peamiste takistustena oma artiklite avatud juurdepääsuga avaldamisel toodi välja kaks kesket kaalutlust: avatud juurdepääsuga ajakirjade kvaliteet ning neis avaldamise rahastusmudel.

Ajakirja mainet tuuakse küsitluse vastanute seas esile peamise põhjusena, miks nad avaldavad avatud juurdepääsuga ajakirjades (viiendik vastuist avaldab seal põhimõtte pärast). Kvaliteedi hindamisel ilmnevad erisused teadusvaldkondade vahel. Mõne vastanu sõnul ongi nende valdkonna parim ajakiri *open-access*. Ja kvaliteedi puudumist („need on meie erialal siiski rämp”) toodi küsitluses avatud vastusena välja sageli neis mitteavaldamise põhjusena.

Paludes hinnangut väitele „avatud juurdepääsuga teadusajakirjade kvaliteet on sama kõrge kui tavalistel teadusajakirjadel“ jagunesid hinnangud viiepallisüsteemi hinnete vahel pea võrdselt (keskmine 3,18). Vaadates tulemusi aga valdkondade kaupa, näeme erisusi. Kõige kõrgem keskmine hinne (4,02) on sellele väitele humanitaarteaduste ja kunsti valdkonna vastajatel, järgnevad arsti- ja terviseteadused (3,44). Kõige madalama hinne (2,89) annavad tehnika ja tehnoloogiavaldkonna esindajad. Hinnang peegeldub ka publitseerimisaktiivsuses: sama järjestus on näha, kui vaadata, kui suur osa valdkonna vastajatest ütleb, et avaldab kõik või enamuse oma artiklitest *open-access* ajakirjades.

Nii intervjuudes kui ka küsitluse avatud vastustes tõid teadlased välja, et nende jaoks on oluline avaldatud artiklite nähtavus. Valdkonnast sõltuvalt võib see nii soosida kui ka takistada avatud juurdepääsuga ajakirjades avaldamist. Nii põhjendati avatud juurdepääsuga ajakirjades avaldamist muu hulgas vastustega “suurem tsiteeritavus ja levik”, “peamine kaalutlus on maksimeerida lugejaskond”, “loetavus - lihtsale lugejale on need kättesaadavamad. Ei taha, et teadus jääks ainult ülikooli”. Sarnast mõtet väljendati ka ühes intervjuus:

*“Teadlaskond leiab [artiklid] nagunii üles, kui ta on huvitatud. Aga võib-olla Eesti kontekstis on tähtsam isegi populariseerimine või levitamine /.../ võib-olla oleks olulisem, et tulemusi just nimelt kogukonnale kättesaadavaks teha.”* (arsti- ja terviseteadused)

Teisalt toodi ka välja, kui valdkonnas on välja kujunenud teatud ajakirjad, mida kõik teadlased jälgivad, siis püütakse avaldada kindlasti nendes, hoolimata, kas ajakirjad on avatud juurdepääsuga või mitte. Mõjufaktorina väljendatav ajakirja tajutud kvaliteet on endiselt oluline tegur avaldamiskoha valikul.

*“Mul on küll olemas kogemus, et mingid asjad saab [avatud juurdepääsuga ajakirjades] kiiremini avaldatud, samal ajal me vaatame ikkagi väga rangelt, et impact oleks vähemalt 3, alla selle me ei avalda.”* (arsti- ja terviseteadused)

Lisaks kohtab umbusku avatud juurdepääsuga ajakirjade suhtes, mille tõttu nähakse seal avaldatud artikleid alaväärsemana (need on “ostetud”) või siis tekitab vastumeelsust ajakirjade ärimudel.

*“Mul on eelarvamus, et [see on] nagu makstud asi, et raha eest saad kiiremini ja paremini sisse ennast osta.”* (loodusteadused)

*“[Avatud juurdepääsuga avaldamise eesmärk pole] mitte see, et meie saame oma teadust paremini avaldada, küsimus oli, et kirjastused hakkasid välja surema ja nüüd mõeldi välja kaval nipp.”* (sotsiaalteadused)

Fookusgruppides osalejaid tõid ka näiteid, kus selline umbusk on välja löönud teadlaste hindamisel ja tõdevad, et sellised hoiakud takistavad *open-access* ajakirjade eelistamist.

*“Olen ise näinud, kuidas nõukogu otsustab doktoritöö kaitsmisele lubamist selle järgi, et tal on seal üks PLoSOne [artikkel] ka, siis peaks ikka ühe veel kirjutama äkki. Et ei ei, selle ta on ostnud, see on väga kahtlane.”* (loodusteadused)

*“Ja kui sa taotled oma töökohta tagasi, siis vaadatakse, mis ta on pannud [publikatsioonide] nimekirja. [Avatud juurdepääsuga ajakirjades avaldamisel] peaks olema siis hoopis teine renomee ju sellisel juhul.”* (loodusteadused)

*“Jah, täpselt, see asi hakkab hindamise juures pihta. Aga see on meil endi kätes.”* (loodusteadused)

Sellised hoiakud võivad ühelt poolt peegeldada teatud vastajarühmade jäikasad mõttemustreid, mis ei luba *open-access* ajakirjasid omaks võtta, kuid näitab kindlasti ka avatud juurdepääsuga ajakirjade ebaühtlast arengut ja teadusvaldkondade erinevat motivatsiooni järgida avatud teaduse põhimõtteid.

Keskne aruteluteema *open-access* avaldamise puhul oli siiski raha: kuna enamasti seostatakse *open-access* avaldamist mudeliga, kus avaldamistasu maksab artikli autor, on teadlaste tavapärane esimene reaktsioon küsida, kes need kulud hüvitab. Juhul kui avatud juurdepääs teadusartiklitele saab Eestis rahastuse üheks nõudeks, võib eeldada, et sellest küsimusest tuleb üks kesksemaid debatikohti. Teadlased väljendasid intervjuudes selget ootust, et kui rahastaja nõuab, et tulemused oleksid avaldatud avatud juurdepääsuga ajakirjades, siis eraldatakse selle jaoks ka täiendavaid ressursse.

Otsustajate sõnul on aga teadusrahastus tõusnud igal aastal, grandilepingutes on praegugi olemas publitseerimiseks ette nähtud eelarverida ning spetsiaalselt avatud teadusele mõeldud lisarahastust ei tule, ülikoolidel ja teadlastel on võimalus olemasoleva eelarve raames teha avatud teaduse alaseid valikuid.

Teadlased omakorda viitasid praeguse teadusrahastuse väiksusele, mis muudab valikud keerukaks.

*“Kui [open-access avaldamise raha] tuleb teadussüsteemist, siis on teadlasel valida, kas saada vähem palka ja [avaldada] open-access või rohkem palka ja mitte open-access. Sel juhul mina valin palga.”* (loodusteadused)



*“Kui eelarvet koostada, siis on küsimus selles, kuhu raha läheb. Et kas maksta 1000 eurot artikli open access’i eest või läheb see 1000 eurot millegi muu peale, on see siis konverentsi külastus, andmete kogumine, midagi veel.”* (humanitaarteadused ja kunstid)

*“Eesti enda rahastus on ju suhteliselt tillukene, sealt siis see 2000-3000 eurot panna open-access’i peale, see paneb mõtlema ja mitte iga artikli puhul ei ole saanud seda luksust lubada. Kui uurimisgrupp publitseerib aastas ühe artikli, siis on see okei, aga kui sul on 4-6 artiklit, siis see on päris suur summa, mis ühest uurimisrühma grandist selle peale kulub.”* (arsti- ja terviseteadused)

Selle tulemusel võib suurened ka ebavõrdsus publitseerimisel väiksemate ressursidega riikide kahjuks. Küsitluses sai kõige kõrgema toetuse (4,17) väide „ilmumiskulude katmise kohustus autorite poolt soodustab ebavõrdsust publitseerimisel“.

Fookusgruupiintervjuudest tuli siiski välja ka see, et kui teadlase on piisav motivatsioon *open-access* avaldamiseks, leitakse selle jaoks ka ressursse, kuigi teinekord tuleb selle jaoks teha kompromisse: “nutad ja maksad”, nagu üks teadlane väljendus. Üks biomeditsiini valdkonna esindaja tõi esile, et pärast mainekat granti, mille nõue oli *open-access* avaldamine, sai see nende rühmale harjumuseks ning nüüd leitakse enamasti võimalus avaldamiskulude eest tasumiseks. Biomeditsiini toodi ka näitena valdkonnast, kus *open-access* avaldamine on saanud valdkondlikuks normiks, tänu sellele, et seda hakkasid nõudma rahastajad ja juhtivad teadusrühmad hakkasid seda edendama. Valdkonna liidrite mõju olulisust mainisid ka teised:

*“Kui hakkab muutuma trendiks, et juhtivad konkreetse valdkonna grupid hakkavad [open-access] avaldama, siis tulevad teised kaasa ja siis see muutubki normiks.”* (loodusteadused)

Ükski fookusgruppides osalenud teadlaste rühm ei pidanud realistlikuks mõtet loobuda senistest teadusajakirjade tellimustest ja suunata need vahendid avatud juurdepääsuga ajakirjades avaldamiskulude katteks. See tooks nende hinnangul siiski kaasa senisest oluliselt halvema juurdepääsu teadusinfole, kuna oluline osa ajakirjadest jäävad ilmselt praeguse tellimiskeskse ärimudeli juurde.

Isearhiveerimine, mida saab käsitleda alternatiivina *open-access* ajakirjades avaldamisele, on küsitletute seas levinum kui avatud juurdepääsuga ajakirjades avaldamine. Küsitluses märkisid rohkem kui pooled, et on enamuse või kõik oma artiklitest isearhiveerinud (või on seda teinud nende kaasautorid), enim humanitaarteadustes. Ligikaudu kolmandik täpsustas, et arhiveerib artiklid ResearchGate’is, ArXivis, oma kodulehel või muus sarnases avalikus paigas. Paljud märkisid isearhiveerimise paigana ETISE, kuid nende puhul ei ole alati mõistetav, kas nad tõlgendavad isearhiveerimisena<sup>29</sup> täisteksti kättesaadavaks tegemist või selle lihtsalt üles laadimist, avalikuks tegemata.

Isearhiveeritud artiklid teadusinfo allikana on vähetähtsad enam kui poolte vastanute jaoks. Toodi välja, et artiklite isearhiveeritud versioone hakatakse otsima siis, kui “ajakiri pole

---

<sup>29</sup> Kuigi ‘isearhiveerimise’ mõiste definitsioon oli küsitluses antud, näitasid vastused, et termin võib olla eksitav, sest ei seostu inimestele avalikuks tegemisega, vaid ainult säilitamisega. Nii olid mõnedki vastused, et oma artikleid arhiveeritakse arvuti kõvakettal. Soovitame kaaluda uue mõiste välja töötamist, mis selgitaks tegevuse olemust intuiitivsemalt.

ülikooli poolt tellitute hulgas ja pole ka avatud juurdepääsuga”, kuigi neis olukordades kasutatakse ka otse autori poole pöördumist.

Mõnes valdkonnas (nt osakestefüüsika, majandus, osa bioloogiast) on valdavaks saanud eelretsenseerimata artiklite avaldamine ja levitamine (nt *pre-print* või *working paper*’i formaadis), ajendatuna peamiselt pikkadest publitseerimisjärjekordadest ajakirjades ja/või vajadusest oma tulemusi kiirelt jagada. Ka need täidavad vähemalt teadlasringkondade endi jaoks vaba juurdepääsu eesmärgi (eeldades kasutajalt samal ajal suuremat asjatundlikkust, et hinnata töö kvaliteeti).

*“Sellise working paper’i versioonis teeme praegu [open-access’i] nii kui nii. Ja meile ei maksa keegi selle eest midagi ja meie ka ei maksa. See on ka oluline, et see open-access on tasuta.”* (sotsiaalteadlane)

Samas on läbi rääkimata, kuidas suhestuvad sellised artiklid rahastajapoolsete publitseerimise ja avatud juurdepääsu tingimustega.

Avatud juurdepääsu nõuete järgmine isearhiveerimiskeskondade abil eeldab täpsemat kursisolekut artikli litsentsitingimustega (võimaliku embargo järgimine, millist versiooni on lubatud üles laadida jms) ning selle tõttu ka teadlaste koolitamist selles valdkonnas. Isearhiveerimise puhul tunnistavad paljud, et seaduslikkuse mõttes on sageli tegu halli alaga, kuna teinekord rikub nende üles riputamine artiklite litsentsitingimusi. Neist mööda minemiseks saab levitamiseks kasutada artikli varasemaid või toimetuse küljendusega versioone, kuid mõned ei pea seda kõige paremaks lahenduseks:

*“Mulle meeldib see kõige puhtam versioon. Ma tahaks, et kõik loeksid seda, mitte minu vahepealset versiooni. Samal ajal ma arvan, et parem jagada ilmselt seda, kui üldse mitte midagi. Kõige parem on see, kui saad jagada lõppversiooni.”* (arsti- ja tervise teadused)

Samuti poleks seetõttu nende keskkondade abil võimalik täita ETAgil ekspertrühma sõnastatud strateegilist eesmärki, et vähemalt pooled artiklid oleks koheselt kättesaadavad *lõplikul kujul*. Seetõttu peljatakse, et:

*“kirjastused ei lepi sellega. Kui [isearhiveerimine] toimub mingisuguses proportsioonis ja kirjastuste tulud ei vähene, siis on see ok. Aga sellele ei saa üles ehitada absoluutset lähenemist.”* (otsustaja)

Sama diskussioon kerkis seoses artiklite avaldamisega ETISes. Paljud intervjuueeritavad nägid, et ETISel on potentsiaal ja tehnilised võimalused olla Eesti keskne artiklite repositoorium ning pidasid seda kõige loogilisemaks arenguks. Kuna projektidega seoses laadivad teadlased nagunii sinna enda artiklid üles, oleks hõlbus need seal ka kohe kättesaadavaks teha (kui need juba ei ole kusagil avaldatud *open-access*). Küll polnud vastanute seas üksmeelt, kes peaks selle juures jälgima litsentsitingimusi, kas ETISe administraator või teadlane ise.

Lisaks ei suuda ETISesse üles laetud artiklid praegu oluliselt kaasa aidata artiklile suuremale mõjukusele, sest need ei ole indekseeritud levinumates teadusartiklite otsingusüsteemides ehk nende rahvusvaheline nähtavus on väga madal (see on lubaduse kohaselt küll muutumas).

**Kokkuvõttes:** kuigi üldised hoiakud on artiklite vaba juurdepääsu soosivad, juhib teadlaste avaldamisstrateegiaid peamiselt põhimõtte, et nende artikkel oleks hõlpsalt leitav ja (oma valdkonna inimeste poolt) võimalikult laialt loetav. Sõltuvalt valdkonnast võib see tähendada avaldamist valdkonna tippajakirjades (sõltumata selle litsentsitingimustest), teadlikke valikuid *open-access* ajakirjade kasuks, laialdast isearhiveerimist, alternatiivsete publitseerimisvormide eelistamist (nt monograafiad, konverentsiannaalid) või hoopis tulemuste kiiret levitamist *working paper*'ite või *pre-print*-artiklite vormis. Mõni valdkond on artiklite vaba juurdepääsu põhimõtted laialdaselt omaks võtnud, teine näeb potentsiaalses nõudes levitada artikleid eelistatud *open-access*-formaadis vastuolu nende valdkonnas tunnustatud parimate publitseerimistavade ja seeläbi ka ohtu enda töö nähtavusele.

Artiklitele vaba juurdepääsu kujundavaid tingimusi mõjutavad oluliselt globaalsed aspektid, mis ei ole kohalikul Eesti tasandil mõjutatavad – eelkõige valdkondlikud trendid eelistatud publitseerimistavade osas (mis võivad olla nii *open-access*'i soosivad kui taunivad), kirjastajate käitumine (enamasti piiranguid seadev) kui ka suurte teadusrahastajate seatud tingimused (enamasti *open-access*'i eelistavad).

Kohaliku tasandi tegevused (koolitused, teavitustegevus, rahastajapoolsed nõuded) saavad toetada globaalset suundumust avaliku juurdepääsu suunas ning kujundada siinsete teadlaste ja teiste huvirühmade hoiakuid ja käitumist.

Peamine kommunikatsioonialane väljakutse on leida ühine arusaam iga osapoole kohustustest olukorras, kus teadlastel on selge ootus, et avatud juurdepääsuga ajakirjades avaldamiskulu ei tule nende senise rahastuse arvelt. Selle juures on oluline tutvustada ka alternatiive *open-access* ajakirjades avaldamisele, mis lähtuvad iga valdkonna võimalustest ja vajadustest.

### 2.2.2 Teaduse avaandmed

Eesti Teadusagentuuri avatud teaduse ekspertkomisjoni sõnastatud Eesti strateegiline eesmärk aastaks 2020 on: „Avaliku sektori poolse rahastamise toel kogutud teadusandmed on kõigile vabalt kättesaadavad ja taaskasutatavad, kui seda ei piira seadusest tulenevad kitsendused. Teadusandmed hoiustatakse kindlatele standarditele vastavates ja avatud repositooriumites ning need tehakse kättesaadavaks esimesel võimalusel.“

Varem on avatud teadusandmete jagamise alaseid hoiakuid ja praktikaid teadlaste seas uurinud Muuli (2014) ja Neerut (2016). Nende kohaselt on teadlaste seas endiselt peamise säilitamisviisina levinud hoiustamine jagamist mittesoodustavatel vahenditel (tööarvutis, välisel andmekandjal). Andmete säilitamise kuludest ei olda teadlikud. Andmeid jagama ja teiste andmeid kasutama ollakse valmis, kuid ollakse vähe kursis näiteks metaandmetega varustamise teemaga.

Ka selle küsitluse tulemused näitasid, et rohkem kui pooled teadlastest hoiustavad andmeid peamiselt enda (töö)arvutis või infokandjal. 30 protsendi andmed on asutuse infosüsteemis ja vaid 5,5 protsendil valdkondlikus repositooriumis<sup>30</sup>. Seejuures ollakse valitud hoiustamisviisiga enamasti rahul, probleeme tuuakse enamasti välja isiklikul seadmed

---

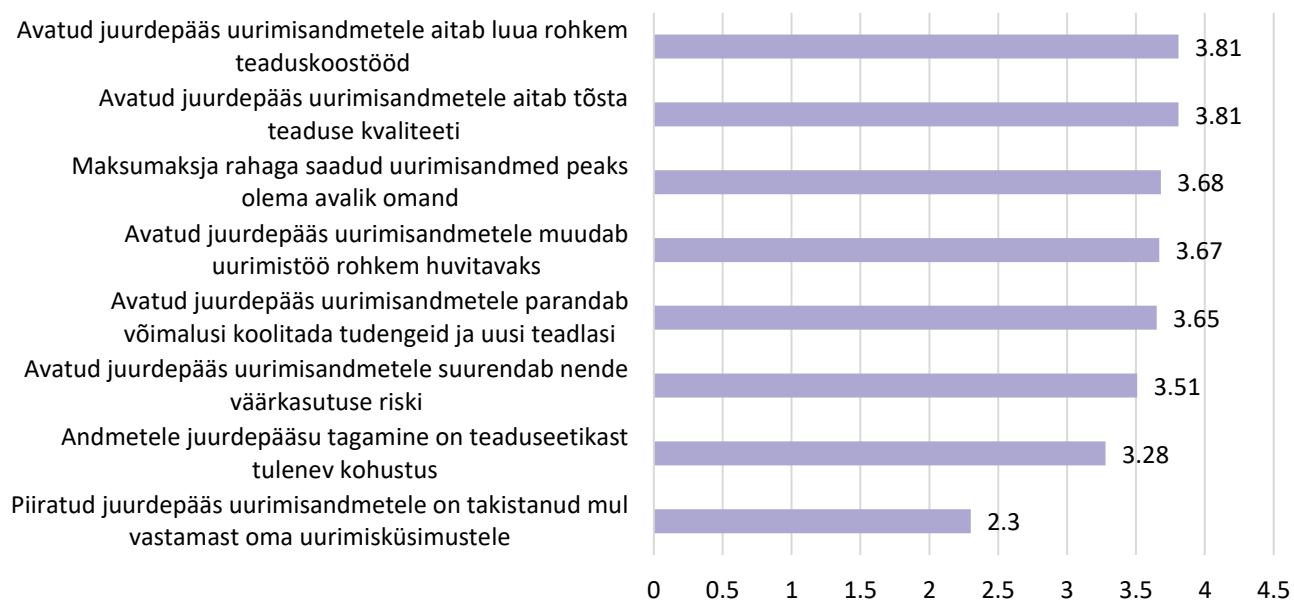
<sup>30</sup> Meie küsitlusuuringu puudus oli võimalus valida andmeid puudutavate küsimuste puhul ainult üks vastusevariant (kuigi iga küsimuse juures oli võimalus vastust laiendada). Seega ei saa küsitlus hästi peegeldada juhte, kui teadlasel on eri andmetega eri kogemusi ja praktikaid ning tulemused on pigem üldisi trende indikeerivad.

hoiustamise kohta, mainides näiteks andmete kadumise ohtu, vähest mahtu, korrastamatust ja keerukust andmete jagamisel. Mitmel korral oodati suuremat institutsiooni panust sobilike hoiustamissüsteemide pakkumisel.

Küsitlusele vastanute seas oli neid, kelle sõnul on kõik nende teadusandmed avalikult kättesaadavad, 5,6protsenti, ja nende osakaal, kelle andmed pole üldse kättesaadavad, 12,8 protsenti. Ligikaudu pooled kinnitasid, et kui teised teadlased küsivad, on nad valmis oma andmeid jagama. Kättesaadavaks mitte tegemise peamiste põhjustena toodi välja isikuandmete kaitse ning kartus, et jagamine ohustab teadustöö eesmärkide saavutamist.

Kaks kolmandikku vastanud teadlastest usub, et teised teadlased või teadusväline huvirühm võiks olla huvitatud nende kogutud andmete kasutamisest. Oma andmeid ei ole viimase kolme aasta jooksul jaganud 17 protsenti vastanutest ega teiste omi kasutanud 20 protsenti. Laiemat jagamist kui vaid oma asutuse ja riigi kolleegidele mainib peamisena 45 protsenti vastanutest, rahvusvaheliste kolleegide ja teadusväliste rühmade kogutud andmeid on omakorda kasutanud umbes pooled. (Võib eeldada, et olulise osa rahvusvahelise jagamise ja kasutuse juhtudest moodustavad koostööprojektid, mille käigus vahetatakse vastastikku andmeid, kuid täpsemat hinnangut, mil määral ei ole andmete jagamine ja saamine ajendatud koostööst, on selle küsimustiku põhjal võimatu anda).

