

Julkaisusyöttöpalvelu tietojen syöttämiseksi ja tallentamiseksi VIRT-julkaisutietopalveluun

Toteutussuunnitelma, luonnos 24.5.2016

CSC – Tieteen tietotekniikan keskus Oy (CSC) on valmistellut käsillä olevan toteutussuunnitelman julkaisusyöttöpalvelun toteuttamisesta. Suunniteltu palvelu mahdollistaisi julkaisujen tallentamisen VIRT-julkaisutietopalveluun sellaisille organisaatioille, joilla ei ole omaa tutkimustietojärjestelmää tai julkaisurekisteriä.

Sisältö

1	Mikä on VIRT-julkaisutietopalvelu?	2
1.1	Julkaisutietojen kerääminen ja raportointi tutkimusorganisaatioista	2
1.1.1	Yliopistot	2
1.1.2	Ammattikorkeakoulut	3
1.1.3	Tutkimuslaitokset	3
2	Julkaisusyöttöpalvelu	4
2.1	Palvelun toiminnallisuudet	4
2.1.1	Perustoiminnallisuudet	4
2.1.2	Lisätoiminnallisuudet	4
2.2	Palvelun hyödyt	4
2.2.1	Kustannushyödyt	4
2.2.2	Toiminnalliset hyödyt	5
2.3	Palvelun kustannukset	5
2.3.1	Toteutuksen kustannukset	5
2.3.2	Kustannusten jakautuminen	6
3	Eteneminen	6
	Liite 1. Ammattikorkeakoulujen julkaisumäärät 2012-14	7

1 Mikä on VIRTA-julkaisutietopalvelu?

Vuonna 2016 käyttöönotettu Korkeakoulujen valtakunnallisen tietovarannon VIRTA-julkaisutietopalvelu kokoaa yhteen tutkimusjulkaisujen tiedot kaikista suomalaisista tutkimusorganisaatioista. Palvelun avoimen lukurajapinnan avulla tietoja voidaan hyödyntää erilaisissa tutkijoille suunnatuissa palveluissa ja prosesseissa. Tiedot voidaan siirtää esimerkiksi tutkimusrahoittajien palveluihin, tutkimusaineisto-, tutkimusinfrastruktuuri- tai muihin tutkijoiden käyttämiin palveluihin, jolloin tutkijan ei tarvitse syöttää tietoja enää käsin.

Opetus- ja kulttuuriministeriön (OKM) vuosittainen julkaisutiedonkeruu toteutetaan vuoden 2016 tiedoista alkaen kokonaisuudessaan VIRTA-julkaisutietopalvelun kautta. Lisäksi korkeakoulut ja muut tutkimusorganisaatiot saavat VIRTA-julkaisutietopalvelusta ajantasaista vertailutietoa omasta ja muiden organisaatioiden julkaisutoiminnasta.

VIRTA-julkaisutietopalvelussa korkeakoulut, valtion tutkimuslaitokset sekä yliopistolliset sairaalat säilyttävät kopiota osasta oman tutkimustietojärjestelmänsä tai julkaisurekisterinsä tietosisältöä. Tutkimusorganisaatiot vastaavat omien julkaisutietojensa oikeellisuudesta. Lisäksi organisaatioiden toimittamia tietoja rikastetaan tietovarannossa tietyiltä osin. Tietovarannossa esimerkiksi tunnistetaan ja yhdistetään organisaatioiden väliset yhteisjulkaisut, määritellään julkaisuille [Julkaisufoorumi](#)-kanava sekä -tasoluokitus.

VIRTA-julkaisutietopalvelu on OKM:n omistama palvelu, jonka teknisestä toteutuksesta ja kehityksestä vastaa CSC. OKM vastaa VIRTA-julkaisutietopalvelun kustannuksista, mutta kukin organisaatio vastaa omien tietojensa keruusta, siihen liittyvistä järjestelmistä ja tiedonkeruun kustannuksista.

Ks. lisää: <https://confluence.csc.fi/display/VIR/VIRTA-Julkaisutietopalvelu>

1.1 Julkaisutietojen kerääminen ja raportointi tutkimusorganisaatioista

OKM on tähän asti kerännyt julkaisutiedot korkeakouluilta vuosittain. Vuoden 2016 julkaisutiedoista alkaen vuosittainen CSV-tiedostojen raportointi päättyy ja tiedonkeruussa siirrytään VIRTA-julkaisutietopalveluun tapahtuvaan jatkuvaan XML-tiedostojen siirtoon organisaatioiden omista lähdejärjestelmistä.

