

Tartu ülikooli raamatukogu panus teadusandmete säilitamisse ja kättesaadavaks tegemisse

Liisi Lembinen

Tartu ülikooli raamatukogu direktori kohusetäitja

Teadlased loovad igapäev suurtest kogustes digitaalseid publitseerimata teaduse toorandmeid, mille säilitamise ja kättesaadavaks tegemisega pole siiani intensiivselt tegeletud. Enam ei piisa artikli kirjutamisest, vaid nüüd soovitakse juurdepääsu ka andmetele, mille põhjal uurimus on teostatud. Teaduse avaandmeteks (*open data*¹) nimetatakse teadustöö käigus kogutud andmed, mis tehakse internetis avatuna kättesaadavaks. Teadmineavatud andmete järjest kasvavast olulisusest on hakanud jõudma nii teadlaste kui ka Eesti riigini, samas on sellealane teavitustöö alles algjärgus. Sellele viitab ka Eesti Teadusagentuuri (ETAG) 2014. aastal läbiviidud uuring. Tulemused näitasid, et 43 protsenti teadlastest hoiavad jätkuvalt teaduse toorandmeid oma isiklikes arvutites ning kõigest 15 protsenti teadlastest on oma toorandmed teinud avalikult kättesaadavaks. Siiski selgus antud uuringust, et üle 80 protsenti teadlastest pigem pooldab oma teadusandmete avalikuks tegemist ja säilitamist, mis peaski olema motiveerivaks aluseks ülikoolidele, raamatukogudele ning riigile antud teemaga tegelemiseks.

ETAGi poolt 17.10.2014 korraldatud teadusandmete seminari² arutelu käigus jäi kahjuks ühe mõttena kõlma, et teadusandmete säilitamise ja avalikustamise temaatikaga tegelemiseks aega on ning vajadust kiirustada ei ole. Sama lähenemine võeti ka avatud juurdepääsu poliitika puhul. Sellest lähtuvalt ei olnud Eestis seni vastu võetud ühtset lähenemisi viisi, kuigi temaatika on juba aastaid vana. Väljaspool Eestit võetakse iseenesest mõistetavana, et asutused ja riik toetavad avatud juurdepääsu poliitikat, ka ei arutata enam teemal, kas avatud juurdepääsu poliitikat üldse on vaja. Samal ajal kui Eesti ei ole veel jõudnud avatud juurdepääsu poliitikat toetavate riikide hulka, on välismaailm jõudnud juba järgmise etappi – tähelepanu keskmes on nüüd juba teaduse toorandmete avaldamine, säilitamine ning andmehalduspoliitika laiemalt ehk arutletakse selle üle, kuidas teaduse toorandmeid pikaajaliselt säilitada ning milline saab selles osas olema eri maade riiklik ja institutsionaalne poliitika.

Andmestiku jagamine ja taaskasutamine on teaduslikus uurimistöös järjest olulisem, seepärast on teadusraamatukogude poolt oluline suutapakkuda andmete leidmiseks, kättesaamiseks ja kasutamiseks efektiivseid vahendeid. Euroopa ja Ameerika Ühendriikide teadusraamatukogud on teaduse toorandmetega tegelenud juba mitu aastat ning selge on, et teadusandmete kättesaadavus on hakanud mängima järjest suuremat rolli eelkõige teaduses endas. Enam ei küsita mitte seda, kas teaduse toorandmeid säilitatakse ning tehakse avalikuks, vaid kus ja kuidas neid säilitatakse ning avaldatakse. Euroopa ja Ameerika Ühendriikide teadusraamatukogud on selle teema propageerimisel võtnud juhtiva rolli, mille kõrval on ka

¹ *Open Data*, Tartu Ülikooli Raamatukogu. http://www.utlib.ee/index.php?e_id=411&e=1

² ETAGi seminar "Teadusandmed: kellele ja milleks?" 17. oktoober 2014. <http://www.etag.ee/analuus/seminar-teadusandmed-kellele-ja-milleks/>

enamik teadusraamatukogusid ühendavaid organisatsioone on väljendanud oma seisukohta teadusandmete haldamise, kättesaadavaks tegemise ning säilitamise osas.