Küsitletud näevad teaduste avaandmetel positiivseid mõjusid. Nende kohta käivaid väited (joonis 11) saavad üsna ühtlaselt soosivaid hinnanguid, kuigi pisut madalamaid, kui samalaadsed väited avatud juurdepääsuga artiklite kohta. Siiski ollakse üsna nõus (viiepalliskaalal 3,51) ka väitega, et “avatud juurdepääs uurimisandmetele suurendab nende väärkasutuse riski”.



**Joonis 11.** Küsitletud näevad teaduste avaandmetel positiivseid mõjusid

Küsitlusest nähtuvast soosivast hoiakust hoolimata näitab intervjuudes ja küsitluses kogutu, et andmete ka tegelikult kättesaadavaks tegemise ees on mitmeid barjääre.

Valik küsitluses vaba vastusena antud takistustest, mis võtavad kokku sagedamini kordunud teemad:

*“Ülikoolis puudub infra andmete avalikustamiseks. Teisest küljest, minu toorandmed ei huvita mitte kui kedagi. Andmete teistele arusaadavaks tegemiseks tuleks need varustada ulatusliku dokumentatsiooniga, mille valmimiseks kulub meeletu aeg. Isegi siis on huvi nende vastu küsitav.”*

*“Kõige suurem oht on andmete väärinterpreteeringus. Kui andmeid jagada vabalt, peaks kõik nende andmete kasutamised ka ikkagi kooskõlastatud olema, et andmekoguja saaks nõ oma “*

*“Osa originaalandmete põhjal pole ise veel teadustööd avaldanud ja kõigile vabalt kättesaadavaks tehes võiks nende kogumiseks nähtud vaevast tulenevat kasu lõigata keegi teine, aga mina jääksin ilma (publish or perish).”*

*“Kui tegu on sensitiivsete patsiendi andmetega, siis peab veenduma, et nende andmete abil ei oleks võimalik isikuid tuvastada. Teiseks ei olda piisavalt hästi kursis avalike andmebaasidega kuhu oma andmeid on võimalik üles laadida ja teiste kolleegidega jagada. Lisaks on andmete jagamine üsna tüütu ja pikk protsess millega võibolla ei taheta tegeleda.”*

*“Takistusi ei ole”*

Intervjuudes laiendati samu takistusi, seega võib peamised barjäärid kokku võtta järgnevalt:

- motivatsiooni puudumine;
- otsustamine, millised andmed ja millal tuleks kättesaadavaks muuta;
- praktilised oskused ja ressursid andmete kättesaadavaks muutmiseks.

Motivatsiooni aspekt toob taas esile selged valdkondlikud erinevused. Eriti mõnes loodusteaduste valdkondades on kujunenud tavaks, et ajakirjad nõuavad koos artikliga algandmeid, mis publitseeritakse koos artikliga. See moodustab osa teaduse kvaliteedikontrollist. On ka erinevad suurprojektid (nt CERNi eksperimendid, Euroopa Sotsiaaluuring), mille põhimõtete seas on samuti andmete avalikult kättesaadavaks tegemine. Seega on rühm teadlasi, kelle jaoks on andmete kättesaadavaks tegemine tavapraktika, mis lähtub valdkonna üldistest tavadest ja nõuetest.

Enamjaolt ei ole andmete kättesaadavaks tegemine aga juurdunud tava ning selle vajadust ja kasu tunnetatakse vähem, võrreldes artiklite vabalt kättesaadavaks tegemisega. See rühm leiab enamasti, et andmete kättesaadavaks tegemise ressurssikulu ja sellest saadav kasu ei ole nende jaoks tasakaalus.

*“Andmed on ressurss, aga ressurss ei teki iseenesest. Ressurss tekib mingi töö tulemusena. Ja töö tähendab aega. Ja aeg on omakorda jälle piiratud ressurss ja siis tekib valik. Ühe tee peal on see, kus sul on seal vastas null: raha selle eest ei maksta, ainult aeg kulub. Siis sa pead võtma teised teed, kus on mingi number peal. Ja kulutad oma aega selle tee peale. Ja siis tekib valik, sul vaatavad vastu viis haru. Siit vaatab null.” (sotsiaalteadused)*

Tuuakse ka välja, et nende kogukonnas ei ole seni andmetele juurdepääsuga olnud olulisi probleeme: kui on tekkinud vajadus, on otse autorilt küsitud ja saadud, seega pole tarvis

rajada selle jaoks eraldi infrastruktuuri. Või siis ollakse veendunud, et alg- või toorandmetel ei ole teiste jaoks väärtust - see, mis andmetes on olulist, antakse edasi nende põhjal valminud publikatsiooniga.

Kui väärtus on andmetes olemas, nähakse jälle riski, et teised uurimisrühmad jõuavad neid kasutades publitseeriised andmete kogujast ette, mistõttu jääb andmete koguja teaduse konkurentsituatsioonis kehvemasse seisu.

Teadlase motivatsiooni kahandab ka tema vastutustunne andmete eest, eelkõige seoses sellega, et andmete tõlgendamine nõuab palju taustateadmisi. Kui andmed pole piisavalt kirjeldatud (või polegi seda võimalik teha), siis peljatakse, et nende põhjal "avaldatakse mind kaasamata ja valesti/asjatundmatult tõlgendades".

Teadlaste endi poolt nähti andmete kättesaadavaks tegemisel teatud eeliseid teadusele (võrdlusuuringud, vigade kergem paljastamine, korduva andmekogumise vältimine, negatiivsete tulemuste jagamine), kuid, mõned eestkõnelejad välja arvata, ei tajutud neid piisavalt tugevatena, et õigustada olulist ressurside lisapanustamist. Rühmana, kes andmekogude kättesaadavaks tegemise vastu kõige enam huvi võiks tunda, toodi välja oma elutööd korrastavad emeriteerunud teadlased.

Peeti küll oluliseks, et kujuneks välja tunnustussüsteem teiste andmete kasutamise eest, kuid teadlikkus sellest oli madal. Mitmel korral arvati, et andmetele viitamine saab võimalik olla ainult koos artikliga viitamisega:

*"Sellega peab ikka mingi artikkel ka seotud olema, niisama andmete peale ei saa."*  
(loodusteadused)

Kui rakendub andmete viitamise süsteem (*a la* Data Citation Index), on teadlaste motiveerimiseks siiski vajalik, et tema hindamises võetaks neid kaalukana arvesse. Hetkel ei ole Eesti otsustajad valmis pakkuma selget või kaalukat lahendust:

*"[Andmekogumist väärtustada] saab rahastaja ikkagi läbi grandi hindamiskriteeriumite, aga olulisem on ikkagi see, kuidas ta hakkab rahvusvaheliselt välja kujunema. Kui hakkab seal välja kujunema mingisugune mudel, on seda Eestis üsna lihtne juurutada."* (otsustaja)

*"Andmete tootmine näib minu silmis paigutuvat teaduse tegemise väärtusahelas mitte päris samale pulgale selle tööga, mis on seotud analüüsi ja mõtestamise ja järelduste tegemisega."* (ülikooli esindaja)

Teine peamine takistuste kogum, mille üle teadlased intervjuudes pikemalt arutlesid, on seotud kättesaadavaks tehtavate andmete olemuse ja ajastamisega. Üksmeel oli selles, et isikuandmed peavad olema kaitstud. Selle kõrval tekkisid osalenud teadlastel küsimused, mida üldse lugeda algandmeteks (kas ka nt loodud koodijupid, arhiiviandmed) ning millal peaks andmed tegema kättesaadavaks. Üldiselt pooldati põhimõtet, et andmed saab teha kättesaadavaks, kui nende põhjal valminud artikkel või artiklid on publitseeritud (ehk joonisel 1 kujutatud andmepüramiidi ülemine aste).

Püramiidi alumistel astmetel olevate andmetüüpide osas oldi ettevaatlikud ning pigem toodi välja nende kättesaadavaks tegemise riske, alates tarbetust ressursikulust kuni konkurentsi jaoksitõlgendusteni.

Toodi välja ka andmete kasutatavuse (ja sellest tulenevalt kasulikkuse) erinevat iseloomu tulenevalt valdkonnast. Kui näiteks loodus- ja täppisteadustes on meetodikate ühtlustamise korral võimalik võrrelda eri riikides kogutud andmeid, siis sotsiaal- ja humanitaarteadustes on uuringud enamasti palju kontekstispetsiifilisemad ja nende väärtus tuleb teadlaste hinnangul välja ainult autori tõlgenduses.

Eraldi teemana tõstatati küsimus pikaajaliste ja seireuuringute kohta: mis hetkel peaksid need saama avaandmeteks? Peeti oluliseks, et teadlasele jääks andmete sisulistest omadustest tulenev vabadus otsustada, millised andmeid ja millal kättesaadavaks teha. Kohustuse korral teha andmed kättesaadavaks, soovivad nad selgeid ning võimalikke ohte arvestavaid reegleid, millised andmed ja millistel juhtudel peaks vabalt kättesaadavaks tegema.

Selliseid küsimusi peaks käsitlema projekti alguses koostatav andmehaldusplaani. Sammuna teaduse avaandmete suunal soovib näiteks Horisont 2020 programm koos granditaotlusega saada projektiga loodavate andmete haldamise plaani (s.o. tegevused andmete säilitamiseks ja võimalusel avaldamiseks). Ka Eesti Teadusagentuur liigub sinna suunas, paludes praegustes granditaotlustes kirjeldada vastavaid tegevusi ja kavandades täiemahulise andmehaldusplaani koostamise nõude sisseviimist lähitulevikus. Fookusgrupiintervjuudest selgus, et vähestel on veel sellise plaani koostamisega kogemusi ning teadmisi selle olemusest, ka abi pole saadud.

*“Õiget vastust ei olnud võimalik küsimusest välja lugeda, kuidas peaks vastama. Et kas nad eeldavad seda, et ma panen [andmed avalikult üles], kuidas ja millal.”*  
(loodusteadused).

*“Nägin esmakordselt seda kirja ühes järel doktoriteemas. Mina, lihtne inimene, võtsin kohe telefoni, helistasin ETAGi, sain väga umbmääraseid vastuseid sealt. Siis ma vaatasin ja mõtlesin, kirjutasin, et andmed on juba ETAGi serveris, jagame koostööpartneritele ja nii edasi. Kas see on õige või ei ole, selle kohta ei öeldud midagi. Ja konkreetne inimene, kes oleks võinud vastata, ei osanud ka vastata.”*  
(loodusteadused)

Siinkohal nägid raamatukogud endil rolli teadlaste toetamisel. Oluline on aga, kas teadlane saab esitatud plaanile ja selle täitmisele tagasisidet või on tegu dokumendiga, mille puhul jälgitakse vaid selle esitamist. Viimasel puhul, viitasid intervjuud, suhtuvad teadlased ka nende koostamise formaalselt.

Oluliseks barjääriks hinnati ka töömahtu, mis kaasneb toorandmete avaldamiskõlblikuks tegemisega (nt anonümiseerimine, metaandmetega varustamine, sobivasse formaati viimine). Sedagi tajuvad teadlased lisakohustusena, mille jaoks napib nii aega kui teinekord ka oskusi ning mis ei too neile selgeid kasusid. Jällegi paisati välja arvamus, et kui grantide osaks saab nõue teha algandmed kättesaadavaks, peaks sellega kaasnema ka seda võimaldavad ressursid, samas kui teised huvirühmad, nt rahastajad leiavad, et andmete korrastamine ja kättesaadavaks tegemine peaks saama teadlaste töö loomulikuks osaks.

Võimalustest teadlasi selles töös toetada mainiti kõige rohkem toetust andmehaldusplaani koostamisel, andmeteadlaste koolitamist ning teadusrühmade või ülikoolide juurde spetsiaalsete andmespetsialistide tööle võtmist. Esimese osas on valmisolek kõige suurem, nii raamatukogude kui ülikoolide TAOde poolt, samas näevad nad enda rolli siiski vaid toetajana, mitte sisuliselt andmetega tegutsejana. Ülikoolid on valmis hakkama andmeteadlasi õpetama,

kuid hetkel ei ilmuta valmisolekut võtta tööle projekte toetavaid ainult selle funktsiooniga spetsialiste.

Võrreldes kirjastamisega (ning teatud määral ka isearhiveerimisega) on avaandmete juures praegu vähem ärihuvisid (mis on näiteks kirjastamise juures üks teadlastes tõrkeid tekitavaid aspekte). Kui need ärihuvid tekivad (andmekaeve jms näol) – mida mõned eestkõnelejad ka eesmärgina näevad – võib see samuti kaasa tuua potentsiaalse vastuseisu oma andmete jagamisele.

Kokkuvõttes on näha, et teaduse avaandmete vallas on Eesti teadlaste seas ühelt poolt kujunenud välja selged eestkõnelejad ja fokuseeritud projektid, teisalt valitseb palju teadmatust ja kartusi. Teadlased on motiveeritud algandmeid kättesaadavaks tegema juhtudel, kus on näha valdkondlik loogika ja potentsiaalne kasu ning kasutajad. Kuid selline selge positiivne mõju on praegu vaid teatud tüüpi (valdavalt kvantitatiivsete loodusteaduslike) andmete korral, ülejäänud valdkondades on palju enam kahtlusi, mille hajutamine või millele valdkonnaspetsiifilise lahenduse leidmine on peamine kommunikatsioonialane väljakutse.

Üks võimalik viis muuta hoiakuid ja soodustada andmete kättesaadavaks tegemise praktikaid, on alustada andmete säilitamise alasest teavitustööst. Nagu küsitlus näitas, on palju andmeid, mida praegu hoitakse personaalarvutites ja andmekandjatel, koos sellega kaasnevate riskidega. Teadlased väljendasid huvi institutsionaalsete või riiklike andmehoidlate vastu ning saab järeldada, et andmete asjakohase hoiustamise vastu oli neil hetkel suurem huvi kui avalikult kättesaadavaks tegemise vastu. Et aga esimene on vajalik samm teise võimaldamiseks, tasub panustada teadlikkuse tõstmisse ja säilitusvõimaluste tutvustamisse. Selline tegevus vajab ka väärtustamist, mida ühe raamatukogu esindaja sõnul hetkel pole:

*“Ei ole mõistlikku seisukohta, kuidas hinnatakse seda, kui keegi oma andmekogud üldse kuskil alles hoiab ja siis omaette, kui nähtavad ja viidatud need on. Täna, kus meil tehniline võimekus on olemas, puuduvad poliitilised otsused.”*

Tihti toodi andmete kättesaadavaks tegemise takistusena, et pole selle jaoks sobilikku kohta. Raamatukogude hinnangul on aga olemasolevate repositooriumite hulk ja omadused juba praegu piisavad, katmaks eri teadusvaldkondade vajadusi. Nii saab andmete hoiustamiseks (ja avaldamiseks) valdkonnast sõltuvalt valida Eestis loodud valdkondlikke repositooriume, rahvusvahelisi valdkondlikke repositooriume ja raamatukogupõhiseid repositooriume.

Samas ei taga repositooriumi olemasolu vastavat käitumist, rõhutab üks otsustaja:

*“Meil on mitmes valdkonnas olemas infrastruktuur, mis lubaks kõik avalikuks teha, aga kus see kultuur pole juuri alla võtnud hästi. /.../ Võib üsna turvaliselt väita, et kultuuri ja poliitikate pool on kriitilisem, raud tuleb järgi.”*

Eraldi aruteluteema oli võimalik keskne Eesti repositoorium siinsete andmete jaoks, mis ühtlasi võimaldaks tagada andmete kättesaadavaks tegemise nõude järgimise kõigile, kes seda vajavad. Mõte tekitas vastakaid arvamusi, vajadust nähti (ka selle jaoks, et koondada kokku näiteks riigi tellitud uuringute andmed ja tulemused), kuid eraldi institutsiooni loomist peeti ebaotstarbekaks. Võimalike variantidena pakuti välja Eesti Teadusarvutuste Infrastruktuuri (ETAIS) või ETISE arendamise sellise võimekuse pakkujaks või ka taristu välja arendamine ülikoolide juures. Samuti leiti, et peaks olema koht, nt ETIS, kuhu on koondatud lingid andmekogudele või pakutud muu moodus neid hõlpsaks leidmiseks.



Ühe olulise teemana kerkis üles teaduse avaandmete suhestumine riigi avaandmetega. Teadlaste hinnangul annaks parem juurdepääs riigi andmetele meie teadusele palju enam kui teadusandmete avamine, seega oleks riigi eeskuju oluline ka teadlaste motiveerimisel. Senised kogemused riigi kogutud andmete kasutamisel on teadlastel olnud pigem frustrerivad:

*“Milleks siis Statistikaamet kogub andmeid, miks ta viib riigi rahaga läbi uuringuid, kui pärast selgub, et praktiliselt sul võimalik ainult... noh, teed oma töö ära ja tuleb, mis tuleb, aga andmeid sa tegelikult üldse ise kuidagimoodi kätte ei saa.”*  
(Sotsiaalteadused)

Ka riigi ja selle allasutuste tellitud uuringute andmete puhul puudub praegu ühtne poliitika nende kättesaadavuse osas. Küsitletud teadusnõunike kohaselt ei ole üheski ministeeriumis praegu nende tellitud uuringute algandmed vabalt kättesaadavad. Ka nende andmete kättesaadavaks tegemise võimalikkuse puhul kohtab sarnaseid kaalutlusi (isikuandmete kaitse, ressursikulu) nagu teadlastel:

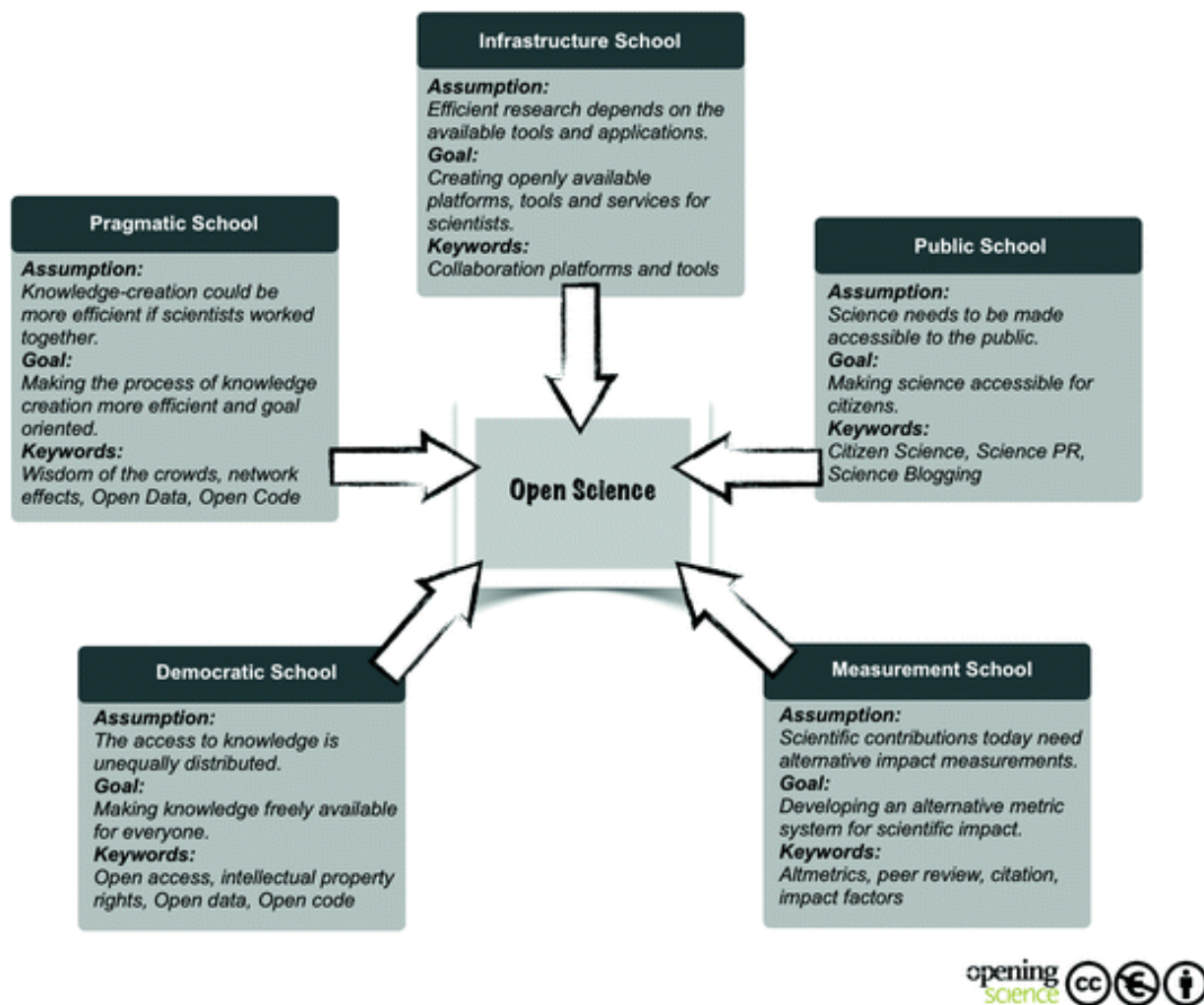
*“Kõikide riigi tellitud uuringute/andmete kättesaadavaks tegemine on mõistlik suund, kuid tuleb arvestada seda, et kõiki ei saa teha avalikult ja keskselt kättesaadavaks (mõistliku argumendi, ratsionaalsuse kaalutlused). Lisaks kui need asuksid ühes keskses süsteemis, peaks see olema lihtne ning mitte-dubleeriv. Nt on olemas ETIS, kuid selle kasutamine ei ole kasutajasõbralik. Kui teha süsteem keskseks ning kõigile kättesaadavaks ei pruugi tähendada see seda, et oluliselt rohkem kasutaksid erinevad ühiskonnagrupid olemasolevaid andmeid, kuid halduskoormus ametnikele/teadlastele jne võib saadava kasu asemel kasvada. Nt on olemas Eurostat, kuid kui paljud ühiskonna-grupid seda kasutavad? Lisaks julgeoleku või ärikaalutlused.”*  
(teadusnõunik)

Üldiselt siiski ollakse seda meelt, et “riigi ja teaduse avaandmed võiksid asuda ühtses süsteemis” (kõik vastanud teadusnõunikud “jah” ja “pigem jah”).