Palvelua pilotoitiin keväällä 2016 ja maaliskuussa 2017 kaikki julkaisutiedonkeruuseen osallistuvat organisaatiot raportoivat vuoden 2016 julkaisutietonsa kokonaisuudessaan VIRTA-julkaisutietopalvelun kautta.

1.1.1 Yliopistot

Useimmat yliopistot keräävät omat tietonsa CRIS-tutkimustietojärjestelmiin. Tiedonsiirto VIRTA-julkaisutietopalveluun voidaan toteuttaa siirtämällä tiedot automaattisesti joko suoraan näistä lähdejärjestelmistä tai tietovarastoista.

1.1.2 Ammattikorkeakoulut

Ammattikorkeakouluilla julkaisujen määrä on yliopistoja huomattavasti vähäisempi, joten varsinaisia tutkimustietojärjestelmiä ei juuri ole käytössä. Syksyllä 2013 tehdyn selvityksen mukaan yhdessä ammattikorkeakoulussa on käytössä tutkimustietojärjestelmä, ja muissa tiedot kerätään joko sähköisellä lomakkeella, kirjastojärjestelmään tai muuhun järjestelmään. Kukin ammattikorkeakoulu on järjestänyt tiedonkeruun parhaaksi katsomallaan tavalla. Pääsääntöisesti tutkijat syöttävät tiedot jollekin sähköiselle lomakkeelle, josta kirjasto tms. taho poimii tiedot esimerkiksi vuosittain excel-tiedostoon ja käy ne läpi manuaalisesti. Osassa korkeakouluja tiedot viedään tämän jälkeen esimerkiksi Voyager-kirjastojärjestelmään.

Tiedonkeruumenetelmä	Ammattikorkeakoulut (lkm)
SoleCRIS	1
Voyager	5
Muu (esim. HR-järjestelmä, projektinhallintajärjestelmä)	7
Ei järjestelmää	7
Yhteensä	20

Lähde: Tutkijan identifiointi -esiselvitys 2013: <https://confluence.csc.fi/display/tutki/Tutkijan+identifiointi+esiselvitys>

Ammattikorkeakoulujen OKM:lle raportointien julkaisujen määrä on ollut selvässä kasvussa. Vuonna 2012 julkaisuja raportoitiin ammattikorkeakouluista yhteensä 3 568, kun vastaava luku vuonna 2014 oli 5 450 (ks. Liite 1).

1.1.3 Tutkimuslaitokset

Julkaisutiedonkeruuta ollaan asteittain laajentamassa myös tutkimuslaitoksiin ja yliopistollisiin sairaaloihin. Näiden organisaatioiden vuoden 2015 tiedonkeruu toteutetaan kokonaisuudessaan VIRTJA-julkaisutietopalvelun kautta syyskuussa 2016. Sairaaloilla sekä useilla tutkimuslaitoksilla on julkaisutiedonkeruuseen tarkoitettuja tietojärjestelmät, mutta erityisesti pienissä laitoksissa tietoja kerätään osin manuaalisesti.

Useimmat tutkimuslaitokset ovat liittymässä julkaisutiedonkeruuseen vasta vuodesta 2015 alkaen, joten niiden osalta julkaisujen lukumäärätietoja ei ole vielä saatavilla.

Tiedonkeruumenetelmä	Tutkimuslaitokset (lkm)
Tutkimustietojärjestelmä	1
Julkaisurekisteri	8
Julkaisuarkisto	7
Muu (esim. Word, Excel, SharePoint, verkkosivut, kirjasto- tai projektinhallintajärjestelmä)	13
Yhteensä	19

Lähde: Valtion tutkimuslaitosten ja yliopistollisten sairaaloiden julkaisutiedonkeruu -esiselvitys 2014: <https://confluence.csc.fi/pages/viewpage.action?pageId=37818643>

2 Julkaisusyöttöpalvelu

Suunniteltu palvelu korvaisi nykyisen julkaisutiedonkeruun sellaisissa ammattikorkeakouluissa ja tutkimuslaitoksissa, joissa ei ole käytössä siihen tarkoitettua järjestelmää. Palveluun syötettävät julkaisut siirtyisivät automaattisesti VIRT-julkaisutietopalveluun, josta ne poimitaan vuosittain OKM:n tiedonkeruuseen ja josta niitä voidaan hyödyntää myös muihin valtakunnallisiin palveluihin. Suunniteltu palvelu tarjoaisi aikaisempaa paremmat mahdollisuudet tietojen reaaliaikaiseen tarkistukseen sekä näyttämiseen.