Euroopa teadusraamatukogude Ühenduse LIBER häälekandja LIBER Quarterly³ avaldas sel aastal Aarhushi ülikooli raamatukogu artikli, mis kajastab taanlaste nägemust raamatukogu uuest rollist teaduse toorandmete haldamisel, selle tegevuse arengusuundadest ning tulevikuvisionist. Seoses teaduse toorandmete haldamise, taaskasutuse ning jagamisega seonduvate tegevuste muutumisega on uuenenud ka teaduse toorandmete kontseptsioon. Taanis ollakse seisukohal, et teaduse toorandmete säilitamine peaks olema reguleeritud keskselt riigi tasandil ning neid peaks säilitama ühtses rahvuslikus (või rahvusvahelises) repositooriumis, mitte kohalikes institutsionaalsetes repositooriumides. Mitmes Euroopa riigis on juba kesksed andmearhiivid loodud, näiteks üks võimsamaid asub Hollandis ning kannab nime DANS. Taani ülikoolid võtnud justkui äraootava seisukoha, eeldates ühelt poolt kesket koordineeritust, teisalt mõistes, et teadlaste ootused vajavad ülereguleerimist vältides täitmist. Ka on taanlased seisukohal, et raamatukogu on teaduse toorandmetega seotud tegevuste süsteemi oluliselt loogiline osa. Mitmed raamatukogud on uute teenuste pakkumiseks uuendanud oma oskuste pagasit ja koolitanud töötajaid.

Raamatukogude tegevused teaduse toorandmetega seotud tegevustes jagunevad üldjuhul kolmeks: 1) nõustamine ja koolitamine – andmete säilitamise ja kirjeldamise osas; 2) säilitamine – andmekogude säilitamine repositooriumides, metaandmete lisamine ning identifikaatorite väljastamine (*DOI*); 3) andmehaldusplaanid (*data management plans* ja *research data management e RDM*) – raamatukogud aitavad koostada teadlasgruppidel teadusandmete haldusplaan. Juba 3-4 aastat tagasi otsiti erinevaid võimalusi, kuidas raamatukogud saaksid panustada andmehaldusesse (*RDM*) ning leiti, et nendeks on nõustamine rahastuse leidmiseks, uurimisgruppidele kirjandusülevaadete tegemine või arengutest teavitamine, bibliomeetria, *impact factor*, institutsionaalse repositooriumi teenuste pakkumine, andmehaldusplaanide koostamine (nõustamine), avatud juurdepääsu propageerimine, andmeanalüüs, autoriõigusealane nõustamine, andmete arhiveerimise ja metaandmete alane nõustamine jne. Raamatukogudele on välja pakutud ka kümneastmelist mudelit partnerluses teadlastega andmete halduses: 1) rahastajate määratud andmehaldusnõuete kogumine (koostöös teaduskondadega); 2) andmehalduse planeerimine läbi toetuse, abistamise ja treeningute igal tasemel; 3) tehniline nõustamine metaandmete ning andmeformaatide osas; 4) andmete tsiteerimine; 5) koolitused (alustades doktorantidest); 6) andmete litsentsidega seotud küsimused; 7) teaduse toorandmete hindamisel osalemine (milliseid andmeid peaks või/ja võiks säilitada); 8) andmete säilitamine (repositooriumid); 9) andmete kättesaadavaks tegemine; 10) teadusandmete mõju (*impact*) hindamine koostöös teadus- ja arenduspersonaliga.⁴ Hea meel on tõdeda, et Tartu ülikooli raamatukogu (TÜR) pakub enamust eelpool mainitud teenustest.

³ Kruse, Filip; Thestrup, Jesper Boserup (2014). "Research libraries' new role in research data management, current trends and visions in Denmark." *LIBER Quarterly* 23 (4), <http://liber.library.uu.nl/>

⁴ Cox, Andrew M; Pinfield, Stephen (2013). "Research data management and libraries: Current activities and future priorities." *Journal of Librarianship and Information Science*, doi:10.1177/0961000613492542. <http://www.lis.sagepub.com>

Ka Horizon2020 on võtnud pilootprojektina uue suuna, kus teatud teadusvaldkondades on rangelt soovituslik lisada taotlustele plaanid, milles märgitakse ära teaduse toorandmete osas tehtavad tegevused.⁵ See on nõue, millest mitmed Euroopa raamatukogud on saanud innustust ning asunud töötama välja teenuseid andmehaldusplaanide koostamiseks.

Tartu Ülikooli raamatukogu tegevused teadusandmete suunal

Kuna TÜR on juba üle viie aasta tegelenud teavitustööga avatud juurdepääsust, on ta eelnevat arvesse võttes võtnud proaktiivse hoiaku tegeleda ka teadusandmetega. Teaduse toorandmete teema on leidnud käsitlemist kahel viimasel aastal peetud Avatud Juurdepääsu Nädalal oktoobris.

2014. aasta kevadel liitus Tartu Ülikool (eesotsas raamatukoguga) rahvusvahelise organisatsiooniga DataCite (www.datacite.org). DataCite on ülemaailmne raamatukogude, andmesäilituskeskuste ja teiste teaduslike uurimisasutuste võrgustik, mille tegevuse üheks eesmärgiks on teaduse avaandmete olulisuse osas teadlikkuse tõstmine. Avaandmed on teaduse seaduspärane osa ja oluline on muuta need identifikaatorite abil tsiteeritavaks, panustades teadustulemuste laialdasse levikusse. DataCite'i liikmetel on õigus väljastada teadusandmetele unikaalseid DOI numbreid (DOI – *Digital Object Identifier*). 2015. aasta seisuga on Tartu ÜlikoolEestis väljastanud 130 DOI numbrit keemikutele, loodusteadlastele ning keeletehnoloogidele.