Lisaks haakub teaduse avaandmete teema ka kogude digiteerimisega, mida eelkõige mainisid raamatukogud.

### 2.2.3 Avatud teaduse võimalikud mõjud

Põhimõttena toetavad avatust pea kõik vastanud ja huvirühmad. Nagu näitavad Fecher ja Friesike (2013), on avatud teadust võimalik mõista ja käsitleda aga eri nurkade alt, eri mõttesüsteemides (Joonis 12). Neis lähenemistes on eri fookused, arusaamad, mida ja mille jaoks tuleb avatud teaduses teha, ja sellest tulenevalt ka eri arusaam eesmärkidest ja kasudest.



## Joonis 12. Avatud teaduse viis mõttesuunda (Fecher ja Friesike 2013)

Ka selle uuringu raames tehtud intervjuud näitasid, et arusaam sellest, miks Eestis on avatud teadus vajalik, ei ole ühtne.

Põhjendustena tõi otsustajad välja vajaduse käia kaasas rahvusvaheliste arengutega (“mõtestada paratamatust”) ja olla isegi avatud teaduse liikumise esirinnas. Otsustajate ja valdkonna eestkõnelejate sõnul aitab avatud teadus tõsta ka Eesti teaduse mõju (“Millisel määral, on raske öelda, aga vaevalt et ta seda ei tee, sest kui asjad on rohkem kättesaadavad, siis vähem neid ikka ei võeta kasutusele”) ning soodustada innovatsiooni.

Intervjuud näitasid, et need põhjendused ei kipu alati teadlasi kõnetama. Avatud teadust edendava käitumise vähesust saab ühelt pool selgitada napi teadlikkusega, teisalt aga motivatsiooniküsimustega. Motivatsiooni oluline komponent on tegevuse eesmärgipärasus, kasude ja tulemuste tajumine.

Analüüsi tulemused näitavad, et teadlased mõtleavad avatud teaduse kasudest peamiselt teadusekeskselt ja pragmaatiliselt (toob parema koostöö, tagab tulemuste laiemat nähtavust, parandab kvaliteeti jne). Seda peamiselt seoses avatud juurdepääsuga artiklitele, vähem aga teaduse avaandmete puhul, sest teaduskogukonnas on andmete omavaheline jagamine (omavahelise suhtluse vahendusel) niigi toimunud. Olukordades, kus avatud juurdepääsule üleminek ei too silmnähtavat kasu neis aspektides, ilmnebki tõrksus võimalike sellekohaste kohustuste vastu.

Avatud teaduses nähakse pea kõigi poolt küll potentsiaalset kasu ka teistele, teadusvälistele huvirühmadele (nt otsustajad, ametnikud, ajakirjanikud, õpilased, vabaühendused):

*“Üldiselt on hea kui andmed, mis ei ole väga tundlikud on kättesaadavad. Järjest enam on näha, et erinevad kodanikualgatused tahavad olla protseessides tihedamalt kaasatud ning mida rohkem neil on andmeid, seda paremini saavad nad seda teha.”* (teadusnõunik)

*“Kogu ühiskonnale. Kindlasti ajakirjanikele ja poliitikutele, aga ka nõunikele, kooliõpetajatele, meditsiinihuvilistele, protestimeelsele aktivistile jne. Ei saa pahaks panna, et inimesed levitavad rumalusest valearusaamisi, kui teaduskirjandusele ligipääs puudub.”* (loodusteadused/tehnika ja tehnoloogia)

Samas ollakse tihti ka skeptilised nende potentsiaalsete kasutajate suutlikkuses artikleid või andmeid mõista ja töödelda.

*“Teoreetiliselt näen väga suurt kasu oleks ühiskonnal sellest, kui kas just andmed, aga nendest andmetest tekkiv teadmine oleks ühiskonnas selgelt levinud, kättesaadav. Aga seni ei ole ühiskonnas seda teist poolt, kes oleks valmis vastu võtma neid andmestikke. See peaks tähendama seda, et vähemalt doktorikraadiga inimeste osakaal ühiskonnas peaks oluliselt oluliselt suurenema.”* (loodusteadused)

*“Ma näen, et avalikult kättesaadavad ajakirjad on ammendamatuks materjalideks kõikvõimalikele diletantidele üksikute teadustulemuste väärtõlgendamise arendamisel. Alternatiivsete raviskeemide pooldajad, vaksineerimisvastased, imevähivijad, toitumisgurud - kogu see seltskond on hakanud oma väidete tõestuseks massiliselt kasutama kontekstist välja rebitud teadustulemusi. Teadusartiklite lugemine on oskus, mida tuleb õppida, oskamatule ei anna need mingit mõistlikku tulemit. Teadusartikli mõistmiseks tuleb minimaalselt omada vastava valdkonna magistriharidust.”* (loodusteadused)

Selliste hoiakute murdmisel on oluline, et teadusvälised huvirühmad ise oma tegevusega näitaksid, et suudavad avatud teadust kasutada ja selle abil tulemusi saavutada.

Ka huvirühmad ise tunnevad põhimõttelist huvi avatud juurdepääsu vastu. Nii teadusnõunikud kui vabaühendused olid küsitluses päri, et vastavalt riigivalitsemine ja vabaühenduste tegevus peaks olema teaduspõhisem. Mõlemal rühmal oli enim mainitud teadusinfo allikas, mida soovivad oma töös rohkem kasutada teadusajakirjades avaldatud artiklid, ning peamine takistus nende kasutamisele puudulik ligipääs. Ühe vabaühenduse esindaja sõnul aitaks parem juurdepääs teadusartiklitele neil:

*“teavitustööd teha n-ö faktipõhisemalt, olla tõsiseltvõetavam suhtluses sihtgrupi ja ka partneritega. Aitaks ka valida, millistele teemadele keskenduda (kus on piisavalt Eesti kontekstis relevantset tausta) ja millistele mitte.”*

Teadusallikate mõistmise keerukust hindas takistuseks kaks vabaühenduste esindajat ning kolm teadusnõunikku seitsmest. Teadusandmetele juurdepääsu võimalusi enda töö parandamisel hinnati madalamalt:

*“Aitaks kaasa küll, ent pole nii suure osatähtsusega kui ligipääs analüüsidele, artiklitele ja aruannetele (algandmete läbianalüüsimiseks napiks tõenäoliselt aega aga kardetavasti ka oskusi).” (teadusnõunik)*

Üks vabaühenduste esindaja võttis teema kokku:

*“Teaduspõhise lähenemise vajalikkust on vaja selgitada nii sihtrühmadele kui ka tavainimesele; vabaühendusi tuleks õpetada, kuidas nad uuringutest endale vajalikku ja kasulikku infot, järeltuleid leiavad. Mitmed vabaühendused ei oska seda infot enda huvides kasutada.”*

Need tulemused näitavad, et vähemalt huvirühmade teadlikumas osas (sest küsitlusele vastasid suurema tõenäosusega just sellised isikud) tajutakse avatud teaduse kasusid, on valmisolek teadusinfot rohkem kasutada ning tuuakse juurdepääsu puudumine esile ühe peamise takistusena.

Seega on eri sidusrühmades olemas arusaam avatud teaduse kasulikkusest, kuid fookuste erinevuse tõttu (infrastruktuuri arendamine, reeglite sõnastamine, kasu tajumine laiemalt teaduse jaoks või kitsamalt teadlase tööle, võimalik kasu ühiskonnale) ei leia osapooled tihti vastastikku ühiseid vaatenurki, millel on mõju nende tegutsemise aktiivsusele.

#### 2.2.4 Kokkuvõtte huvirühmade kaupa

Raamatukogud on motiveeritud teemaga tegelema ning näevad enda rolli eelkõige harija ja teavitajana, mõnel juhul ka taristu (repositooriumite) pakkujana. Samas kurvad raskusi ressurside (eelkõige inimeste ja aja) puuduse üle.

Teadlased tunnevad hirmu, et avatud teaduse nõuetest saab järjekordne lisakoorem, mis pidurdab nende võimalusi pühenduda teadustegevusele – nii raha kui ka muude ressurside osas. Valdkonni on erinevad praktikad ja suhtumine avatud teadusesse, on tekkinud ja laienemas teema eestvedajate rühm.

Ülikoolide jaoks ei ole teema esmase tähtsusega. Nad on valmis toetama teadlasi sedavõrd, et vastata rahastajate nõuetele ja teadlaste endi olulisematele soovidele, kuid ei taha võtta avatud teaduse suunas suunaja rolli (näiteks teadlase töö hindamiskriteeriumite kaudu).

Eesti Teadusagentuur on teema üks algatajaid Eestis ning liigub vastavate nõuete grandilepingutesse sisse viimise poole. On valmis olema teadlastele ja institutsioonidele toeks, näiteks koolituste ja juhendmaterjalidega.

Haridus- ja teadusministeerium – on poliitika sõnastaja ning juhib infrastruktuuri projektide arendamist. Ei kavanda eraldi sihtrahastust avatud teaduse tegevuste jaoks.

Teised huvirühmad – huvi teema vastu on üksikutel, kuid seni laialt teadvustamata.

#### Soovitused:

- Sõnastada avatud teaduse positiivsed mõjud eri ühiskonnatasanditel, koos näidetega, ning kasutada neid teadlikkuse tõstmisel ning avatud teaduse suunal liikumise vajalikkuse põhjendamisel;

- Kasutada teadlaste hoiakute ja käitumise suunamiseks sama valdkonna esindajaid, kes on avatud teaduse eeskõnelejad;
- Arvestada avatud teaduse alaste nõuete kehtestamisel valdkondlike erinevustega;
- Toetada avatud teaduse alaste oskuste suurenemist kõigis seotud institutsioonides (nt andmespetsialiste koolitades, lisades teema doktorantide õppekavva);
- Pidada osapoolte vahel pidevat sisulist arutelu avatud teaduse eesmärkide ja vahendite osas (näiteks andes teadlastele sisulist tagasisidet andmehaldusplaanide kohta);
- Pakkuda ülikoolidele motivatsiooni institutsionaalsete avatud teaduse praktikate paremaks juurutamiseks.

### 3 Avatud teadusandmete säilitamise maksumust mõjutavad tegurid

*Lähteülesanne: vaadeldakse varjatud kulusid Open Access publitseerimisele ja avatud teadusandmete rakendamise kulusid erinevate tegurite lõikes.*

Digiajastu mõjutab jätkuvalt nii teaduse tegemist kui selle tulemuste saamist ja levitamist ning tõstatab üha uusi küsimusi teadlaste, teadusasutuste ja teaduse rahastajate ette. Teadusandmed on praegu üks tõsisemaid proovikivisid – mitte üksnes teaduspublikatsioonide analüüside tõenduseks olevad andmekogumid, vaid kogu andmestu, mis kogutakse ja luuakse uurimistöökäigus. Need andmed on väärtuslikuks allikmaterjaliks mitte üksnes teistele teadlastele, vaid kõigile huvitatud isikutele ja organisatsioonidele nii avalikus- kui erasektoris. Teadusandmete avalikustamine aitab kaasa uurimuste korratavusele ja tõstab teadustulemuste usaldusväärsust. Kuid teadusandmete taaskasutamine eeldab terve taristu ökosüsteemi olemasolu – riistvara, tarkvara ja teenused, kuid ka poliitikad, korrad ning organisatsioonid, teadusrühmad ja teadlased. Üha rohkem andmeid, mis teaduse tegemise käigus luuakse, tõstavad nõudeid eelkõige just tehnilisele taristule, kuhu andmed talletatakse ja mida kasutatakse nende töötlemiseks. Teadusandmed ei ole reeglina ühetaolised, vaid pigem heterogeensed, mistõttu peab nende haldamise taristu toime tulema keerukate nõuetega kõigi teadusandmete elukäigu etappide puhul: loomine, kogumine, valimine, analüüsimine, töötlemine, salvestamine ja säilitamine.

Teaduse ja ka teadusandmete maastik on kirev – erinevad on nii asutused, rahastusvõimalused, koostöövormid kui erialad ja valdkonnad. Teadusandmete arhiveerimise ja taaskasutamise korraldamine pidevalt muutuv ja arenevas kogukonnas on olnud keeruline. Pikka aega domineerinud entusiastlike teadlaste ja kogukondade initsiatiivide kõrvale on üha rohkem tekkinud asutuse-põhiseid lahendusi, riiklikke- ja üle-Euroopalise teenuseid. Tulemuseks on suhteliselt ebaühtlase tasemega ja vähe-konsolideeritud teenuste „turg“, kus esineb kattuvusi ja tühimikke. Üks selle põhjuseid on eri tasandi asutuste eesmärkide, poliitikate ja huvide erinevused, mis on suunanud teadusandmetega tegelema eelkõige asutuse vajadustest lähtudes. Üle-Euroopaliselt, ja ka ülemaailmselt, on asunud teadusandmetega seotud poliitikaid uuendama ja neile vastavaid ühiskasutuse taristuid looma, mis lähtuvad vähem eri teadusharude erisustest ning keskenduvad teadusandmete haldamise vajadustele ja taaskasutamisest tekkivale kasule teaduses ja ühiskonnas laiemalt.

*Good data management is not a goal in itself, but rather is the key conduit leading to knowledge discovery and innovation.*

*(The FAIR Guiding Principles for scientific data management and stewardship)*

Järgnev ülevaade tugineb rahvusvahelise teadusandmete avaldamise ja arhiveerimise praktika kirjeldusele kirjanduses. Esmalt kirjeldatakse vajaliku tugi-taristu maksumust mõjutavaid tegureid, mis on omavahel seotud ja koosmõjus. Analüüsi teises osas hinnatakse erinevate andmete avaldamise ja arhiveerimise teenuste omadusi. Seejärel vaetakse teadusandmete avaldamise ja arhiveerimise tasuvuse arvutamise võimalusi ning lõpetuseks tuuakse ettepanekud (teadusandmete arhiveerimise osas) järgmise viie aasta tegevuste teekaardina.

### 3.1 Teadusandmete säilitamise taristu maksumust mõjutavad olulisemad komponendid

Teadusandmete haldamist ja arhiveerimist suunavad poliitikad ja korrad keskenduvad enim järgmistele aspektidele, kuna need otsused määravad kõige enam koguteenuse maksumust.

#### 3.1.1 Milliseid andmeid avaldada ja arhiveerida?

Teadusprojekti käigus kogutakse ja luuakse kirev ja väga kõikuva tasemega kirjeldatud andmete hulk, mida iseloomustavad versioonide ja vormingute paljusus, kattuvus, ebavajaliku info rikkus, vähesed kontekstiandmed, kompressioon ja mõnikord krüpteerimine. Teadusprojekti elukäigu jooksul satuvad erinevad andmehulgad erinevatesse andmehoidlatesse, näiteks teadusartikli aluseks olnud andmed võivad jõuda kirjastaja digihoidlasse juba enne teadusprojekti lõppu. Rahvusvahelise praktika põhjal saab välja tuua vähemalt neli stsenaariumi:

1. teadusartikliga seotud andmed, mis reeglina talletatakse artikli avaldanud kirjastuse digihoidlas (nt. PLOS One<sup>31</sup>);
2. distsipliini või valdkonna edenemise seisukohalt olulised uued andmed, mis avaldatakse pigem oma erialases publikatsioonide ja/või andmete portaalis (nt. GenBank<sup>32</sup>);
3. rahvusvahelise või erasektori rahastusega uurimisprojektide tulemandmed, millel on kokku lepitud (nt. projekti rahastaja poolt ette nähtud) avaldamise ja arhiveerimise reeglid, nt. kasutatakse riigiüleseid taristuid (nt. Eudat<sup>33</sup>) või kindlaid erialaseid andmehoidlaid;
4. teadusprojekti andmehaldusplaanis näidatud projekti tulemandmed, mis antakse üle kas institutsionaalsesse digihoidlasse (nt. ülikooli raamatukogus) või riiklikku teadusandmete arhiivi (nt. HEPDATA<sup>34</sup>);
5. muud andmed, mis jäävad teadusprojekti veebisaidile, institutsioonide arvutivõrkudesse ja osalenud teadlaste isiklikesse arhiividesse ning mida süstemaatiliselt ei arhiveerita.

Andmete vaates on tulemuseks paljuharuline mudel, kus erinevad versioonid projekti tulemandmetest on talletatud eri asukohtades (Joonis 13).

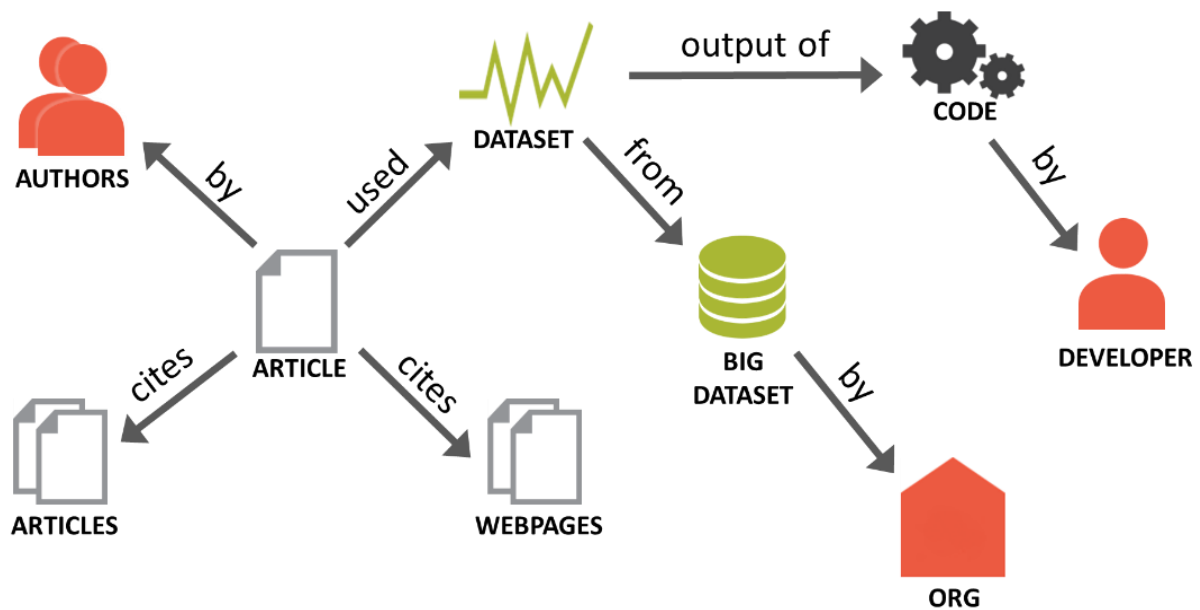
---

<sup>31</sup> <http://journals.plos.org/plosone/s/data-availability>

<sup>32</sup> <https://www.ncbi.nlm.nih.gov/genbank/>

<sup>33</sup> <https://www.eudat.eu/services-support/>

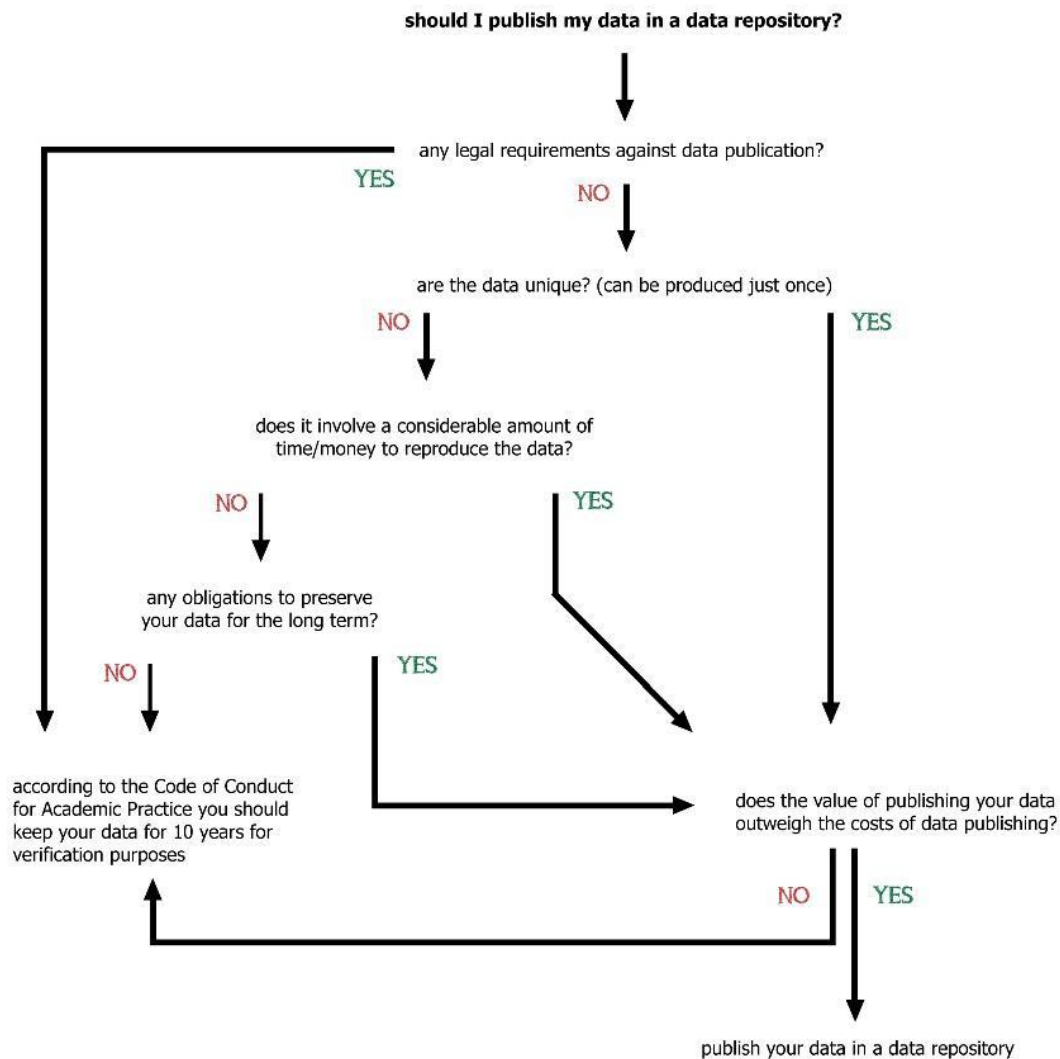
<sup>34</sup> <https://hepdata.net/>



**Joonis 13.** Teadusprojekti andmete publitseerimise protsess (RMap Project 2016; [http://rmap-project.info/?page\\_id=98](http://rmap-project.info/?page_id=98))

Teadusprojekti lõppedes või hiljem tagant järele on raske otsustada, milliseid andmeid või nende osa oleks vaja projekti tulemusena avaldada ja arhiveerida. Kõigi projekti tulemite arhiveerimine oleks enamasti ebaotstarbekas, kulukas ja ajamahukas, eelkõige täiendava kirjelduse loomise vajaduse tõttu. Valiku tegemisel tuleb vastata küsimusele: „Mida pidada kõige väärtuslikumaks osaks projekti andmetest?“. Ühtseid, eri distsipliine katvaid arhiveerimise reegleid on raske kehtestada ja seepärast püüdleval paljud riigid selle suunas, et projekti andmehaldusplaan seaks paika andmed või valiku kriteeriumid, milliseid andmeid projekti tulemusena avaldatakse ja arhiveeritakse. Ülikoolid ja andmearhiivid on asunud temaatilisi juhiseid oma teadlastele andma, mis abistavad arhiveerimist väärivate andmete valikul (nt. Whyte & Wilson, 2010; Tjalsma, Rombouts, 2011; <http://www.dt-toolkit.ac.uk/researchscenarios/archiving.cfm>; Joonis 14).