2.1 Palvelun toiminnallisuudet

2.1.1 Perustoiminnallisuudet

Ensimmäisessä vaiheessa toteutetaan palvelun perustoiminnallisuudet:

- Julkaisutietojen tallentaja (tutkija tai muu organisaation henkilöstöön kuuluva) tunnistautuu palveluun esimerkiksi Haka-tunnistautumisen avulla.
- Tallentaja syöttää julkaisun tiedot syöttölomakkeelle, joka sisältää kaikki OKM:n tiedonkeruun edellyttämät tiedot vaadittavien määritysten mukaisesti
- Syöttölomake on integroitu Julkaisuforumin julkaisukanavatietokantaan, jolloin lomakkeella valittavan julkaisukanavan (lehti, konferenssi tai kustantaja) tiedot saadaan automaattisesti, eikä niitä tarvitse syöttää käsin
- Organisaation pääkäyttäjä(t) saavat tiedon puutteellisista tiedoista sekä mahdollisista duplikaateista ja voivat palvelussa tarkistaa ja korjata tiedot
- Tarkistamisen jälkeen julkaisun tiedot siirtyvät automaattisesti VIRT-julkaisutietopalveluun
- Organisaation pääkäyttäjä(t) voivat määrittää, onko julkaisu VIRT-julkaisutietopalvelun rajapinnan kautta luettavissa ja voidaanko julkaisun tiedot näyttää JUULissa
- Organisaation pääkäyttäjä(t) voivat halutessaan poistaa julkaisun VIRT-julkaisutietopalvelusta

2.1.2 Lisätoiminnallisuudet

Palvelua kehitetään vaiheittain ja alla mainittuja lisätoiminnallisuuksia sekä muita uusia toiminnallisuuksia lisätään palveluun käyttäjäorganisaatioiden tarpeiden ja käytössä olevien resurssien mukaan:

- Syöttöpalvelu integroidaan organisaation omaan julkaisuarkistoon siten, että tallentaja voi julkaisun tietojen syöttämisen yhteydessä tallentaa sen kokotekstin julkaisuarkistoon (esim. DSpace)
- Tekijöiden henkilötietojen luetaan lomakkeelle organisaation HR-järjestelmästä
- Syöttölomake integroidaan ORCID-tunnistejärjestelmään, jolloin tutkijoiden tietoja ja tunnisteita voidaan hakea ORCIDista automaattisesti

2.2 Palvelun hyödyt

2.2.1 Kustannushyödyt

Julkaisusyöttöpalvelun käytöstä syntyy sitä käyttäville organisaatioille säästöjä, kun

- Organisaatiolla ei ole enää tarvetta ylläpitää omaa lomaketta, tietokantaa ja portaalia tietojen keräämiseen, raportointiin ja näyttämiseen

- Tiedonkeruuseen, tarkistuksiin ja tietojen käsittelyyn kuluva työmäärä vähenisi arviolta puoleen nykyisestä (nykyisin 1-2 htkk / vuosi)
- Organisaation ei tarvitse rakentaa prosessia tietojen siirtämiseksi VIRTA-julkaisutietopalveluun
- OKM:n muuttuvia tiedonkeruuvaatimuksia ei tarvitsisi päivittää kunkin organisaation omiin tiedonkeruujärjestelmiin

2.2.2 Toiminnalliset hyödyt

	Nykytilanne	Syöttöpalvelun mahdollistama tavoiteta
Tietojen näyttäminen	Tiedot JUULissa jopa 1,5 vuoden viiveellä	Tiedot reaaliaikaisesti näkyville JUULiin
Tiedon laatu ja yhteismitallisuus	Kullakin organisaatiolla oma lomakkeensa ja tarkistusprosessinsa	Yhteinen syöttölomake varmistaa, että tiedot yhteismitallisessa muodossa
Tietojen tarkistus	Tarkistetaan vuosittain käsin	Syöttölomake tarkistaa useimmat tiedot automaattisesti ja muut tiedot voidaan tarkistaa heti.
Duplikaattien hallinta	Tallentaja ei näe, onko julkaisu jo syötetty	Tieto duplikaatista tallentajalle automaattisesti syöttövaiheessa
Tietojen hyödyntäminen omissa järjestelmissä	Tiedot viedään vuosittain käsin kirjastojärjestelmään tms.	Tiedot avoimen rajapinnan avulla konekielellisesti luettavissa
Tiedonsyötön automatisointi	Julkaisukanavan, tekijöiden jne. tiedot syötetään käsin	Julkaisufoorumi-tietokannan ja Haka-tunnistautumisen integrointi mahdollistaa tietojen automaattisen tuonnin syöttölomakkeelle
Tiedonsiirto VIRTaan	Organisaatio tuottaa XML-skeeman mukaisen tiedoston omasta järjestelmästä ja siirtää sen VIRTaan	Tiedot automaattisesti suoraan VIRTaan

2.3 Palvelun kustannukset

2.3.1 Toteutuksen kustannukset

Ensimmäisen vaiheen (ensimmäinen sopimusvuosi) kustannukset ovat yhteensä noin **100 000 euroa**, mikä kattaa julkaisusyöttöpalvelun suunnittelun, perustoiminnallisuuden toteutuksen, tarvittavat integraatiot, testauksen, käyttöönoton, sopimushallinnan sekä käyttäjien tuen.