Vajadus ja idee DataCite'ga liitumiseks ning DataCite Eesti loomiseks tekkis Tartu Ülikoolil juba 2012. aastal ning idee algatajateks olid loodusmuuseumi teadlased. Kogu DataCite Eesti loomise protsessi ongi TÜ raamatukogu jaoks oluliseks teinud teadlaste initsiatiiv, kuid seejuures mitte humanitaarteadlaste poolt, kes on üldjuhul alati raamatukogu suhtes positiivsemalt häälestatud, vaid just loodusteadlaste huvi raamatukoguga koostööd teha.

DataCite Eesti on nelja ülikooli (Tartu Ülikool, Tallinna Ülikool, Tallinna Tehnikaülikool ja Eesti Maaülikool) konsortsium, mille aastamaks tagab nende ülikoolide teadlaste teaduse toorandmetele piiramatu koguse DOI numbreid. Samuti on loodud DataCite Eesti veebiportaal⁶, mida sisaldab üldisemat teavet teadusandmete haldamise ning DataCite organisatsiooni kohta. DataCite Eesti projekti elluviimist koordineerib Tartu Ülikooli raamatukogu, kes pakub kohalikule teaduskogukonnale teadusandmete haldamisega seotud teenuseid ning toimib vahendaja klientorganisatsiooni ja DataCite'i laiema taristu vahel. Võimaldades uurimisasutustel oma andmestikele DOI-sid lisada, aitab raamatukogu kaasa Eesti teadusandmestiku rahvusvahelisele kasutajaskonnale nähtavaks ja kättesaadavaks tegemisele. DataCite Eesti konsortsiumi loomist rahastab Eesti Teadusagentuur.

Tartu Ülikooli raamatukogu avas ka eraldi teaduse toorandmete repositooriumi (<http://datadoi.ut.ee/>), toetamaks Tartu ülikooli teadlaste andmete säilitamist. Tartu ülikooli teadusandmete repositoorium on mõeldud eelkõige üksikteadlase toetuseks, kes ei soovi

⁵ *Guidelines on Data Management in Horizon 2020* (2013). Euroopa Komisjon, (11. detsember). http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

⁶ DataCite Eesti. www.datacite.ut.ee

säilitada teaduse toorandmeid oma isiklikus arvutis ning kes pole oma andmed arhiveerinud mõnes teises andmearhiivis. Kõik TÜ DataDoi repositooriumis paiknevad andmekogud saavad DOI numbrid ning on leitavad DataCite'i otsinguportaali kaudu. Esimene aasta teaduse toorandmetele juurdepääsu loomises on Eestis sujunud edukalt – konsortsium on loomisel ning esimesed DOI numbrid on väljastatud.

2014. aasta augustist töötab raamatukogus kaks uut ametinime kandvat ja tööülesandeid täitvat raamatukoguhoidjat – teadusandmete spetsialisti -, kes on end viimaste kuude jooksul viinud kurssi teadusandmete temaatikaga. Open Access nädala raames oli TÜ raamatukoguhoidjatel võimalus läbida teadusandmetealane koolitus, mida viis läbi just selle ala välisekspert Stuart MacDonald Edinburgi ülikooli raamatukogust, kellel on aastatepikkune kogemus andmetega tegelemisel. Edinburgi ülikooli raamatukogu on spetsiaalselt raamatukoguhoidjatele välja töötanud MANTRA programmi, mille abil omandada andmehaldusalaseid teadmisi ja oskusi.⁷ Tartu ülikooli raamatukogu teadusandmete spetsialisti ülesandeks on pakkuda teadlastele andmete haldamisel igakülgset tuge, alates metaandmetest, standarditest ja litsentsidest kuni andmehaldusplaanide koostamise ning teavitustöö tegemiseni. Eesmärgiks on jõuda teadlaseni enne, kui teadlane hakkab mõtlema oma teadusandmete haldamise ja säilitamise peale. Raamatukogu soovib olla teadlasele esmane ning efektiivselt toimiv partnerteaduse toorandmete käitlemisel.

Teadusandmete haldamine ja säilitamine on Eestis alles siiski algjärgus, kuid keegi peab astuma esimese sammu ning avama ukse. Tartu Ülikool raamatukogu juhtimisel on selle ukse avanud.

⁷ Research Data Management Training MANTRA. Edinburghi ülikool (2014). <http://datalib.edina.ac.uk/mantra/>