**Joonis 14.** Andmete valik arhiveerimiseks. (Wagenigen University & Research. <https://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/When-and-why-to-publish-your-dataset.htm>)

Eesti sihtrühmade küsitlus näitas (vt. ptk.2), et:

- teadlaste huvi on avaandmete jagamiseks teiste teadlastega on suur, kuid see toimub pigem küsijatele kättesaadavuse põhimõttel;
- teadlaste huvi teiste teadlaste (eelkõige rahvusvaheliste kolleegide) andmeid kasutada on suur;
- andmete kasutajate huvi väljaspoolt teadust on suhteliselt madal ning erineb oluliselt sihtrühmade lõikes.

**Soovitus:** Riiklikul tasandil on otstarbekas reguleerida nõue, et andmehaldusplaanis näidatakse ära kriteeriumid, mille alusel projekt plaanib oma andmed arhiveerida. Andmehaldusplaani koostajate jaoks tuleb lisada vastav juhendmaterjal.

### 3.1.2 Andmete kvaliteet avaldamiseks ja arhiveerimiseks

Enne andmete avaldamist ja üleandmist säilitajale tuleb veenduda, et andmed on nõutud kvaliteedis. Näiteks, kas andmete vorming ja võimalik vajalik tarkvara on kasutatavad; kas andmed on piisavalt kirjeldatud, et nende taaskasutamine on võimalik ka konteksti otseselt mitte tundval inimesel; kas kõik vajalikud õigused andmetega seoses on välja selgitatud ja võimaldavad andmete üleandmist ning taaskasutamist?

Andmete puhastamine, kirjeldamine ja konverteerimine kujule, mis tagaks nende pikaajalise taaskasutatavuse võib osutada kulukaks ja aeganõudvaks. UK Data Service on välja töötanud juhendmaterjali andmete arhiveerimiseks ettevalmistamisega seotud tegevustega ja ka vastavate kulude hindamiseks (UK Data Service, 2015).

Sellest, kui põhjalikult on andmekogu kirjeldatud, sõltub otseselt selle kasutatavus tulevikus (Open Research ..., 2017, lk. 38):

*Re-usability depends critically on recording comprehensive information about why, when, where and how data was created, gathered or measured, and how it was subsequently manipulated or modified: data documentation on provenance. Documenting provenance in detail is essential to enable users to understand the data (names, labels and descriptions for variables, code and so on) and to verify that the research was conducted as described.*

Rõhutatakse kirjelduse, õiguste ja litsentsi-andmete põhjaliku dokumenteerimise vajadust, et toetada teadusandmete taaskasutamist mitte üksnes inimeste, vaid ka automaatsete süsteemide poolt, on paljud organisatsioonid asunud toetama n.n. FAIR andmete põhimõtteid (Wilkinson et al. 2016):

- Findability – leitavad
- Accessibility – kättesaadavad
- Interoperability – koosvõimelised
- Re-usability – taaskasutatavad.

FAIR andmete põhimõtetele on oma toetuse andnud nii Euroopa Komisjon<sup>35</sup> kui G7 riikide Open Science Working group (G7 Science..., 2017) ning mitmed koostöökogud on asunud välja töötama vahendeid FAIR andmete loomiseks (nt. DTL FAIR Data tools<sup>36</sup>, WDS/RDA Assessment of Data Fitness for Use Working Group<sup>37</sup>, FAIR Metrics Group<sup>38</sup>). Taoliste töövahendite abil on võimalik hinnata, kas projektis loodud andmed vastavad FAIR-andmete kvaliteedi ootustele. Täpsemad andmed ja hinnangud maksumuse kohta, mis andmete viimine FAIR põhimõtetele vastavale kvaliteedi tasemele endaga kaasa toovad, seni puuduvad, kuid vastavaid uuringuid on juba käivitatud.<sup>39</sup>

**Soovitus:** Seada FAIR põhimõtteid järgivate andmete arhiveerimine soovitavaks kvaliteedi tasemeks riiklikku teadusrahastust saavatele uurimisprojektidele.

<sup>35</sup> <http://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupDetail&groupID=3464>

<sup>36</sup> <https://www.dtls.nl/fair-data/find-fair-data-tools/>

<sup>37</sup> <https://www.rd-alliance.org/groups/assessment-data-fitness-use>

<sup>38</sup> <https://www.dtls.nl/fair-data/fair-metrics-group/>

<sup>39</sup> vt. <https://www.dtls.nl/categorie/fair-data/>

### 3.1.3 Kus andmeid avaldada ja arhiveerida?

Teadusprojektide rahastamisel pannakse tihti paika nõuded projekti tulemandmete arhiveerimiseks avaldamiseks. Lisaks ette nähtud kanalitele on projektil võimalik oma tulemandmeid avaldada veel läbi:

- üldiste rahvusvaheliste avatud teaduse platvormide (nt. Figshare, Zenodo, Dryad, jt.)
- valdkondlike rahvusvaheliste andmearhiivide (nt. CLARIN, GESIS, PubMed Central, jt.)
- riiklike andmearhiivide (nt. ANDS), mis võivad olla jaotatud valdkondlikeks (nt. NATARC ja ESTA Eestis; Ühendkuningriigis UK Data Archive sotsiaalteaduste jaoks, ADS arheoloogia andmete jaoks, National Chemical Database Service, jpt).
- teadusasutuse digihoidla (DataDOI, Edinburgh DataShare, Bodleian ORA-Data, jpt.)
- projekti veebileht ja informaalne andmete jagamine (nt. kolleegilt-kolleegile).

Lisaks avaldamisele ja hoiustamisele pakuvad eri kanalid erinevaid lisateenuseid, s.h. püsiidentifikaatori lisamine, kindlatele sihtrühmade teavitamine, embargo-periood avaldamisele, pikaajaline säilitamine, aruandlus projekti rahastajale, bibliomeetria ja altmeetria, sidumine teadusartiklitega, nõustamine, andmete kasutajate päringutele vastamine, jt. tugiteenuseid. Suutlikkus neid pakkuda eeldab aga andmete kvaliteedinõuete täitmist ja kulusid, mis on seotud andmete üleandmise ja arhiveerimisega. Üldine reegel on, et mida usaldusväärsemat säilitusteenust ja konkreetse valdkonna sihtrühmale suunatud avaldamist andmearhiiv pakub, seda kulukamaks kujuneb andmete arhiveerimine neis. Mida universaalsem on andmete avaldamise platvorm (s.t. ei teeninda kindlat valdkonda või sihtrühma), seda madalamad on sinna andmete arhiveerimise nõuded, aga tulemusena ka andmete nähtavus ja mõju. Kui andmearhiivid on üha enam panustamas usaldusväärse tööendamisele läbi auditite, millega kontrollitakse nende digitaalse säilitamise tegevuste jätkusuutlikkust (nt. DRAMBORA; CoreTrust; DIN 31644; ISO 16363), siis teadusandmete teenusplatvormid piirduvad enamasti salvestusruumi pakkumisega, kuid ei võta endale vastutust andmete pikaajalise kasutatavana säilitamise eest (nt. ETAIS<sup>40</sup>). Ajutiste projektide veebilehtede kaudu jagatavad andmed on kõige suurema kadumise riskiastmega – kuigi nende talletamiseks on võimalik kasutada veebiarhiveerimise teenuseid, panustatakse sellesse vähe.

Läbi viidud kvalitatiivse uuringu andmetel on Eesti teadlaste eelistus arhiveerimisel arvuti või väline kõvaketas (üle poole vastanutest), millele populaarsuselt järgneb institutsionaalne andmehoidla (29% vastanutest). Sealjuures on küsimustikule vastanud märkinud, et andmete taastamine kõvaketta purunemisel on keeruline (53%) ning mitmed vastajad on märkinud, et täiesti võimatu (12%) (vt. ptk. 2).

### 3.1.4 Kuidas tagada teadlastele tugi ja nõustamine?

Teadlaste ja projektide nõustamine ning toetamine andmete loomisel ja haldamisel on üha mahukam teenus, mida pakuvad nii ülikoolide raamatukogud, andmearhiivid (nt. UK Data Service) kui üleriiklikud kompetentsikeskused (nt. Digital Curation Centre Ühendkuningriigis, Research Data Netherlands Hollandis, jt.). Mõned keskused pakuvad ka andmehalduse

---

<sup>40</sup> <http://etais.ee/>

teenuseid käimasolevatele projektidele (nt. SurfDrive Hollandis, DataShare Šotimaal, jt.). Ühtset väljakujunenud mudelit, kuidas avatud teaduse põhimõtete rakendamise ja andmete haldamise juhendamist riigis korraldada, ei ole välja kujunenud. Enamuses Euroopa riikides on riik panustanud mõne teenuse väljaarendamisse, kuid selle kõrval on alt-üles initsiatiivina aktiivsed ka üksikud institutsioonid ja üha enam institutsioonide koostöös hallatavad teenuseplatvormid (nt. JISCi poolt toetatud RDMRose, 4TU.Centre for Research Data Hollandis, jt.). Teadusandmete haldamise toetamine projekti teostamise faasis on olulise tähtsusega teenus, millele tehtud kulutus aitab kokku hoida hilisemas faasi arhiveerimisele ja säilitamisele tehtavaid kulutusi.

Kvalitatiivne uuring näitas, et Eesti teadlaste peamiseks mureks on andmete avalikuks tegemisega seotud suur töömaht ning tundlikud andmed, samuti vajavad teadlased tuge metaandmete loomisel ning andmehaldusplaani koostamisel. (vt. ptk. 2)

#### *3.1.4.1 Eesti teadusraamatukogude tegevus avatud teadusandmete kontekstis*

Ideaalstruktuurina peaks teadusraamatukogude tugisüsteem hõlmama koolituste korraldamist, juhendmaterjalide koostamist ning konsultatsioone. Esmajoones peaks teadlastele olema tagatud tugisüsteem, mis abistaks neid teadusandmete haldamisel ja selle planeerimisel.

Koolituste sihtrühmaks on peamiselt doktorandid ja projektidega alustavad teadlased. Juhendid on abiks kõigile, kes vajavad infot teadusandmete haldamise ja andmehaldusplaanide (*Data Management Plan*) koostamise kohta. Konsultatsioonid hõlmaksid nii andmehaldusplaanide koostamist kui andmete hoiustamise võimalusi. Raamatukogudes peaks olema võimalik arendada ka kompetentsi, et osaleda teadusprojektides andmete haldamises.

Teine oluline pool on tehnilise võimekuse tagamine – repositooriumid. Ülikoolide raamatukogud peaksid osutama abi teadusandmete arhiveerimisel. Ühest küljest luues ja arendades repositooriume, kus teadlased saaksid hoiustada/arhiveerida uurimistöö andmeid ning teisalt nõustama sobivate, juba olemasolevate, repositooriumite leidmisel. Lisaks sellele saavad raamatukogud aidata metaandmete loomisel ja standardite kasutamisel.

Pidades silmas Eesti mastaape ja olemasolevaid ressursse oleks ilmselt otstarbekas luua repositoorium koostöös kõigi ülikoolidega DataCite Eesti konsortsiumi egiidi all.

Teadusandmetega tegelemiseks on loodud konsortsium DataCite Eesti, kuhu kuuluvad Tartu Ülikool, Tallinna Tehnikaülikool, Eesti Maaülikool ning Tallinna Ülikool. Ülikoolide esindajatena tegutsevad konsortsiumis liikmete raamatukogude töötajad. Antud hetkel on vaid Tartu Ülikooli raamatukogu võimeline pakkuma ulatuslikku tugiteenust ja tehnilist abi. TÜ Raamatukogus on teadusandmete spetsialiste kokku neli. Samuti on loodud võimalus arhiveerida teadusandmeid repositooriumis DSpace. Ülejäänud raamatukogudel puudub hetkel ressurss tugiteenuste ulatuslikumaks väljaarendamiseks. On olemas spetsialist, kes tegeleb teadusandmete küsimustega, kuid seda lisaks muudele tööülesannetele.

### 3.2 Vabakasutus-kirjastamisega seotud kulud.

Uuringu käigus koostati ülevaade teadusajakirjade väljaandmisega seotud kulude kohta Eesti Teaduste Akadeemia Kirjastuse näitel (Ülo Niine, Eesti Teaduste Akadeemia Kirjastuse direktor).

Teaduste Akadeemia Kirjastus annab käesoleval ajal välja seitset rahvusvaheliselt eelretsenseeritavat ja juhtivates andmebaasides kajastatud teadusajakirja (Tabel 4). Kõigil ajakirjadel on rahvusvaheline toimetuskolleegium. Eesti Teadusagentuuri poolt hallatavas Eesti Teadusinfosüsteemis on kõik ajakirjad paigutatud kategooriasse 1.1.

**Tabel 4. Teaduste Akadeemia Kirjastuse ajakirjade kirjastamise mahud**

Nimetus	Numbrite arv aastas	Maht lehekülgedes
Acta Historica Tallinnensia	1	142
Estonian Journal of Archaeology	2	184
Estonian Journal of Earth Sciences	4	258
Linguistica Uralica	4	320
Oil Shale	4	392
Proceedings of the Estonian Academy of Sciences	4	474
Trames	4	430
Kokku	23	2200

Alates 2006. aastast on kõigi ilmunud ajakirjade artiklite täistekstid koos informatsiooniga ajakirjade kohta vabalt kättesaadavad kirjastuse kodulehel [www.kirj.ee](http://www.kirj.ee). Täistekste levitavad ka tuntud elektroonilised kirjastused ja portaalid EBSCO, C.E.E.O.L., The Gale Group Inc., ProQuest LLC, H. W. Wilson (ühines EBSCO-ga), Digital Publication with the Leading Asian Distributor (Airiti Inc.) ning Join CNKI Scholar (Hiina). Nende portaalide jaoks toodab ja kannab neile kirjastus artiklite järgmised elektroonilised versioonid: pdf-failid, kaht tüüpi sgml-failid, kolme tüüpi xml-failid ja spetsiaalse faili elektroonilise raamatukogu C.E.E.O.L. jaoks.

Ajakirjade varasemate aastakäikude tekstid on Google'i vahendusel digitaliseeritud ja üles pandud. Kirjastuse kodulehelt on tehtud vastavad lingid.

TLÜ Akadeemiline Raamatukogu on digitaliseerinud Eesti Teaduste Akadeemia väljaanded ajavahemikus 1945–1991. Algselt olid need kättesaadavad ainult raamatukogu sisevõrgus, kuna ©/(copyright) kuulus kirjastusele Perioodika, viimane on eksistentsi lõpetanud. 2016. aastal õnnestus välja selgitada kirjastuse Perioodika juriidiline õigusjärglane, kes lubas raamatukogul avada kõik digitaliseeritud väljaanded avalikus internetivõrgus, nüüd on need kõigile kättesaadavad.

Kõigi ajakirjade teadusartiklitele on pandud doi-indeksid (Digital Object Identifier). Crossrefi statistika näitab, et artiklite otsitavus doi järgi suureneb: on kuid, kus see ületab 8000 piiri.

Selline täistekstide kättesaadavuse paljus loob paremad võimalused artiklitega tutvumiseks, suurendab loetavust ja tsiteeritavust. Ajakirjades avaldatud artikleid kajastavad kümned rahvusvahelised referatiivajakirjad, teadusveebid ja andmebaasid. Allpool mõned autoriteetsemad.

Kõik kirjastuse poolt välja antavad ajakirjad on lülitatud paljudes riikides teadustöö produktiivsuse mõõdikuna kasutatavasse kirjastuse Elsevier BV poolt hallatavasse andmebaasi SCOPUS®. Vastav vorm (Coverage Form) 2017. aastaks on järjekordselt allkirjastatud.

Sageli teaduse hindamisel bibliomeetrilise informatsiooni alusena kasutatavas andmebaasis Thomson Reuters Web of Science® Core Collection on kajastatud kuus ajakirja:

1. Acta Historica Tallinnensia
2. Estonian Journal of Archaeology
3. Estonian Journal of Earth Sciences
4. Proceedings of the Estonian Academy of Sciences
5. Oil Shale (sh Current Contents®'is)
6. Trames. A Journal of the Humanities and Social Sciences

Kaks ajakirja on andmebaasis ERIH:

1. Estonian Journal of Archaeology
2. Linguistica Uralica

2016. aastal ilmus seitsme ajakirja 23 numbrit, kokku 162 kirjutist, nendest 145 teadusartiklit ja 17 lühiteadet. Elektrooniline versioon ilmus keskmiselt üks kuu enne väljakuulutatud paberkandja tähtaega, mis tähendab sisuliselt elektroonilise versiooni eelilmumist.

Autoreid oli kokku 302, neist Eesti autoreid 179 ja välismaalt 122. Enamik artikleid ilmus inglise keeles, v.a ajakirjas Linguistica Uralica, kus ilmus ka vene- ja saksakeelseid artikleid. Inglise keeles ilmus 11 kaastööd (38% mahust), vene keeles 26 (56% mahust), saksa keeles 3 (6% mahust). Inglisekeelsetest kaastöödest olid 9 teaduslikud uurimused ja 2 retsensioonid. Vene keeles ilmus peale teaduslike uurimuste ka enamik retsensioone ja ülevaateid, saksa keeles ilmusid 2 teadusartiklit ja 1 ülevaade. Et ajakirja autoritest moodustavad suure osa Venemaa soome-ugri rahvaste keeleteadlased, siis on vene keele määr täiesti loomulik. Acta Historica Tallinnensia ilmub eesti keeles põhjalike ingliskeelsete kokkuvõtetega ja ilmus ka üks ingliskeelne artikkel.

Teadada on kirjastuse eelarvelised võimalused ja kulutuste struktuur. Samas on raske määratleda iga ajakirja väljaandmise omahinda ja ühe lehekülje maksumust. Käsikiri käib läbi paljude töötajate, vt toimingud käsikirjaga. Tuleb määratleda töötaja osakaal mingi töö tegemisel, töö keerukusaste (siin tekib ilmselt üksjagu subjektiivsust), peab arvestama töötasu suurust, muude kulude suurust. Trükikulude osa on lihtne, see on teada.

Toimingud käsikirjaga

1. Registreerimine ja kontrollimine.

2. Retsenseerimine – 2 (mõnikord 3) retsensenti, üks retsensent välismaalt, teine Eestist.
3. Retsensioonide saatmine autorile paranduste tegemiseks.
4. Parandatud käsikirja ja paranduste kontrollimine, vajadusel kordusretsenseerimine.
5. Toimetamine (keel, stiil), kooskõlastamine autoriga.
6. Toimetaja paranduste sisestamine – sisestaja (osadel ajakirjadel teeb toimetaja).
7. Artikli küljendamine – küljendaja.

I korrektuur:

8. Korrektor kontrollib, et kõik toimetaja parandused on sisse viidud ning uusi vigu ei ole tekkinud.
9. Keeletoimetaja kontrollib küljendatud artiklit ja saadab selle autorile.
10. Keeletoimetaja teeb autori vastuste põhjal vajalikud parandused.
11. Küljendaja viib parandused sisse.

II korrektuur:

12. Korrektor võrdleb parandatud ja küljendatud artikleid, teeb vajalikud parandused.
13. Keeletoimetaja loeb veel kord artikli läbi ja kinnitab korrektori parandused.
14. Küljendaja viib viimased parandused sisse.

III korrektuur:

15. Korrektor kontrollib artikli üle.
16. Tehniline toimetaja teeb maketi (kaaned, leheküljenumbrid, paneb doi-numbrid jms).
17. Küljendaja trükib maketi välja.
18. Makett ja pdf failid saadetakse trükikotta.
19. Artiklite ülespanek kodulehele. Eelnevalt artiklite ettevalmistamine elektrooniliseks ülespanemiseks, viidete töötlemine – igale referentsile DOI lingi väljaotsimine (juhul kui on). Ülespaneku kontroll. Xml-failide ja sgm-failide loomine. DOI-numbrite aktiveerimine Crossrefis, selleks vastava xml faili allalaadimine. Pdf-, xml- ja sgml-failide saatmine partneritele: DIGAR, Gale, Proquest, (Wilson), Digital Publication with the Leading Asian Distributor (Taiwan), Join CNKI Scholar (Hiina), AGRICOLA jt.

20. Toimingud trükiversiooniga:

Ajakirjad transporditakse trükikojast tellijatele – Eesti Post, TLÜ Akadeemiline Raamatukogu, KE Kirjastus (vahendab saatmist välismaale), Ajaloo Instituut (2 ajakirja), erinumbrid konkreetsetele tellijatele, kirjastusest toimub müük poodidele (viiakse ise kohale) ja otsemüük soovijatele nii posti teel kui kohapeal.