Jatkokehitysvaiheessa (ensimmäisen sopimusvuoden jälkeen) kustannukset ovat yhteensä noin **100 000 e / vuosi**, mihin sisältyvät tarvittavien lisätoiminnallisuuksien toteutus, mahdollisten tiedonkeruumuutosten päivittäminen, palvelun ylläpito, sopimushallinta sekä käyttäjien tuki.

2.3.2 Kustannusten jakautuminen

Suunniteltu palvelu on sitä käyttäville organisaatioille maksullinen. Organisaation vuosimaksu määräytyy julkaisujen määrän perusteella.

Alla on arvioitu palvelun vuosittaiset kustannukset ja alustavasti määritelty niiden jakautuminen organisaatioille, mikäli käyttäjäorganisaatioita olisi 19:

Julkaisujen määrä	Organisaatioiden lkm	Maksu euroa / vuosi
0-150	4	4 000
151-300	6	5 000
300-500	9	6 000
Yhteensä	19	100 000

3 Eteneminen

CSC päättää palvelun toteuttamisesta ammattikorkeakouluille ja tutkimuslaitoksille lähettämänsä kyselyyn 19.8.2016 mennessä tulleiden vastausten perusteella. Mikäli palvelun käyttöönottavia organisaatioita on riittävästi, palvelun toteutus käynnistetään syksyllä 2016 ja se on organisaatioiden käytössä keväällä 2017.

CSC on yhteydessä kyselyllä kiinnostuksensa ilmaiseisiin organisaatioihin neuvottelujen käynnistämiseksi, mikäli syöttöpalvelu päätetään toteuttaa. Neuvotteluissa tarkennetaan muun muassa vuosimaksun, sopimuskauteen ja palvelun käyttöehtoihin liittyvät asiat.

Mikäli palvelu toteutetaan, CSC järjestää sekä toteutuksen että jatkokehityksen aikana käyttöönottaville organisaatioille säännöllisiä tapaamisia, joissa organisaatiot voivat tuoda esille palvelun kehitykseen liittyviä tarpeita.

Liite 1. Ammattikorkeakoulujen julkaisumäärät 2012–14

Lähde: Vipunen (www.vipunen.fi)

Ammattikorkeakoulu / julkaisujen lkm	2012	2013	2014	Yhteensä	Muutos v:sta 2012 v:een 2014
Arcada	80	115	139	506	74%
Centria-ammattikorkeakoulu	79	115	104	417	32%
Diakonia-ammattikorkeakoulu	61	110	114	461	87%
Haaga-Helia ammattikorkeakoulu	118	204	218	735	85%
Humanistinen ammattikorkeakoulu	112	126	159	540	42%
Hämeen ammattikorkeakoulu	160	234	297	896	86%
Jyväskylän ammattikorkeakoulu	284	290	291	1 269	2%
Kajaanin ammattikorkeakoulu	42	40	61	230	45%
Karelia-ammattikorkeakoulu	162	160	200	758	23%
Kemi-Tornion ammattikorkeakoulu	111	159		415	
Kymenlaakson ammattikorkeakoulu	152	287	330	884	117%
Lahden ammattikorkeakoulu	144	181	163	613	13%
Lapin ammattikorkeakoulu			406	406	
Laurea-ammattikorkeakoulu	168	220	280	867	67%
Metropolia ammattikorkeakoulu	191	324	351	1 303	84%
Mikkelin ammattikorkeakoulu	203	310	393	1 311	94%
Oulun seudun ammattikorkeakoulu	193	161	176	799	-9%
Rovaniemen ammattikorkeakoulu	104	111		347	-100%
Saimaan ammattikorkeakoulu	105	119	148	563	41%
Satakunnan ammattikorkeakoulu	100	186	286	763	186%
Savonia-ammattikorkeakoulu	136	84	221	697	63%
Seinäjoen ammattikorkeakoulu	176	208	295	984	68%
Tampereen ammattikorkeakoulu	154	230	291	862	89%
Turun ammattikorkeakoulu	366	337	341	1 604	-7%
Vaasan ammattikorkeakoulu	25	16	33	160	32%
Yrkeshögskolan Novia	142	143	153	639	8%
Yhteensä	3 568	4 470	5 450	19 029	53 %