Järgnev näide on aasta 2013 põhjal. Ilmus üheksa ajakirja, 32 numbrit kogutiraažiga 3860 eks, siit keskmine 120 eks. Täpsem maksumus on välja toodud tabelis 5.

**Tabel 5. Teaduste Akadeemia Kirjastuse ajakirjade hinnanguline väljaandmise maksumus**

Nimetus	No arv aastas	Maht (lk)	Teisendus- koefitsent ühisele formaadile	Maht formaadi kohta (lk)	Aasta hind (EUR)	Ühe lehekülje maksumus (EUR)
Acta Historica Tallinnensia	1	184	1	184	12 657	68,79
Estonian Journal of Archaeology	2	156	1	156	11 299	72,43
Estonian Journal of Earth Sciences	4	244	1,4	341	31 768	93,16
Estonian Journal of Ecology	4	294	1	294	21 280	72,38
Estonian Journal of Engineering	4	338	1	338	24 255	71,76
Linguistica Uralica	4	320	1,1	352	23 767	67,52
Oil Shale	4(5)	386(564)	1,1	425	33 737	79,38
Proceedings of the Estonian Academy of Sciences	4	284	1,4	398	37 643	94,58
Trames	4	448	1	448	30 486	68,05
Kokku	32	2200		2936	226 892	

Ühe lehekülje keskmine maksumus oli 77,3 eurot. Arvutused on tehtud ilma erinumbriteta ja ajakirja ülemahuta. Sama näitaja 2014. Aasta kohta on 83 eurot, mille juures ilmus kaheksa nimetust, 29 numbrit kogutiraažiga 3680 eksemplari. Hinnanguliselt on 2015 ja 2016 lehekülje maksumus jäänud ligikaudu samale tasemele kui 2014. aastal.

### 3.3 Teadusandmete haldamise ja arhiveerimisega seotud kulud

Teadusandmete arhiveerimise ja pikaajalise taaskasutatavuse tagamise üks peamisi takistusi on olnud lühiajalise projektirahastuse ja pikaajalise säilitamise reaalsete kulude konflikt. Vaatamata andmemahude üha kiiremale kasvamisele, ei tõuse andmeid säilitavate asutuste eelarved samas tempos. Seetõttu on parema planeerimise võimaldamine läbi andmehaldusplaanide andmearhiivide ühine huvi teadlastega.

Andmete haldamise plaani koostamine ja kulude prognoosimine on teadusprojektide taotlemisel suhteliselt uus nõue. Vähestel teadlastel on seni kogemusi andmehalduse hinnastamisel, enamasti piirdub see andmesalvestuse (ingl. k. *storage*) mahtude ja sellega seotud kuludega. Teadlaste abistamiseks teadusandmete elukäigu haldamisega seotud kulude arvestamisel on välja töötatud mõned kalkulaatorid.

**Andmehaldusega** seotud kulude kalkulaatoritest on seni põhjalikum UK Data Archive välja töötatud „UK Data Service Data Management costing tool“ (UK Data ..., 2015)<sup>41</sup>, mis tugineb

<sup>41</sup> [http://www.data-archive.ac.uk/media/247429/costing\\_v3.docx](http://www.data-archive.ac.uk/media/247429/costing_v3.docx)



kahes osas läbi viidud uuringul *Keeping Research Data Safe*.<sup>42</sup> Töövahend jaotab kulud järgmistesse kategooriatesse:

- Andmete kogumine
- Andmete sisestamine ja transkribeerimine
- Andmete valideerimine ja dokumenteerimine
- Andmete arhiveerimiseks ettevalmistamine
- Andmete taaskasutamine.

Nendest etappidest kõige kulumahukamaks osutusid sotsiaalteadustes kvalitatiivsete andmete transkribeerimine ja anonümiseerimine ning kvantitatiivsete andmete puhastamine ja valideerimine. Paljud teised andmetega seotud tegevused on samas integreeritud tüüpilise uuringu ja andmeanalüüsi tegevustesse ja nende maksumuse eraldi hindamine osutus keerukaks. Kulukalkulaatori kasutajate ühine tõdemus oli aga, et mida hiljem teadusprojekti elukäigus andmete ettevalmistamine avaldamiseks ja arhiveerimiseks ette võetakse, seda kulukamaks see kujuneb.<sup>43</sup>

Veelgi praktilisema suunitlusega ja detailsem andmehalduse kulude kalkulaator on värskelt valminud Hollandi ülikoolide koostöös (Westerhof et al., 2017).<sup>44</sup> Kalkulaatori jaotisteks on tüüpilised teadusuuringu tegevused, mida hinnastatakse tööaja ja reaalsete kulude vaates. Kalkulaator sobib kasutamiseks mitte üksnes teadlastele, vaid ka teadusprojektide rahastajatele ning aitab välja tuua tegevused, mis on vähem kulukad, kui need teostatakse projekti elukäigus võimalikult vara.

**Andmearhiivide** kulumudeleid on uuritud kauem kui andmehalduse plaanide maksumust (vt. Evaluation of Cost Models ..., 2014). Viimase viie aasta põhjalikem kuluarvutus valmis Hollandi andmearhiivil DANS (Palaiologk et al., 2012), mis kombineerib tasakaalus tulemuskaardi meetodi ja ABC kulumudeli (*activity-based costing*) rakendamisest andmearhiivi tegevusele. DANSi loodud *Cost Model for Digital Archiving* kalkulaatori kasutamine teiste riikide õigusruumides, kus raamatupidamislike kulu- ja varade arvestuse reeglid on teistsugused, osutus aga mõnevõrra keerukaks. Ühendkuningriigi erinevad teadusnõukogud on hinnanud oma kulutusi valdkondliku andmearhiivi pidamisele ca 1,2-1,5%-le oma teaduskulude eelarvest (The Royal Society, 2012, lk. 68).

Euroopa Komisjoni kaasrahastusega FP7 projekt 4C – *Collaboration to Clarify the Cost of Digital Curation*<sup>45</sup> – projekt analüüsis erinevaid digitaalse säilitamise mudeleid ja töötas välja kontseptuaalse referentsmudeli kuluarvestuse mudelite jaoks (Cost Concept Model ..., 2014). Selle alusel loodi digitaalse säilitamise kulude võrdlemise töövahend,<sup>46</sup> kuhu praeguseks on oma kuluandmed sisestanud 20 andmearhiivi. Kogutud andmetest nähtub, et kõige kulukam andmearhiivi tegevusetapp on reeglina andmete töötlemine ja kirjeldamine enne vastuvõtmist ning säilitamine (vt. Joonis 15). Kõige olulisem kuluartikkel (ca 65% kõigist

---

<sup>42</sup> <https://beagrie.com/jisc/>

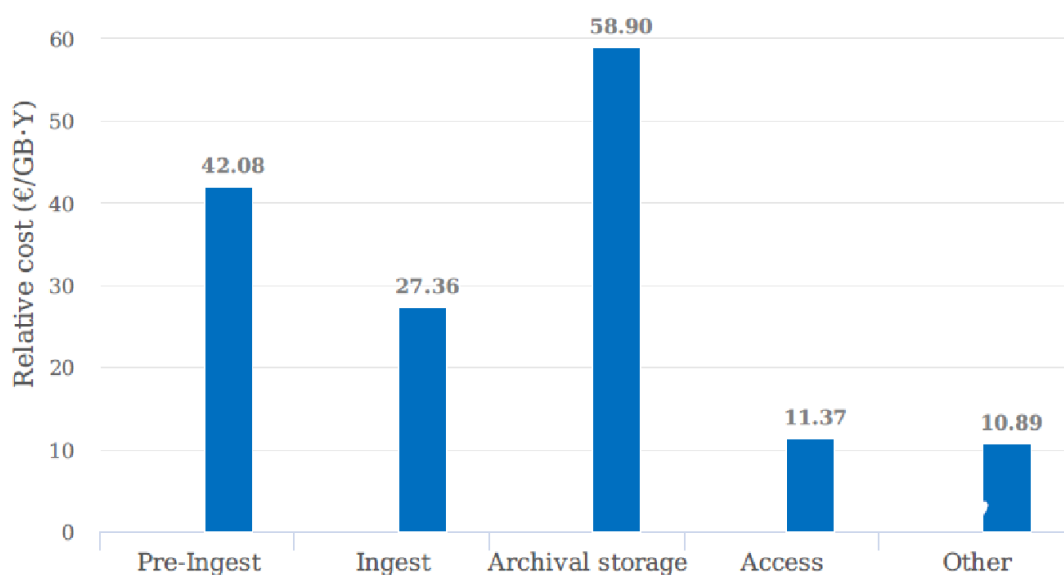
<sup>43</sup> <https://www.ukdataservice.ac.uk/manage-data/plan/costing>

<sup>44</sup> [https://www.uu.nl/sites/default/files/datamanagementcostguide\\_june2017.pdf](https://www.uu.nl/sites/default/files/datamanagementcostguide_june2017.pdf)

<sup>45</sup> <http://4cproject.eu/>

<sup>46</sup> <http://www.curationexchange.org/>

kuludest) on personalikulud, mis kasvavad vastavalt riigi üldisele palgatasemele ja maksupoliitikale.



**Joonis 15.** Andmearhiivide kulude jaotus tegevuste vahel (1 GB maksumus aastas) (allikas: CCeX).

Andmearhiivi kogukulud (s.h. personalikulud, riist- ja tarkvarakulud, teavitustöö ja üldkululõiv) taandatuna võrreldavale ühikule – gigabaidi andmete säilitamine aastas – annavad portaali Curation Cost Exchange tulemuseks (Euroopa andmearhiivide baasilt) keskmiseks maksumuseks 30 € / GB aastas. Sellele lisanduvad veel vajalikud investeeringud ja majandamiskulud, millega koos oleks ühe väiksema, ca 20 TB andmeid säilitava andmearhiivi aastane eelarvevajadus ca 100 000 €. Ühendkuningriikide Kuninglik Selts võrdles eri kategooria andmearhiivide reaalseid eelarvekulusid järgmiselt:<sup>47</sup>

- riiklikud andmearhiivid ja globaalsed platvormid – alates 2 milj € aastas (nt. UKDA 6 milj €);
- erialased andmearhiivid – alates 280 000 € aastas;
- institutsionaalsed digihoidlad – alates 200 000 € aastas.

4C projekti tööd jätkas spetsiifiliselt sotsiaalteaduse andmearhiivide kontekstis CESSDA SAW projekt, mis töötas välja kasumlikkuse tutvustamise vahendite paketi andmearhiividele.<sup>48</sup> Täpseid kulunumbreid projekt siiski andmearhiivide kohta ei juurde ei lisanud, vaid tugines Curation Cost Exchange andmetele.

**Avatud teadusandmete tasuvusuuringuid** on seni põhjalikumalt läbi viidud Ühendkuningriikides ja Austraalias. *Keeping Research Data Safe* projektid jõudsid erinevate distsipliinide andmearhiivide kulud ja kasutatavust analüüsides tulemuseni, et iga andmearhiivi investeeritud euro annab ühiskonnale tagasi 5,4 euro ulatuses väärtust (Beagrie

<sup>47</sup> <https://royalsociety.org/topics-policy/projects/science-public-enterprise/digital-repositories/>

<sup>48</sup> <https://www.cessda.eu/Projects/All-projects/CESSDA-SaW/WP4/Cost-Benefit-Advocacy-Toolkit>

& Houghton, 2014). Arvutuse aluseks on erinevate mõju hindamise meetodite kombinatsioon (vt. Joonis 16):

Investment & Use Value (Direct)	Contingent Value (Stated)	Efficiency Impact (Estimated)	Return on Investment (Estimated)
<b>ESDS</b> Based on Depositor & User Counts Investment Value £23m per annum Use Value £24m per annum	Based on User Count Willingness to Pay £25m per annum Willingness to Accept £82m per annum (£111m per annum)	Based on User Count Survey Community £68m-£112m per annum Wider Community £100m plus per annum	Based on Data Spend Increased Rol [From Additional Use] £58m - £233m 2.5 to 10-fold Rol
<b>ADS</b> Based on Deposit & Download Counts Investment Value £1.2m per annum Use Value £1.4m per annum	Based on Download Count Willingness to Pay £1.1m per annum (Constrained by capacity to pay) Willingness to Accept £7.4m per annum	Based on User Count User Community [ADS data use] £13m per annum User Community [All activity time] £58m per annum	Based on Data Spend Increased Rol [From Additional Use] £2.4m - £9.7m 2.1 to 8.3-fold Rol
<b>BADC</b> Investment Value £2.8m per annum Use Value £2.3m per annum	Willingness to Pay £5.2m per annum Willingness to Accept £16m per annum	User Community [BADC data use] £10m per annum User Community [All activity time] £58m per annum	Increased Rol [From Additional Use] (Non Recreate) £11m - £34m 4 to 11-fold Rol

**Joonis 16.** Kolme andmearhiivi kulude ja loodava väärtuse võrdlus (Beagrie & Houghton, 2014, lk. 14).

Austraalias uurisid kaks majandusteadlast avalikest andmetest saadavat hüve nii riigiasutuste (sh. riikliku statistika andmed, kaardiandmed, hüdrooloogilised andmed) kui Austraalia riikliku teadusandmete arhiivi ANDS kuluandmete põhjal (Houghton, 2011; Houghton & Gruen, 2014). Tasuvusanalüüsis kasutati järgmist valemit (vt. Joonis 17):

$$\frac{\text{Benefit}}{\text{Cost}} = \frac{\text{Agency savings} + \text{Users' savings} + \text{Increased returns to annual expenditure on PSI production}}{\text{Agency costs} + \text{Users' costs}}$$

**Joonis 17.** Teadusandmete tasuvuse arvutamise valem Austraalias läbi viidud tasuvusanalüüsis.

Üleriikliku hinnangu saamiseks arvestasid autorid kulude poolel riigieelarve kulutusi teadusele ja selle taristule tervikuna; tulude poolele kanti ka hinnanguline 80% teadusandmetest, mis on nõuetekohaselt arhiveerimata ja avaldamata ja millest seetõttu jääb oodatav kasu saamata. Võrreldes ja mudelleerides andmete kasutust enne ja peale nende avaldamist, näitas Austraalia uuring, et eri andmeliikide puhul on võimalik andmetest saadav kasu üsna erineva suurusega (nt. geo- ja ilmastiku andmete taaskasutusväärtus on märksa suurem kui folkloristika andmetel), kuid kõigi liikide puhul ületab saadav kasu (projekteerituna järgneva

20 aasta peale) andmete arhiveerimisele ja avaldamisele tehtavate kuludega. Üldine keskmine andmete avaldamise kasumlikkuse määr on sarnane Ühendkuningriikide uuringuga – 5,3 kordne kasu.

Sarnase uuringu läbiviimine **Eestis** eeldaks esmalt täpse ülevaate saamist olemasolevatest teadusandmetest, nii andmearhiivides ja digihoidlates talletatud kui teadlaste ja teadusasutuste valduses olevatest andmetest. Kuna kõiki teaduse valdkondi hõlmav andmete kaardistus kujuneks mahukaks, siis oleks otstarbekas esmalt käivitada katseprojekt, mille käigus võib kasutada taoliste kaardistuste jaoks mujal välja töötatud vahendeid (nt. Data Asset Framework<sup>49</sup>). Käesoleva uuringu käigus läbi viidud intervjuude põhjal oleks sobiv valdkond sotsiaalteadused, kus andmete haldamise ja avaldamise praktika on Eestis praegu väga ebaühtlane, samas andmete taaskasutusväärtus on oluline.

**Soovitus:** Käivitada katseprojekt sotsiaalteaduse andmete mahtude ja nendega seotud kulude kaardistamiseks ning kasutada saadud tulemusi kogu Eesti teadusandmete tasuvusanalüüsi teostamiseks.

---

<sup>49</sup> <http://www.data-audit.eu/>

## 4 Teadusandmete säilitamise korraldamine: stsenaariumid

*Lähteülesanne: alternatiivsete institutsionaalsete ja rolliliste perspektiivide analüüs. Aluseks dokumendianalüüs, küsitluse analüüs, intervjuude tulemused.*

Teadusandmete arhiveerimisest, avaldamisest ja säilitamisest tekkivad kasud on hästi dokumenteeritud ja analüüsitud teaduskirjanduses (nt. Vines et al., 2013; RIN, 2011). Läbi andmehaldusplaanide rakendamise on teadlikkus neist kasudest jõudsalt levimas teadusasutustes ja teadlaskonna seas, kes on korraga nii andmete loojad kui ka kasutajad. Eri distsipliinide andmekogumise, -kasutuse ja -halduse praktikad on erinevad, mistõttu on kõiki teaduse valdkondi võrdselt hästi teenindavaid andmearhiive keerukas tagada nii rahvusvahelisel, riigi kui institutsiooni tasandil. See peatükk vaeb erinevate huvirühmade vajadusi ning erinevate arhiveerimislahenduste eeliseid ja miinuseid.

Soovitus: Töstmaks kogukonna teadlikkust, tuleks Eestis vastu võtta Ühendkuningriikidega sarnane memorandum avatud teadusandmete haldamise sisuliste põhimõtetega *Concordat on Open Research Data* (2016).<sup>50</sup>

### 4.1 Teenuse komponendid

Teadusandmete haldamise taristu moodustavad koostoimivad:

- tehnoloogiad – nii riist- kui tarkvara, kiired ja suure läbilaskevõimega ühenduskanalid ning kõrge andmetöötluse suutlikkus;
- tehnoloogiatel rajanevad teenused, mis hoiavad andmeid kasutatavana, pakuvad juurdepääsu andmetele ja nende analüüsimise vahendeid;
- organisatsioonid, kes neid teenuseid pakuvad;
- regulatsioonid, standardid ja protokollid, mis juhivad teenuste pakkumist ja kasutamist;
- oskustega inimesed, kes osalevad nii teenuste pakkumises kui nende kasutamises.

Kuna teadus on rahvusvaheline, siis on kõigi nende komponentide sees omakorda keerukas kombinatsioon sellest, mida pakutakse asutuse, riigi ja rahvusvahelisel tasandil.

#### 4.1.1 Teenuse korraldajate ja kasutajate vaade

Teadusandmete arhiveerimise ja avaldamise teenuse korraldamisega otseselt seotud kasutajate ring koosneb:

- teadlased ja üliõpilased
- ülikoolid ja teised teadusasutused
- teaduse rahastajad ja hindajad
- andmearhiivid ja raamatukogud

<sup>50</sup> <http://www.rcuk.ac.uk/documents/documents/concordatonopenresearchdata-pdf/>

- teaduskirjanduse kirjastajad.

Sõltuvalt pakutavate teenuste koosseisust ja korraldusest ning säilitatavate andmete iseloomust jagunevad rollid ja nende täitmiseks vajalikud oskused just nende osapoolte vahel. Näiteks võib teadlaste nõustamine andmete haldamiseks ja arhiveerimiseks olla nii ülikooli, raamatukogu, kirjastaja või andmearhiivi ülesanne; andmete kirjeldamisega võib tegeleda teadusprojektis osalev doktorant või teadlane, kuid mõnede distsipliinide puhul saab selle ülesande delegeerida erialalisele andmearhiivile, kui seal töötavad vastava valdkonna spetsialistid; andmete kasutamise ja viitamise üle arvestuse pidamine võib olla nii ülikooli, raamatukogu kui ka teaduse rahastaja poolt rajatud keskse teenuse roll.

Laiem ring huvitatud osapooli ja kasutajaid on lai, sisaldades näiteks järgmiseid huvirühmi:

- õpetajad
- õppurid
- teadusest huvitatud ühiskonnaliikmed (*citizen-science*)
- ajakirjanikud
- poliitikakujundajad
- asutused
- ettevõtted
- kolmanda sektori organisatsioonid
- andmekaeve ja avaandmete töötlemisega seotud huvirühmad
- andmekaitse ja intellektuaalomandi kaitsega tegelevad organisatsioonid.

Idealne teadusandmete haldamise, avaldamise ja säilitamise platvorm on teadlase, teaduse rahastaja, teadusasutuse või ülikooli ja teiste seotud huvirühmade vaatenurgast erinev. Ühisteks omadusteks on alati lihtne juurdepääs taristule ja teenustele, mis toetaksid teadlaseid nii andmete loojana kui kasutajana, võimaldaksid andmeid töödelda, ilma, et eeldaksid sügavaid eelteadmisi taristu kasutamiseks. Samuti on ootus, et andmete säilitamise platvorm on hästi sidustatud rahvusvaheliste teaduskommunikatsiooni ja publitseerimise kanalitega ja et selle kasutamist on võimalik integreerida uuringute töövoogudega, sõltumata asukohast, kus teadlane oma uuringut läbi viimas on. Laiema kasutajate ringi huvi on saada üha enam töödeldud ja konteksti asetatud kujul andmeid, et neid oma vajadusteks ära kasutada. (*Lõppraportile lisatakse huvirühmade ja teenuste maatriks*)

#### 4.1.2 Andmete vaade

Arhiveerimise, säilitamise ja avaldamise teenuses osalevate andmete ja teenuste omavaheline seos on mitmetahuline. Teenuste sihtrühmast ja koosseisust sõltub reeglina see, milliseid andmeid kogutakse ja avaldatakse; andmete omadustest võib aga omakorda sõltuda see, milliseid teenuseid kasutajatele pakkuda on võimalik. Peamised andmete omadused, mis mõjutavad teenuste koosseisu on:

- teaduse valdkond, mille andmeid arhiveeritakse – sotsiaal-, humanitaar-, täppisteadused, jt.
- konkreetsed distsipliinid, mida teenindatakse – kristallograafia, keeleteadused, ajalugu, jt.
- andmete iseloom – kvantitatiivsed/kvalitatiivsed, kaasaegsed/ajaloolised, jne.
- andmete maht – suurandmed või tavaasuusega andmekogud.

Erinevate distsipliinide andmete ühised vajadused teenuste järele on vähemalt:

- mugavad vahendid andmete üleandmiseks arhiivile
- andmete kvaliteedi ja kirjeldatuse kontrollimine arhiivi ülevõtmisel
- püsiidentifikaatori omistamine
- turvaline hoid ja varundamine
- aktiivne digitaalne säilitamine ja erinevate kasutusvormingute loomine
- andmete provenientsi, tervikluse ja töötlemise andmete järjepidev loomine, et tõendada autentsust
- andmete indekseerimine, metaandmete avaldamine ja linkimine erinevatesse keskkondadesse
- mugavad kasutus- ja töötlemisvõimalused kasutajatele
- andmetele viitamise ja kasutamise üle arvestuse pidamine.

Nende teenuste koostoimel on võimalik tagada andmete turvaline säilimine ja kestev kasutatavus sihtrühmade jaoks.

#### 4.1.3 Teenuste ja tööprotsesside vaade

Sõltuvalt andmete arhiveerimise, säilitamise ja avaldamise teenuste korraldusest on teenused erinevalt pakettidesse kombineeritud. Tüüpiliselt jagunevad andmearhiivi teenused järgmiselt:

##### Põhiteenused:

- Andmete vastuvõtt andmete loojatelt
- Andmete hoid ja pikaajaline säilitamine
- Arhiveeritud andmete kataloogi pidamine, s.h. püsiidentifikaatorid
- Andmetele juurdepääsu tagamine, aktiivne avaldamine ja linkimine
- Andmete kasutamise ja viitamise üle arvestuse pidamine.

Iga põhiteenus koosneb veel tervest reast konkreetsest töövoogudest, mida üldisel tasandil ei ole otstarbekas välja tuua.

### Täiendavad teenused:

- Andmete loojate nõustamine ja toetamine (nt. juhendmaterjalid, koolitused) andmehalduse ja -arhiveerimise teostamiseks
- Andmete kasutajate nõustamine ja toetamine andmete kasutamiseks ja töötlemiseks
- Andmete ja nende kasutusvõimaluste tutvustamine ja oma teenuste turundamine kasutajatele
- Võrgustike loomine teiste andmehoidlate ja kasutuskeskkondadega, juurdepääsu (kasutuslitsentside) soetamine teiste andmearhiivide või kirjastuste andmekogudele
- Avatud juurdepääsu ja avatud teaduse propageerimine
- jne.

#### 4.1.4 Tehnoloogiate vaade

Teadusandmete kogumiseks ja haldamiseks vajalik tehniline taristu on tihti ulatuslik ja kallis. Andmete arhiveerimise, säilitamise ja avaldamise vajadused on suhteliselt tagasihoidlikud, koosnedes tüüpiliselt:

- tarkvara andmete kvaliteedi kontrollimiseks, töötlemiseks ja kirjeldamiseks nende vastuvõtmisel (konkreetsed vajadused sõltuvad andmete tüübist);
- digihoidla tarkvara andmekogude ja nende kirjelduste haldamiseks;
- püsiidentifikaatori omistamise tarkvara või sellega liidestumise võimekus;
- digitaalse säilitamise ja andmete konverteerimise tarkvaravahendid;
- kasutajaportaali andmete otsingu teostamiseks ja teenusepakkuja veebilehe jaoks;
- tarkvara andmete töötlemiseks kasutajateeninduses;
- tarkvaraliidesed andmete ja metaandmete vahetamiseks teiste teenuskeskkondadega;
- kasutajate autentimise tarkvaravahendid;
- kasutajatele pakutavad andmete töötlemise tarkvaravahendid;
- kontoritarkvara ja organisatsiooni toimimiseks vajalik tugitarkvara;
- riistvara digihoidla, arhiivis andmete töötlemise vahendite ja kasutuskeskkondade toetamiseks (võimsus sõltub andmekogude mahust ja analüüsivajadusest);
- arvutitöökohad arhiivi töötajatele;
- riistvara andmete turvaliseks hoiustamiseks ja varundamiseks;
- piisava võimsusega ühenduskanaleid nii andmeandjateni kui andmete kasutajateni.



#### 4.1.5 Regulatsioonide ja standardite vaade

Andmete arhiveerimise, säilitamise ja avaldamise teenuste toimimiseks vajalik õiguslik ja regulatiivne taustsüsteem on kirev ja mitmetasandiline, seal hulgas näiteks:

##### Rahvusvahelisel tasandil:

- Seaduslik taust andmete kogumiseks ja taaskasutamiseks
- Andmesubjektide õiguste kaitse
- Intellektuaalomandi kaitse
- Avatud teaduse ja juurdepääsu põhimõtted
- FAIR andmete põhimõtted
- Teaduse eetika põhimõtted.

##### Riigi tasandil (eelmisele lisaks):

- Teaduse rahastamise ja andmete haldamise reeglid
- Andmehaldusplaani rakendamise kohustus ja põhimõtted
- Teadusandmete taaskasutamise ja viitamise arvestamise reeglid teaduse hindamisel
- Andmete litsentseerimise põhimõtted
- Teadusprojekti rahastamisega kaasnevate andmete arhiveerimise reeglid

##### Teadusasetuse tasandil (eelmistele lisaks):

- Asutuse teadusandmete poliitika, kus näidatakse ära eelistatud andmete arhiveerimise teenus
- Täiendavad motivatsioonivahendid teadlastele andmete arhiveerimiseks ja avaldamiseks

##### Teadusprojekti tasandil (eelmistele lisaks):

- Andmehaldusplaan, mis määrab ära muu hulgas kasutatavad failivormingud, kirjelduse standardid ja reeglid, andmete kasutamise embargoperioodi pikkuse, jmt.

### Andmearhiivi tasandil:

- Andmete kogumise, haldamise ja kasutamise kord
- Andmete säilitamise poliitika, s.h. aktsepteeritavad failivormingud
- Ohuplaan, s.h. varundamise ja taaste kord
- Infoturbe reeglistik
- Andmearhiivi usaldusvärsuse tõendamise kord, nt. auditi materjalid
- Aruandluse reeglid
- jpm.

## 4.2 Teenuse organisatsiooni stsenaariumid

Teadusandmete arhiveerimise ja avaldamise teenuseid pakuvad praegu nii teadusasutused, riiklikud andmearhiivid, valdkondlikud andmearhiivid ja võrgustikud nii riigi kui rahvusvahelisel tasandil, erasektori teenusepakkujad ja kirjastajad, rahvusvahelised organisatsioonid ja teaduskogukonna initsiatiivid. Direktiiviga kehtestatud teadusandmete avaldamise nõude täitmiseks ei ole otseselt vaja mõnda neist teenusepakkujatest kõrvale jätta või suunata andmed vaid ühele teenusepakkujate grupile. Alljärgnev kirjeldab kombineeritud teenuste osapoolte tüüpilisi rolle ja hindab nende sobivust Eesti keskkonda.

### 4.2.1 Institutsionaalsete andmehoidlate võrgustik

Institutsionaalset digihoidlat (*institutional repository*) on defineeritud kui „teenuste paketti, mida asutus pakub oma kogukonna liikmetele asutuse ja selle kogukonna poolt loodud digitaalse ainese haldamiseks ja avaldamiseks“ (Lynch, 2003). Asutuste digihoidlad pakuvad mitmesuguseid hoiu ja avaldamisega seotud teenuseid, vahendades ja turundades asutuse akadeemilist toodangut üleilmselt. Teadusandmed on üha sagedamini arhiveerimisteenuste paketi osa, kuna seeläbi saab asutus toetada oma teadlaseid andmete avaldamise nõude täitmisel. Raamatukogude spetsiifikaks olevad info- ja säilitamisteadused on abiks üha kasvavate andmemahutuste eesmärgipärasel haldamisel. Lisaks on raamatukogud hakanud pakkuma ülikooli teadlastele abi andmehaldusplaanide koostamisel, andmete kirjeldamisel, õiguste (autoriõigus, andmekaitse) ja litsentside reeglites orienteerumisel, andmete haldamisel teadusprojekti käigus, andmete säilitamiseks sobivate failivormingute ja praktikate rakendamiseks, andmete taaskasutamiseks õppetöös, jmt. Institutsionaalsed andmehoidlad toetavad enamasti ka andmetele viitamise ja taaskasutamise aruandlust.

Institutsionaalne andmehoidla võib olla asutatud ülikooli raamatukogu juurde (nt. DataDOI<sup>51</sup> Tartu Ülikooli raamatukogus), eraldi ülikooli allasutusena (nt. Edinburghi Ülikooli DataShare<sup>52</sup>) või ka mitme ülikooli koostöös (nt. 4TU konsortsium Hollandis<sup>53</sup>).

---

<sup>51</sup> <https://datadoi.ut.ee/>

<sup>52</sup> <https://datashare.is.ed.ac.uk/>

<sup>53</sup> <http://researchdata.4tu.nl/en/home/>

Institutsionaalset andmehoidlat teenindab ülikooli raamatukogudes keskmiselt 2,3 töötajat (Hollandi ja Ühendkuningriigi andmete põhjal), kelle peamisteks ülesanneteks on andmete kontrollimine, töötlemine ja kasutajate (nii andmete loojate kui taaskasutajate) teenindamine ja teavitustöö. Eelmises peatükis hinnati institutsionaalse andmehoidla aastase eelarve vajaduseks suurusjärku 150-200 000 €. Rahastuse tagajateks on valdavalt omanik-ülikoolid, kuid mõnel määral kasutatakse ka teadusprojektide eelarvest tasu võtmist pakutavate teenuste eest.

Eesti kontekstis vaadatuna on selle stsenaariumi teostamise SWOT analüüs järgmine:

<b>Tugevused</b>	<b>Nõrkused</b>
<p><b>Lähedus teadlasele, tuttav keskkond</b></p> <p><b>Olemas DataCite võrgustik ülikoolide raamatukogude tasandil</b></p>	<p>Puudub ülikoolide tugi andmehoidlate rajamiseks ja pidamiseks</p> <p>Keerukas katta paljusid erinevaid distsipline oma erinõuetega andmetele ning suuta lisada andmetele lisandväärtust arhiivis</p> <p>Teadusandmete haldamise ja säilitamise oskustega töötajate nappus</p>
<b>Võimalused</b>	<b>Ohud</b>
<p><b>Lihtsam luua aruandlus ülikooli teadusloome ja selle taaskasutamise kohta</b></p> <p><b>Võimalik tugineda ülikooli IT taristule</b></p> <p><b>Võimalus teha ülikoolide vahel koostööd andmehoidla pidamiseks ja teadlaste teenindamiseks</b></p> <p><b>Mõnevõrra lihtsam planeerida taristu laiendamise vajadusi, kuna on teada ülikoolis käimasolevate ja lõppevate projektide arv</b></p> <p><b>Olemas on e-Varamu portaal, mida on võimalik edasi arendada teadusandmete vahendamiseks</b></p>	<p>Ülikoolide digihoidlad on praegu nõrgalt arenenud: andmehoidla olemas vaid TÜR-is, teistel teadusraamatukogudel puudub tarkvaraline ja sisuline tugi andmehoidlale</p> <p>Vajadus luua digitaalse säilitamise kompetents paljudes keskustes</p> <p>Taristu haldamiseks vajalikud kulud tuleb katta kõigil andmehoidlaid pidavatel asutustel</p>

Võimalik näide: institutsionaalsed andmehoidlad andmehaldusplaanide teostajana

Eelkirjeldatud kulu tegureid arvestades võiks institutsionaalsetele andmehoidlatele rajatud avatud teadusandmete taristu kujuneda järgnevast rollijaotusest (NB! kirjeldatud on vaid ühte võimalikku stsenaariumi, mis lähtub taristu sõnastatud eesmärgist).

<b>Tegur</b>	<b>Kirjeldus</b>
Eesmärk	Juurutada lühikese ajaga andmehaldusplaanide kasutamine teadusprojektides ja tagada edaspidi nende raames loodud andmete turvaline arhiveerimine ja teaduslikku taaskasutusse toomine
Andmete fookus	Määratud ajahetkest alates rahastatavate ja algatatavate teadusprojektide andmed, millel on taaskasutusväärtus teiste teadlaste jaoks ning mis on ette nähtud projektide andmehaldusplaanides  Digihoidlasse kogutakse oma teadusasutuses loodud andmeid
Teenuste fookus	Andmehaldusplaanide loomise ja täitmise nõustamine, toetamine töövahenditega ning arhiveerimise ja avaldamise teostamine
Kompetentside fookus	Teadusasutuse või ülikooli uurimissuundade (eripäraste) andmete teenindamiseks vajalikud tehnilised teadmised; suutlikkus nõustada teadusprojekte andmehaldusplaanide koostamisel ja rakendamisel; teadusandmete (teadusvõrgustikes) avaldamiseks vajalikud kompetentsid
Tehnoloogia vajadused	Teadusandmete haldamise taristu teadusprojekti käigus  Digihoidla riist- ja tarkvara erinevates ülikoolides  Andmete avaldamislahendused ja ühendatus valdkondlike võrgustikega
Regulatsioonide fookus	Andmehaldusplaanide esitamise ja järgimise nõue  Andmete arhiveerimise kohustuse kehtestamine  Andmete taaskasutamise ja viitamise teadustegevuse tulemushindamise osaks tegemine
<b>Teenus</b>	<b>Kirjeldus</b>
Andmehaldusplaanide rakendamine	Võimalik teostada kas läbi keskse-/ühisteenus (nt. ETAIS) või ülikoolide ja teadusasutuste kohapealse taristu ja teenuste baasil
Nõustamisteenused	Võimalik pakkuda teadlasele lähedal asuva raamatukogu või teadusasutuse oma teenusepakkuja kaudu

Andmete arhiveerimine	Võimalik jagada vastutus teadusprojekti ja raamatukogu või teenust pakkuva üksuse vahel
Andmete avaldamine	Vajalik lahendada digihoidla ja valdkondlike ning rahvusvaheliste võrgustike koostöös
Aruandlus	Võimalik lahendada digihoidla baasil
<b>Osapool</b>	<b>Roll</b>
Teadlane	Andmehaldusplaani koostamine ja arhiveeritavate andmete valik, kasutuspiirangute (nt. embargoperiood) määramine
Teadusprojekt	Andmehaldusplaani teostamine, andmehalduse töövahendite kasutamine Andmete arhiveerimine projekti käigus ja lõppedes
Ülikooli raamatukogu	Teadlaste ja teadusprojektide nõustamine andmehaldusplaanide koostamisel Andmehalduse teostamise nõustamine teadusprojekti käigus Projekti käigus loodud andmete arhiveerimise juhendamine ja/või teostamine Andmehoidla haldamine Andmete säilitamine Andmete avaldamine teaduslikku taaskasutust tagavates võrgustikes (nii Eestis kui rahvusvaheliselt) Aruandlus andmete taaskasutusest ja viitamisest nii teadlasele kui teadusasutusele
Teadusasutus või ülikool	Digihoidla pidamise lubadus ja vajaliku rahastuse tagamine Teadlase tunnustamine teadusandmete taaskasutamise eest
ETAG	Andmehaldusplaanide ja andmete arhiveerimise nõude kehtestamine ja jälgimine
<b>Tulumudelid</b>	<b>Kirjeldus</b>

Teaduasutuse otsetoetus	Kuna teadustulemite kogumine, avaldamine ja talletamine on ülikooli kohustus ja huvides, on teadusandmete arhiveerimise taristu ja teenuste ülevalpidamine ülikoolile kasulik, pakkudes ühtlasi lisateenuse oma töötajatele
Lõiv andmete arhiveerimise eest	Andmehaldusplaaniga ette nähtud lõiv teadusgrantist, mis tagab projekti tulemandmete korrektse ja turvalise arhiveerimise, hoiu ja kontrollitud kasutamise ning aruandluse andmete kasutamise kohta
Lisandväärtusega andmete kasutuse eest tasu võtmine	Andmehoidla võib avatud andmete teatud kasutusjuhtude puhul pakkuda tasulisi teenuseid andmete nt. pakendamiseks, linkimiseks, konverteerimiseks, anonümiseerimiseks jm viisil töötlemiseks

#### 4.2.2 Riiklik andmearhiiv

Enamus olemasolevaid andmearhiive on spetsialiseerunud mõne valdkonna või distsipliini andmete talletamisele ja vahendamisele. Nende rolliks on koguda erinevatest allikatest pärinevad andmekogud ühte kohta, varustada need kasutamiseks vajalike kontekstiandmetega ja aktiivselt andmeid taaskasutuseks pakkuda. Andmearhiivide tugevuseks on kujunenud tealastest sihtrühma ja kasutajaskonna tundmine ning suutlikkus rääkida kaasa valdkonna uurimisprojektide andmeid puudutavatel teemadel. Andmearhiividest on saanud kompetentsikeskused, mis keskenduvad säilitamisele, edendavad standardeid, juhendavad teadlaseid andmete ja metaandmete teemadel ning nõustavad teaduse rahastajaid strateegilistes andmetega seotud küsimustes. Andmearhiivi püüdlus on olla n.ö. *one-stop-shop* oma valdkonna teadlaste jaoks, mis pakub õppe- ja uurimistööks vajalikke andmeid (s.h. vahendatud andmeid, mille kasutuslitsentse soetatakse teistelt andmeomanikelt), abistab andmete jooksva haldamisega seotud küsimustes, korraldab koolitusi ning pärast arhiveerimist hoolitseb andmete säilitamise, linkimise (s.h. publikatsioonidega), kasutajate toetamise ning turundamise eest, koostades takkapihta ka vajalikke aruandeid, mida teadusprojekt peab esitama projekti rahastajale või teadlane oma instituudi tegevuse hindamiseks. Suutlikkus spetsialiseeruda ühe teadusvaldkonna vajadustele ning tugev baas digitaalse säilitamise teostamiseks, mille usaldusväärsus on asunud auditeerima<sup>54</sup> on põhilised konkurentsieelised institutsionaalsete digihoidlate ees. Paljude valdkondade andmearhiivid on moodustanud üle-Euroopalisi võrgustikke, mille kaudu jagatakse üksteisega nii paremat praktikat kui ka oma riigis kogutud andmeid, nt. CESSDA<sup>55</sup> sotsiaalteaduste andmearhiividele, CLARIN<sup>56</sup> keeleteaduse ressursside arhiividele, SHARE<sup>57</sup> tervise ja vananemise andmebaasidele, jt.

<sup>54</sup> vt. Data Seal of Approval (<https://www.datasealofapproval.org/en/>) ja Core Trust Seal (<https://www.coretrustseal.org/>)

<sup>55</sup> <https://www.cessda.eu/>

<sup>56</sup> <https://www.clarin.eu/>

<sup>57</sup> <http://www.share-project.org/>

Andmearhiivide asutajateks on reeglina teadusagentuurid või teaduste akadeemiad, kuid nende igapäevane pidamine on enamasti usaldatud ülikoolidele, uurimiskeskustele ja laboratooriumidele.

Valdkondlikus andmearhiivis töötab suuremates riikides tüüpiliselt 40-60 inimest, väiksemates riikides 25-30 töötajat (nt. Soome sotsiaalteaduslikus andmearhiivis ca 30<sup>58</sup>). Andmearhiivide aastased eelarved ulatuvad reeglina üle miljoni euro, kuigi kuna kaks kolmandikku kuludest moodustab palgakulu, siis on reaalsed eelarved riigiti erinevad. Andmearhiivide rahastamise mudelid on kombineeritud paljudest allikatest korruga (vt. ka OECD, 2017):

- Osalevate institutsioonide toetus – reeglina andmearhiivi majutava ülikooli tugi, nii rahaline kui tugifunktsioonide ja taristu näol
- Andmeandjatelt võetav tasu andmete arhiveerimise ja säilitamise eest; enamasti ette nähtud andmehaldusplaani kaudu teadusprojekti rahastamisel
- Konkreetsetes valdkonnas tegutsevate teadusasutuste siht-toetus, et tagada valdkonna teaduse jaoks vajalik taristu, s.h. andmete kättesaadavus
- Riiklik rahastus projektide või uute teenuste arendamise kaudu
- Omatulu läbi arendatud äriteenuste, millega vahendatakse töödeldud andmeid eraettevõtetele
- Kasutustasud mahukat töötlust nõudvate andmekogude kasutajateni viimise eest
- Üldkululõiv, mida ülikoolid teadusprojektidelt koguvad ja millest osa eraldatakse andmearhiivi tegevuse toetuseks
- Teaduse rahastajatelt saadav otsetoetus taristu ülalpidamiseks ja investeeringuteks.

Andmearhiivide toimimismudeleid ja rahastusvõimalusi on viimastel aastatel korduvalt uuritud ja kõrvutatud (Ember & Hanisch, 2013; Kitchin, 2015; OECD, 2017) ning antud soovitusi nende rahastusmudeleid ühtlustada. OECD (2017) analüüsis 47 andmearhiivi, millel oli kokku 95 erinevat tuluallikat (vt Joonis 18).

---

<sup>58</sup> <http://www.fsd.uta.fi/en/contact/staff.html>



- 1) Structural (central contract)
- 2) Hosting Support (indirect or direct support through institutional hosting)
- 3) Annual Contract (from depositing institution)
- 4) Data Deposit Fee (may be paid by researcher, RPO or publisher, may originate with funder)
- 5) Access Charge (for the data or for value-adding services)
- 6) Projects (to develop infrastructure or value-adding services)
- 7) Private Contracting (services to parties other than core funder)

**Joonis 18.** Teadusandmete arhiivide rahastusallikad (OECD, 2017, lk. 22).

Eesti vanim valdkondlik andmearhiiv – Eesti Sotsiaalteaduslik Andmearhiiv ESTA<sup>59</sup> – mis asutati 1996. aastal, on viimasel kümnendil vaevelnud eelarvevahendite nappuse käes. CESSDA SaW projekti analüüs Euroopa sotsiaalteaduste andmete arhiveerimise seisundist liigitab Eesti sotsiaalteadusliku andmearhiivi 20 aastat peale selle asutamist ikka kategooriasse „*promising infrastructure building proto-activities*“ (CESSDA, 2017, lk. 24-25). Ühe selle olukorra põhjusena tuuakse raportis välja poliitika ja kordade puudumine, mis ei taga andmearhiivi teenuse jätkusuutlikuseks vajalikku andmete kogust ja rahastuse määra.

Eesti kontekstis vaadatuna on selle stsenaariumi teostamise SWOT analüüs järgmine:

Tugevused	Nõrkused
<p><b>Tuginemine üle-Euroopalistele võrgustikele tagab kõrge kompetentsusastme</b></p> <p><b>Teadusvaldkonna andmete probleemistiku tundmine, mida teenindatakse</b></p> <p><b>Keskendumine lisaks andmete avaldamisele ka nende säilitamisele</b></p>	<p>Praegu vaid paar andmearhiivina toimivat teenusepakkujat: ESTA, Eesti Keeleressursside Keskus</p> <p>Puudub andmearhiivide traditsioon ja vastav regulatsioon, mis motiveeriks teadusprojekte oma andmeid andmearhiivi deponeerima</p> <p>Andmearhiivide jaoks vajaliku taristu ja tarkvara haldamise kogemused on napid</p>
Võimalused	Ohud
<p><b>Andmearhiivide mudel on Euroopas hästi sisse töötatud ja seda on lihtne kopeerida;</b></p>	<p>Piisava rahastuse puudumine andmearhiivi asutamiseks ja/või ülalpidamiseks, kuna</p>

<sup>59</sup> <http://www.psych.ut.ee/esta/>



<p><b>olemas on andmearhiivi asutamise juhendid<sup>60</sup></b></p> <p><b>Mõningaid kulukaid kompetentse, nt. digitaalse säilitamise jaoks, on vaja luua vaid üksikutes keskustes</b></p> <p><b>Olemas on e-Varamu portaal, mida on võimalik edasi arendada teadusandmete vahendamiseks</b></p>	<p>rahastusmudelid on väikeses riigis piiratud</p> <p>Paljudes teadusvaldkondades puutub kriitiline mass andmeid, et õigustada valdkondliku andmearhiivi asutamist</p> <p>Nappide vahendite tõttu ei ole võimalik luua andmearhiive iga teadusvaldkonna jaoks</p>
--	---

Võimalik näide: riiklik andmearhiiv eesti teadusandmete täieliku koguna

Eelkirjeldatud kulu tegureid arvestades võiks riikliku andmeandmearhiivi pidamisel rollide jaotus olla järgnev (NB! kirjeldatud on vaid ühte võimalikku stsenaariumi, mis lähtub taristu sõnastatud eesmärgist).

<b>Tegur</b>	<b>Kirjeldus</b>
Eesmärk	Moodustada ja arhiveerida võimalikult täielik eesti teadusandmete kogu, tagada selle pikaajaline säilimine ning teha see kättesaadavaks nii teadlastele kui laiemale avalikkusele
Andmete fookus	Eelkõige juba olemasolevad teadusandmete kogud, mis seni on arhiveerimata ja mille säilimine ei ole tagatud  Käimasolevate teadusprojektide raames tekkivad andmekogud
Teenuste fookus	Teadusandmete kaardistamine ja arhiiviväärtuslike andmekogude identifitseerimine  Andmete arhiveerimine, sh töötlemine, kirjeldamine Andmehaldusplaanide alusel andmete vastuvõtmine  Andmete aktiivne säilitamine  Andmete avaldamine erinevates kanalites, s.h. riigi avaandmete portaalid, ettevõtted, asutused, teaduse populariseerijad, jt.  Teavitustöö andmete taaskasutamiseks

<sup>60</sup> CESSDA SaW. *Guide for Developing National Data Service Plans* (2016)  
<https://cessda.net/content/download/1065/9156/file/CESSDA%20SaW-3.3-guide.pdf>

Kompetentside fookus	<p>Teadusandmete tekkimise ja haldamise eripärad, arhiveerimise vajadused</p> <p>Digitaalse säilitamise teostamiseks vajalikud teadmised ja oskused</p> <p>Teadusandmete taaskasutuse ja avaandmetega seotud kompetentsid</p> <p>Teadusandmete haldamise strateegilised probleemid</p>
Tehnoloogia vajadused	<p>Vahendid teadusandmete kaardistamiseks ülikoolide ja uurimisprojektide salvestuskeskkondades</p> <p>Paindlikud andmete arhiveerimise ja kirjeldamise lahendused</p> <p>Andmearhiivi tark- ja riistvara</p> <p>Andmete avaldamislahendused</p> <p>Ühendatus valdkondlike ja riiklike võrgustikega</p>
Regulatsioonide fookus	<p>Andmearhiivi statuut keskse teenusepakkujana ja volitused andmeid koguda</p> <p>Andmete arhiveerimise kohustuse kehtestamine</p> <p>Andmete taaskasutamise ja viitamise teadustegevuse tulemushindamise osaks tegemine</p> <p>Ajalooliste teadusandmete taaskasutusse andmise lubade ja kasutuspiirangute kehtestamise regulatsioon</p>
<b>Teenus</b>	<b>Kirjeldus</b>
Olemasolevate teadusandmete kogude kaardistamine	Võimalik kasutada teistes riikides välja töötatud kaardistusvahendeid (nt. DAF) ja tugineda ülikoolide abile
Nõustamisteenused	Keskse andmearhiivi teenusena, piiratud valdkondade eripäradega arvestamise suutlikkus
Andmete arhiveerimine	Põhitöö teostab andmearhiiv, võimalusel abi teadlaselt või uurimisprojekti liikmetelt
Andmete säilitamine	Andmearhiivi põhikompetents ja teenus, võimalik tugineda mõne olemasoleva teenusepakkuja taristul
Andmete avaldamine	Andmearhiivi ja valdkondlike, riiklike ning rahvusvaheliste võrgustike koostöö

Aruandlus	Võimalik lahendada andmearhiivi baasil
<b>Osapool</b>	<b>Roll</b>
Teadlane	Juurdepääs lõppenud teadusprojektide andmetele ja arhiveeritavate andmete valikus ja arhiveerimises osalemine, kasutuspiirangute määramine
Teadusprojekt	Andmete arhiveerimine projekti käigus ja lõppedes
Andmearhiiv	Projekti käigus loodud andmete kaardistamine ja arhiveerimise teostamine Andmearhiivi haldamine Andmete säilitamine Andmete avaldamine kasutusvõrgustikes (nii Eestis kui rahvusvaheliselt) Andmete taaskasutuse propageerimine, andmete pakendamine uuteks kasutusvõimalusteks Aruandlus andmete taaskasutusest ja viitamisest nii teadlasele kui teadusasutusele
Teadusasutus või ülikool	Teadlase tunnustamine teadusandmete arhiveerimise ja taaskasutamise eest
ETAG	Andmehaldusplaanide ja andmete arhiveerimise nõude kehtestamine ja jälgimine Andmearhiivi kui keskse teenusepakkuja staatuse kehtestamine
<b>Tulumudel</b>	<b>Kirjeldus</b>
Riiklik otsetoetus	Andmearhiivi kui teadustaristu keskse teenuse rahastamine tagatakse HTMi eelarve raames
Lõiv andmete arhiveerimise eest	Andmehaldusplaaniga ette nähtud lõiv teadusgrantist, mis tagab projekti tulemandmete korrektse ja turvalise arhiveerimise, hoiu ja kontrollitud kasutamise ning aruandluse andmete kasutamise kohta
Lisandväärtusega andmete kasutuse eest tasu võtmine	Andmearhiiv võib avatud andmete teatud kasutusjuhtude puhul pakkuda tasulisi teenuseid andmete nt. pakendamiseks, linkimiseks, konverteerimiseks, anonümiseerimiseks jm viisil töötlemiseks

Konsultatsioonid ja nõustamine	Andmete töötlemise ja andmeteaduse alased tasulised konsultatsioonid, mida andmearhiiv kompetentsikeskusena võib pakkuda erinevatele sektoritele
Litsentsitasu andmete vahendamise eest	Võimalus andmeid vahendada rahvusvahelistesse võrgustikesse ja erinevatele kasutusalaadele, milleks luuakse tasuliste kasutuslitsentside süsteem (saadud tulu on võimalik jagada ka andmete loojatega)

#### 4.2.3 Rahvusvahelised teenuste platvormid

Erinevaid teadusandmetega seotud teenuseid on hakanud pakkuma nii ärilised kui kasumit mittetaotlevad pakkujad. Neist Figshare<sup>61</sup> pakub tuge andmete haldamisele kogu teadusprojekti elukäigu jooksul ning omistab DOI püsiidentifikaatoreid; teenused on valminud ka nii teadusasutuste kui kirjastajate jaoks. Figshare kasutab n.n. *freemium* mudelit, kus tasuta on kasutada piiratud maht isiklikku kettaruumi, mida on võimalik tasu eest juurde soetada, kui samas piiramatult avalikustatud andmete kettaruumi andmete hoidmiseks. Dryad<sup>62</sup> sai alguse teadusajakirjade tugiteenusena, mis pakkus võimalust hoida bioloogia ja ökoloogia publikatsioonide alusandmeid, kuid on nüüdseks laiendanud nii teenuste kui valdkondade hulka. Zenodo<sup>63</sup> asutasid CERN ja OpenAire projekt, et pakkuda salvestusruumi iga valdkonna teadusandmetele, omistada neile DOI püsiidentifikaatorid ja koostada aruandeid projektide rahastajatele. Scratchpads<sup>64</sup> pakub sarnaseid andmesalvestuse teenuseid kitsamalt bioloogia valdkonnas. Kirjastajate hulgas on Elsevier<sup>65</sup> olnud aktiivseim teadusandmete haldamise jaoks töövahendite väljatöötamises, mille hulka kuulub nüüd ka Mendeley<sup>66</sup>.

Euroopa Liit on võtnud eesmärgiks välja arendada Euroopa avatud teaduse pilve (*European Open Science Cloud*), mis on mõeldud olema usaldusväärne, avatud platvorm, kus teaduskogukond saab oma andmeid ja tulemeid salvestada, jagada ja taaskasutada.<sup>67</sup> Pilvelahendus tugineb suure võimsusega teadustaristul, kuid peaks hakkama liitma ka olemasolevaid riikide andmetaristuid. Esimesed pilootprojektid taolise ühiskasutuse keskkonna valitsemise ja toimimise testimiseks on käivitunud,<sup>68</sup> Euroopa Komisjon on eesmärgiks seadnud EOSC püsiva rahastusmudeli väljatöötamise 2017.a. lõpuks. Sarnane ühistu-põhimõttel toimiv andmete taristu on juba varem loodud Eudat projekti raames.<sup>69</sup>

Kohalike andmearhiivide, aga ka institutsionaalsete digihoidlate kogud on üha enam nähtavad ja lingitud rahvusvahelistesse portaalidesse, mille kaudu soovivad andmekogusid üles leiavad,

<sup>61</sup> <https://figshare.com/about>

<sup>62</sup> <http://datadryad.org/pages/organization>

<sup>63</sup> <http://about.zenodo.org/>

<sup>64</sup> <http://scratchpads.eu/>

<sup>65</sup> <https://www.elsevier.com/about/open-science/research-data>

<sup>66</sup> <https://data.mendeley.com/>

<sup>67</sup> <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

<sup>68</sup> <https://eoscpilot.eu/about-eoscpilot>

<sup>69</sup> <https://www.eudat.eu/what-eudat>

nt. <http://www.databib.org/>, <http://www.re3data.org>, <http://www.opendoar.org>, <https://www.openaire.eu/>, <http://roar.eprints.org/>, <http://www.datacite.org/>, jt.

Teadusandmete arhiveerimiseks, säilitamiseks ja avaldamiseks vajalik teenuste pakett on võimalik kokku kombineerida erinevate teenusepakkujate lahendustest. Ühendkuningriigis on näiteks JISC sõlminud lepingud lausa mitme erineva kommerts-teenusepakkujaga, kes teenindavad teadusasutusi.<sup>70</sup> Sellise kompleks-teenuse tervikmaksumust ei ole paraku võimalik praegu veel välja arvutada, kuid üha suurem Euroopa Liidu aktiivsus sellel suunal loob lähiaastatel uued võimalused, mis võivad kujuneda arvestatavateks alternatiivideks asutuse- ja riigipõhiste lahendustele. Praeguses kontekstis on selle stsenaariumi teostamise tugevused ja nõrkused järgmised:

Tugevused	Nõrkused
<p><b>Teenused rakendavad rahvusvahelist paremat praktikat ja tänu suurele kasutajate arvule on pidevalt uuendamissuutlikud</b></p> <p><b>Talletatud andmekogud on koheselt nähtavad rahvusvahelistes võrgustikes ja võimalus nende taaskasutamiseks on seeläbi suurem kui kohalikus andmearhiivis avaldades</b></p> <p><b>Teenuse mudelid motiveerivad andmeid avaldama projekti käigus esimesel võimalusel (nt. tasuta kettaruum, lisateenused saadaval proportsionaalselt avaldatud andmetega)</b></p> <p><b>Loodud on tooted ja teenused toetamaks andmehaldust ka teadusprojekti või uuringu teostamise käigus</b></p>	<p>Teenusepakkujad ei ole Eestis koha peal, ega suuda pakkuda eestikeelset või kohalikke olusid arvestavat tuge</p> <p>Agregaator-tüüpi teenusepakkujad ei keskendu üksikutele teaduse valdkondadele, mistõttu on kasutatavad vormingud, standardid ja nõuded üldised</p>
Võimalused	Ohud
<p><b>Euroopa Liidu tugi paljudele teenustele on täiendavaks garantiiks teenuse laiale levikule ja kestvusele</b></p> <p><b>Tänu teenuste ja platvormide rahvusvahelisusele saavad Eesti teadustulemid rohkem silma paista</b></p> <p><b>Suund on võetud andmehalduse ja kirjastamisprotsessi tugevamale seostamisele või isegi integreerimisele</b></p>	<p>Rahvusvahelise teenusepakkuja jätkusuutlikkust on raske kontrollida ja ennustada</p> <p>Andmete arhiveerimise, säilitamise ja avaldamise täisteenuse saamiseks tuleb erinevaid teenusepakkujaid omavahel kombineerida ja nende platvormid omavahel liidestada</p>

Võimalik näide: teadusprojekti andmete haldus teenusepakkujate toel

<sup>70</sup> <https://www.jisc.ac.uk/data-archiving-framework>, <https://arkivum.com/jiscarchiving/> ja <https://researchdata.jiscinvolve.org/wp/2017/11/29/research-data-shared-service-well-alpha/>

Eelkirjeldatud kulu tegureid arvestades võiks olemasolevatel teenustel tuginev teadusandmete haldamine kujuneda järgmiseks (NB! kirjeldatud on vaid ühte võimalikku stsenaariumi, mis lähtub taristu sõnastatud eesmärgist).

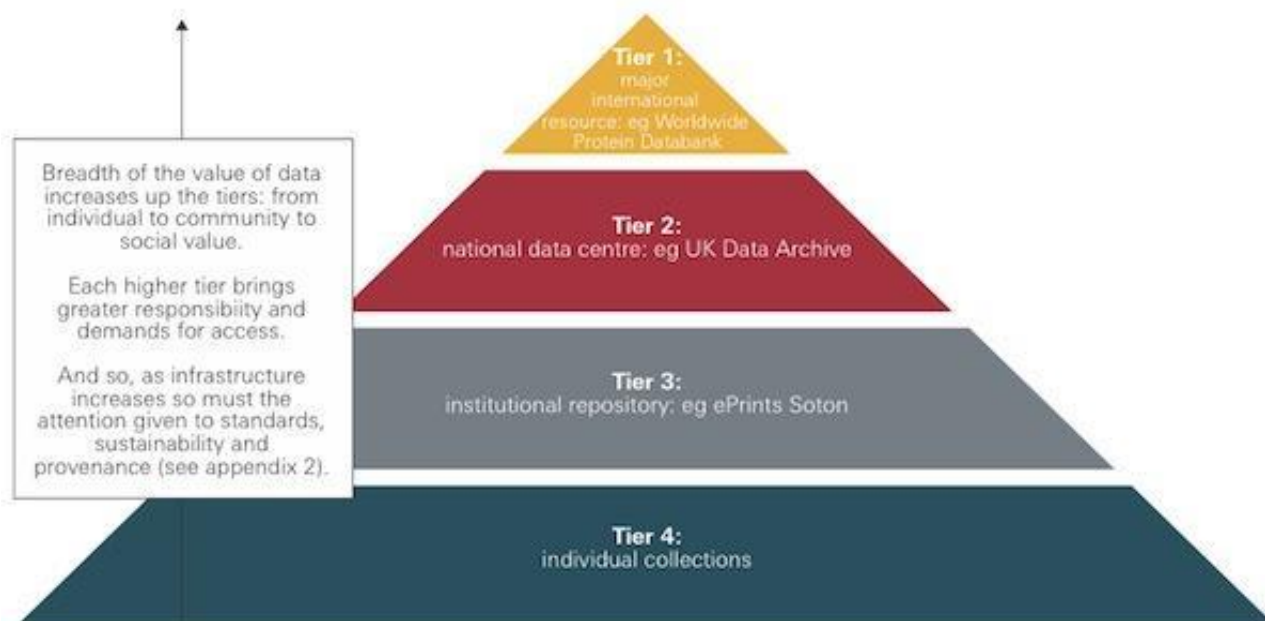
<b>Tegur</b>	<b>Kirjeldus</b>
Eesmärk	Tagada teadusandmete rahvusvaheline kasutatavus nii teadusprojekti jooksul kui selle järgselt
Andmete fookus	Käimasolevate teadusprojektide raames tekkivad andmekogud
Teenuste fookus	Andmehaldusplaanide teostamine ja andmete haldamine uuringu käigus Publikatsioonide sidumine algandmetega Andmete hoiustamine pilvelahendustes Rahvusvahelise andmete nähtavuse ja taaskasutuse toetamine
Kompetentside fookus	Teadusandmete haldamine pilvelahendustes
Tehnoloogia vajadused	Juurdepääs teadusandmete haldamise keskkondadele
Regulatsioonide fookus	Võimalus kasutada rahvusvahelisi teenusepakkujaid teadusandmete haldamiseks, hoidmiseks ja avaldamiseks
<b>Teenus</b>	<b>Kirjeldus</b>
Teadusprojekti andmete haldamine uuringute käigus	Võimalik tugineda ETAISi, Figshare, Zenodo, Mendeley, Dryadi, OpenAire, vmt. teenusepakkuja lahendustel ja taristul
Publikatsiooni allikandmete publitseerimine	Võimalik tugineda kirjastaja andmehoidla taristule
Andmete arhiveerimine	Võimalik kasutada rahvusvaheliste teenusepakkujate (nt. Figshare, Zenodo, Mendeley, Dryadi, OpenAire, vmt.) teenuseid
Andmete säilitamine	Võimalik kasutada rahvusvaheliste teenusepakkujate (nt. Figshare, Zenodo, Mendeley, Dryadi, OpenAire, vmt.) teenuseid

Andmete avaldamine	Võimalik kasutada rahvusvaheliste teenusepakkujate (nt. Figshare, Zenodo, Mendeley, Dryadi, OpenAire, vmt.) teenuseid
Aruandlus	Võimalik kasutada mõnede rahvusvaheliste teenusepakkujate (nt. OpenAire jt.) teenuseid
<b>Osapool</b>	<b>Roll</b>
Teadlane	Tagab andmete dokumenteerituse ja vastavuse rahvusvaheliste teenusepakkujate nõuetele
Teadusprojekt	Tagab andmete dokumenteerituse ja haldamise vastavalt rahvusvaheliste teenusepakkujate nõuetele ja andmehaldusplaanile
Teenusepakkujad	Andmete turvaline hoid ja töötlemine vastavalt oma teenustele
Teadusasutus või ülikool	Teadlase tunnustamine teadusandmete arhiveerimise ja taaskasutamise eest Andmehaldusteenuste kasutamise juhendamine
ETAG	Andmehaldusplaanide ja andmete arhiveerimise nõude kehtestamine ja jälgimine Teenusepakkujate teenuste kasutamise reguleerimine (vajadusel)
<b>Kulumudel</b>	<b>Kirjeldus</b>
Tasu kasutatud kettamahu ja teenuste eest	Teadusprojektile (või teadusasutusele) tuleb vajadusel tasuda täiendava kettaruumi ja lisateenuste eest, mida teenusepakkujalt andmete haldamiseks ja arhiveerimiseks soetatakse

#### 4.3 Visioon avatud teaduse taristu teenusest

Kujundada välja uus avatud teaduse taristu teenus, mis teadlase jaoks oleks korraga nii lokaalne kui globaalne, tagaks kõik andmete vajadused kirjelduse ja pikaajalise säilivuse osas, pakuks erinevatele kasutajarühmadele just neile sobivaid analüüsivahendeid ning ei nõuaks seejuures olulisi lisakulutusi, on tõsine väljakutse. Uut teenust luues tuleb arvestada sellega, et juba väljakujunenud praktika järgi talletatakse osa teadusandmeid ka edaspidi rahvusvaheliste teenusepakkujate juurde. Paljutsiteeritud [teadusandmete](#) klasside hierarhia, mille Royal Society pakkus välja juba mõnda aega tagasi, kehtib endiselt ka praegu (vt. Joonis 19). Riigi tasandil andmete arhiveerimiseks taristu ja teenuste loomisel tuleb otsustada, kas fookus saab olema praegu käimasolevate ja tulevaste teadusprojektide andmete kogumisel ja

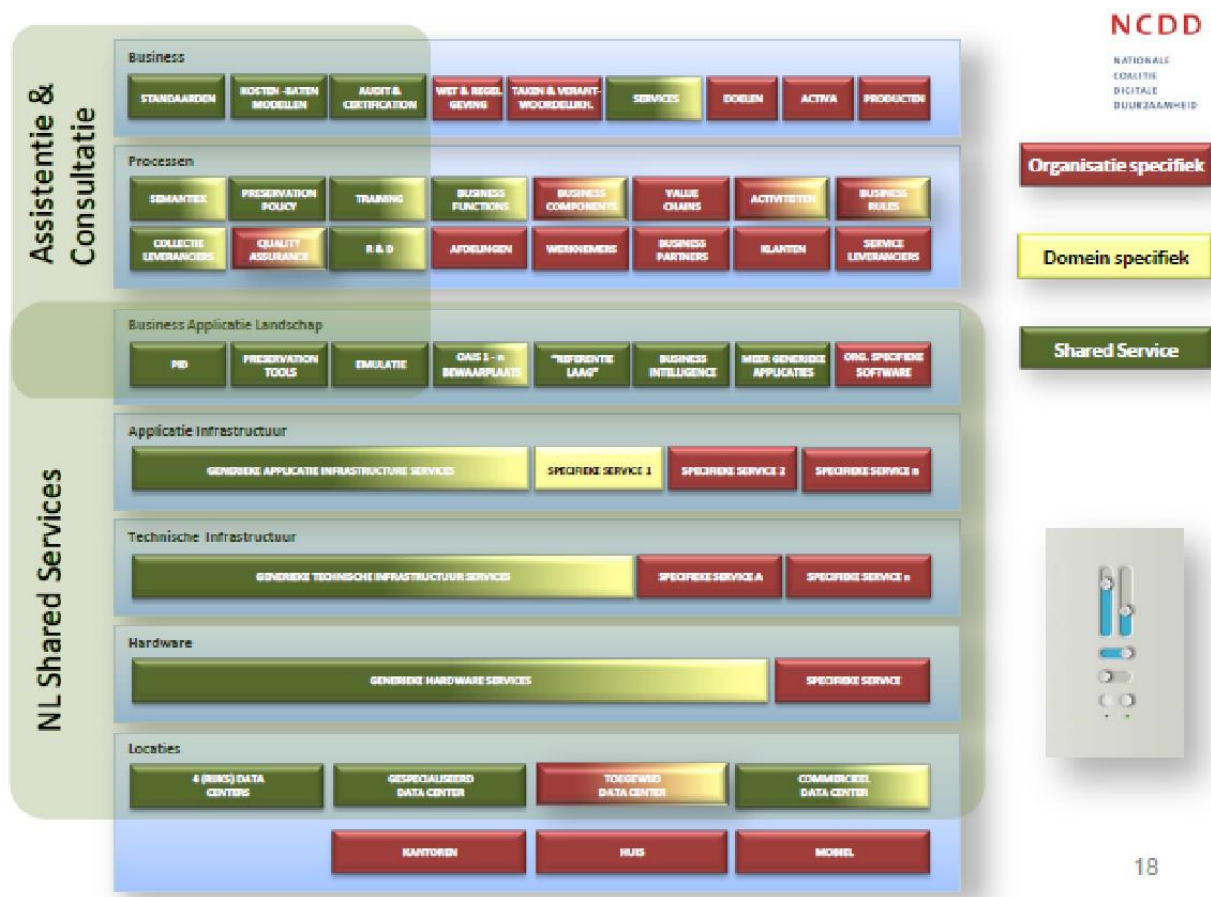
avaldamisel, või tuleks esmalt keskenduda teadlaste ja teadusasutuste valduses olevate, lõppenud ja ajalooliste uuringute andmete kokkukogumisele. Need ei ole päriselt üksteist välistavad stsenaariumid, kuid ajalooliste andmete arhiveerimisega saaks tulemuslikumalt hakkama ülikoolide digihoidlatel tuginev teenus.



**Joonis 19.** Teadusandmete püramiid - kasvava väärtuse ja pikaajalisuse püramiid (The Royal Society, 2012, lk. 60).

Riigid, kus hiljuti on digitaalseid arhiive ja säilitamise teenuseid analüüsitud, on arvutuste tulemusel jõudnud tõdemuseni, et uute organisatsioonide loomise asemel on kuluefektiivsem rajada teenused olemasolevate kompetentside ja teenuste najale hajutatud mudelina. Hollandis on digitaalse säilitamise koalitsioon läbi viinud rea uuringuid ja välja pakkunud visiooni, kuidas riiklikul tasandil koordineerida digitaalse säilitamise teenuseid nii, et ühegi vajaliku valdkonna andmed ei jääks teenusest ilma või oleks hävimisohus. Kuigi analüüsides on lähtutud mäluasutuste teenustest, on ka teadusandmeid vaadatud selle visiooni ühe võimaliku osana ning teadusasutusi teenusepakkujana (vt. Joonis 21).



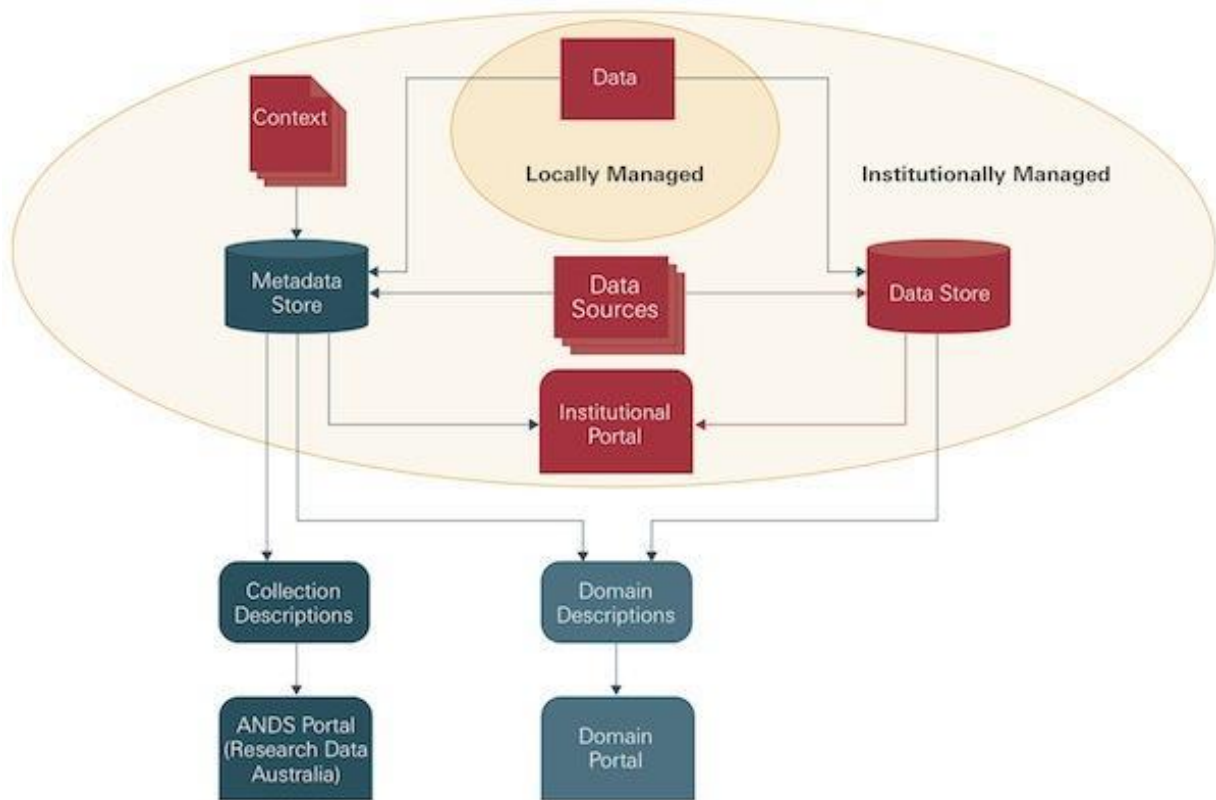


18

**Joonis 20.** Hollandi rahvuslik digitaalse säilitamise taristu (NCDD, 2015, lk.12).

Sarnasel olemasolevate teenuste agregeerimise mudelil tuginevad ka Euroopa Liidu initsiatiivid Eudat ja Open Science Cloud, kus keskselt panustatakse teenuste koordineerimise ja orkestreerimise kihi loomisesse, mõnevõrra hädavajalikku IT taristu koosvõimesse ning töötatakse välja regulatsioonid ja poliitikad teenuste pakkumiseks. Austraalia riikliku teadusandmete arhiivi ANDS<sup>71</sup> asutamisele eelnes pikk jada analüüse, mille tulemusena loodi teenus, mis samuti kombineerib kohaliku, valdkondliku ja riikliku tasandi (Joonis 16).

<sup>71</sup> <http://www.ands.org.au/about-us/what-we-do>



**Joonis 21.** Australian Research Data Commons (2012).

Selleks, et Eestis oleks võimalik sarnane agregeeritud teadusandmete haldamise teenuste pakett kirjeldada, on täiendavalt vaja läbi viia uuring meil leiduvatest ja tekkivatest teadusandmete mahtudest, mida oleks vaja kohalikul tasandil arhiveerida. Uuring peaks andma ülevaate ka andmete iseloomust distsipliinide kaupa ning ootustest nende kasutamisele. Sarnased uuringud on läbi viidud näiteks Soomes,<sup>72</sup> Ühendkuningriigis, Lõuna-Aafrikas,<sup>73</sup> jm. Teiste riikide kogemus näitab samuti üheselt, et teenuse väljaarendamine peab olema täiendatud toetavate tegevustega, et tekiks kultuur, kus andmeid tahetakse arhiveerida ning teadlased on motiveeritud seda tegema.

<sup>72</sup> [http://www.fsd.uta.fi/fi/julkaisut/julkaisusarja/FSDjs07\\_OECD\\_en.pdf](http://www.fsd.uta.fi/fi/julkaisut/julkaisusarja/FSDjs07_OECD_en.pdf)

<sup>73</sup> <http://dx.doi.org/10.17159/sajs.2016/20150316>

## 5 Teekaart

Lähteülesanne: sihtgruppide rollide ja positsiooni kaardistamine; koostatakse enim tõenäoline institutsionaalne lahend. Teekaardi peatükis kirjutatakse põhjalikumalt lahti üks eelnevalt kirjeldatud stsenaariume. Ühe sisendina kasutame ka 4. detsembri arutelu tulemust.

Kes	Mida	Millal				
		2018	2019	2020	2021	2022
Teadlane	Viib ennast kurssi teaduse avaandmete põhimõtetega ning rakendab neid uuringu planeerimisel	●				
Teadusprojekt/ teadusgrupp	Annab sisendit institutsionaalse teadusandmete poliitika kujundamiseks ning planeerib teadusgrupis tööjaotuse lähtuvalt teaduse avaandmete nõuetele (nt andmete haldamise eest vastutav isik on uurimisgrupi liige); viib ennast kurssi valdkondlike andmearhiividega	●				
Ülikool/TA asutus	Kehtestada organisatsiooni (avatud) teadusandmete poliitika või kord ning pakkuda selle rakendamise juhendamist, tugistruktuur teadusandmete haldamiseks	●	●			
Teadus- raamatukogu	Koolitused, juhendmaterjalid ja konsultatsiooniteenus on välja arendatud; institutsionaalsed (või institutsioonide vahelised) repositooriumid töötavad	●	●			
Teadusvälja- annete kirjastaja	Liitub deklaratsiooniga <i>Joint Declaration of Data Citation Principles</i>	●	●	●		
Rahastajad (sh ETAG, KIK, Archimedes, Innove)	Integreerib veelgi tugevamalt andmehaldusplaani projektitaotluse osana. Lisab avaldatud teadusandmed teadlase tegevuse hindamise mudelisse	●	●	●		
Ministeeriu-mid (sh HTM, MKM)	Lisab avaldatud teadusandmed teadusasutuste tegevuse hindamise mudelisse (sh baasfinantseerimine ja evalveerimine)		●	●	●	
	Loob keskkonna andmete (nii teadusandmete kui ka riiklikult kogutava statistika) säilitamiseks, mis võimaldab avatust ja taaskasutust erinevate sihtrühmade poolt		●	●	●	●

## Soovitused

1. Sõnastada avatud teaduse positiivsed mõjud eri ühiskonnatasanditel, koos näidetega, ning kasutada neid teadlikkuse tõstmisel ning avatud teaduse suunal liikumise vajalikkuse põhjendamisel;
2. Kasutada teadlaste hoiakute ja käitumise suunamiseks sama valdkonna esindajaid, kes on avatud teaduse eeskõnelejad;
3. Arvestada avatud teaduse alaste nõuete kehtestamisel valdkondlike erinevustega;
4. Toetada avatud teaduse alaste oskuste suurenemist kõigis seotud institutsioonides (nt andmespetsialiste koolitades, lisades teema doktorantide õppekavasse);
5. Pidada osapoolte vahel pidevat sisulist arutelu avatud teaduse eesmärkide ja vahendite osas (näiteks andes teadlastele sisulist tagasisidet andmehaldusplaanide kohta);
6. Pakkuda ülikoolidele motivatsiooni institutsionaalsete avatud teaduse praktikate paremaks juurutamiseks;
7. Käivitada pilootprojekt sotsiaalteaduse andmete mahtude ja nendega seotud kulude kaardistamiseks ning kasutada saadud tulemusi kogu Eesti teadusandmete tasuvusanalüüsi teostamiseks;
8. Tõstmaks kogukonna teadlikkust, tuleks Eestis vastu võtta Ühendkuningriikidega sarnane memorandum avatud teadusandmete haldamise sisuliste põhimõtetega (*Concordat on Open Research Data*, 2016);
9. Seada FAIR põhimõtteid järgivate andmete arhiveerimine soovitavaks kvaliteedi tasemeks riiklikku teadusrahastust saavatele uurimisprojektidele;
10. Riiklikul tasandil on otstarbekas reguleerida nõue, et andmehaldusplaanis näidatakse ära kriteeriumid, mille alusel projekti meeskond planeerib oma andmed arhiveerida. Andmehaldusplaani koostajatele tuleb lisada vastav juhendmaterjal;
11. Andmete taaskasutamise praktika toetamiseks käivitada rahastusprogramm, mis toetab ainult avatud teadusandmete põhjal teadustöö tegemist;
12. Kehtestada standardid ja metoodika avatud andmekogude loomiseks ja tunnustamiseks (näiteks arvestades neid nii teadlase kui ka institutsiooni tasemel ühe teadustöö mõõdikuna).

## Viidatud kirjandus

1. Beagrie, N., Houghton, J., Palaiologk, A., Williams, P. (2012). Economic Impact Evaluation of the Economic and Social Data Service. JISC.  
<http://www.esrc.ac.uk/files/research/evaluation-and-impact/economic-impact-evaluation-of-the-economic-and-social-data-service/>
2. Beagrie, N., Houghton, J. (2014). The Value and Impact of Data Sharing and Curation: A synthesis of three recent studies of UK research data centres. JISC.  
<http://repository.jisc.ac.uk/5568/1/iDF308> -  
[Digital Infrastructure Directions Report%2C Jan14 v1-04.pdf](http://repository.jisc.ac.uk/5568/1/iDF308-Digital%20Infrastructure%20Directions%20Report%20Jan14%20v1-04.pdf)
3. Bolikowski, L (2014). Situation and Plans on Fostering Open Science in Poland.  
<https://avointiede.fi/documents/10864/18388/Bolikowski/d8f52729-8403-4b6d-9970-b9dc406ac9af>
4. Business Models for Sustainable Research Data Repositories (2017). OECD Science, Technology and Innovation Policy Papers. December 2017 No. 47. [http://www.oecd-ilibrary.org/science-and-technology/business-models-for-sustainable-research-data-repositories\\_302b12bb-en](http://www.oecd-ilibrary.org/science-and-technology/business-models-for-sustainable-research-data-repositories_302b12bb-en)
5. CESSDA SaW. (2017). Report summarising the current state of needs of new data services. CESSDA Strengthening and Widening, deliverable 4.7.  
<http://cessdasaw.eu/content/uploads/2017/11/SAW-D4.7.pdf>
6. COAR (24.01.2017) Confederation of Open Access Repositories Survey of Research Data
7. Cost Concept Model and Gateway Specification. (2014). 4C – Collaboration to Clarify the Cost of Digital Curation, JISC. <http://www.4cproject.eu/d3-2>
8. DCC How to Write a DMP guide: <http://www.dcc.ac.uk/resources/how-guides/develop-data-plan>
9. DCC How to Select What Data to Keep guide: <http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep>
10. DCC How to Licence Research Data guide: <http://www.dcc.ac.uk/resources/how-guides/license-research-data>
11. Davies, T (2014) [Open Data in Developing Countries - Emerging Insights from Phase I Report](http://www.dcc.ac.uk/resources/how-guides/develop-data-plan)
12. EC's Agenda on Open Science: <https://ec.europa.eu/digital-agenda/en/open-science>  
<https://ec.europa.eu/digital-agenda/en/open-science>
13. DMPonline tool: <https://dmponline.dcc.ac.uk/>
14. Ember, C., Hanisch, R. (2013). Sustaining Domain Repositories for Digital Data: A White Paper. Output of the workshop, "Sustaining Domain Repositories for Digital Data," Ann Arbor, MI, June 24-25, 2013. Ann Arbor, MI: Inter-University Consortium for Political and Social Research, <http://dx.doi.org/10.3886/Sustaining-Domain-Repositories-Digital-Data>
15. Evaluation of Cost Models and Needs & Gaps Analysis. (2014). 4C – Collaboration to Clarify the Cost of Digital Curation, JISC. <http://www.4cproject.eu/d3-1>
16. G7 Science Ministers Communiqué (2017). Open Science. Turin, Italy, September 28, 2017.  
<http://www.g8.utoronto.ca/science/2017-science-communique.html>
17. Houghton, J. (2011). Costs and Benefits of Data Provision. Report to the Australian National Data Service. ANDS.

- [http://www.ands.org.au/data/assets/pdf\\_file/0004/394285/houghton-cost-benefit-study.pdf](http://www.ands.org.au/data/assets/pdf_file/0004/394285/houghton-cost-benefit-study.pdf)
18. Houghton, J.; Gruen, N. (2014). Open Research Data. Report to the Australian National Data Service. ANDS. [http://www.ands.org.au/data/assets/pdf\\_file/0019/393022/open-research-data-report.pdf](http://www.ands.org.au/data/assets/pdf_file/0019/393022/open-research-data-report.pdf)
  19. Kitchin, R. et al. (2015), Funding models for Open Access Repositories, Maynooth: Maynooth University. Dublin: the Royal Irish Academy and Trinity College Dublin. <http://dx.doi.org/10.3318/DRI.2015.4>
  20. Lord, P., Macdonald, A., Lyon, L., Giaretta, D. (2004) From Data Deluge to Data Curation. Proceedings of the UK e-Science All Hands Meeting 2004
  21. Lynch, C. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. Association of Research Libraries Report No. 226. <http://old.arl.org/resources/pubs/br/br226/br226ir.shtml>
  22. Mossink, V., Bijsterbosch, M. and Nortier, J. (2013) European Landscape Study of Research Data Management. For SIM4RDM- Support Infrastructure Models for Research Data Management. [\[https://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM%20landscape%20report%20vs1%204\\_14.08.13.pdf\]](https://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM%20landscape%20report%20vs1%204_14.08.13.pdf) (14.12.2016)
  23. Nationale Coalitie Digitale Duurzaamheid (NCDD). (2015). Together where possible, individually where necessary. Constructing a network of nationwide facilities together. [http://ncdd.nl/site/wp-content/uploads/2014/06/summary\\_NCDD\\_research\\_DEFWEB.pdf](http://ncdd.nl/site/wp-content/uploads/2014/06/summary_NCDD_research_DEFWEB.pdf)
  24. Open Research Data Taskforce. (2017). Research Data Infrastructures in the UK: Landscape report. <http://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Documents/ORDTF%20report%20nr%201%20final%2030%2006%202017.pdf>
  25. Palaiologk, A.S.; Economides, A.A.; Tjalsma, H.D.; Sesink, L.B. (2012). An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS. International Journal of Digital Libraries, 12, 195-214. [dx.doi.org/10.1007/s00799-012-0092-1](http://dx.doi.org/10.1007/s00799-012-0092-1)
  26. Szprot, J. (2015) Open Science in Poland 2014. A Diagnosis.
  27. Tenopir, C., Birch, B. and Allard, S. (2012) Academic Libraries and Research Data Services: Current Practices and Plans for the Future: An ACRL White Paper. [\[http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir\\_Birch\\_Allard.pdf\]](http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf) (14.12.2016)
  28. Tenopir, C.jt. (2015) Research Data Services in Academic Libraries: Data Intensive Roles for the Future? Journal of eScience Librarianship, 4, 2. [\[http://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1085&context=jeslib\]](http://escholarship.umassmed.edu/cgi/viewcontent.cgi?article=1085&context=jeslib) (14.12.2016)
  29. Tenopir, C jt. (2016). Research Data Services in European Academic Research Libraries. [\[http://libereurope.eu/wp-content/uploads/2016/10/LIBER\\_RDM\\_SURVEY\\_FINAL\\_MANUSCRIPT.pdf\]](http://libereurope.eu/wp-content/uploads/2016/10/LIBER_RDM_SURVEY_FINAL_MANUSCRIPT.pdf) (14.12.2016)

30. The Research Information Network (RIN); JISC. (2011). Data centres: their use, value and impact. A Research Information Network report.  
[http://www.rin.ac.uk/system/files/attachments/Data\\_Centres\\_Report.pdf](http://www.rin.ac.uk/system/files/attachments/Data_Centres_Report.pdf)
31. The Royal Society. (2012). Science as an open enterprise. The Royal Society Science Policy Centre report 02/12. <https://royalsociety.org/~media/policy/projects/sape/2012-06-20-saoe.pdf>
32. Tjalsma, H., Rombouts, J. (2011). Selection of Research Data. Guidelines for appraising and selecting research data. DANS Studies in Digital Archiving 6.  
<https://dans.knaw.nl/nl/over/organisatie-beleid/publicaties/DANSselectionofresearchdata.pdf>
33. UK Data Service, (2015). Data management planning.  
<https://www.ukdataservice.ac.uk/manage-data/plan/costing>
34. Vines, T.H.; Andrew, R.L.; Bock, D.G.; Franklin, M.T.; Gilbert, K.J.; Kane, N.C.; et al. (2013). Mandated data archiving greatly improves access to research data. FASEB Journal. 27. 1304-1308. <http://dx.doi.org/10.1096/fj.12-218164>
35. Wageningen University & Research (s.d.) Data Management Planning.  
<https://www.wur.nl/en/Expertise-Services/Data-Management-Support-Hub/Browse-by-Subject/When-and-why-to-publish-your-dataset.htm>
36. Whyte, A.; Wilson, A. (2010). How to Appraise and Select Research Data for Curation. DCC How-to Guides. Edinburgh: Digital Curation Centre.  
<http://www.dcc.ac.uk/resources/how-guides/appraise-select-data>
37. Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18
38. Westerhof, A.; Pronk, T.E.; van der Kuil, A.; Mordant, A. (2017) Data Management Cost Guide.  
[https://www1.edugroepen.nl/sites/RDM\\_platform/Financieel1/Data%20Management%20Costs.aspx](https://www1.edugroepen.nl/sites/RDM_platform/Financieel1/Data%20Management%20Costs.aspx)  
<https://doaj.org/> (26.11.20)

