

How to create and fill institutional repositories
Workshop FinnOA:n seminaarissa Hankenilla 9.1.2004

MUISTIO

Workshopin vetäjät:

Ari Muhonen, Teknillisen korkeakoulun kirjasto, FinnOA-työryhmä
Leo Waaijers, SURF, Holland

Kansainväliset vieraat:

Jean-Claude Guedon, University of Montreal
Ingegerd Rabow, Lundin yliopiston kirjasto

Osallistujat:

Anneli Ahtola, Tampereen yliopiston kirjasto
Pekka Aroviita, SPR Veripalvelu
Sini Autio, Geologian tutkimuskeskus
Carita Bjön, Kela, Tutkimusosasto, Tietopalvelu
Hannele Fabritius, HY Kumpulan tiedekirjasto
Riitta Grahn, STAKES FinOHTA
Arja-Riitta Haarala, Tampereen teknillisen yliopiston kirjasto
Turid Hedlund, Svenska handelshögskolan
Timo Hellgren, VTT Tietopalvelu
Maria von Hertzen, Arcada
Jonas Holmström, Svenska handelshögskolan
Gunilla Häkli, HY Oikeustieteellisen tiedekunnan kirjasto
Ilkka Jäppinen, Puolustusvoimien teknillinen tutkimuslaitos
Jukka Kananen, Kuopion yliopiston kirjasto
Esa-Pekka Keskitalo, Helsingin yliopiston kirjasto
Sinikka Koskiala, Teknillisen korkeakoulun kirjasto
Kimmo Koskinen, HY Viikin tiedekirjasto
Ellen Kotanen, Stadia
Kari Mikkonen, Hämeen ammattikorkeakoulu
Ulla Ohvo, Etelä-Karjalan ammattikorkeakoulu
Päivi Pekkarinen, HY Terveystieteiden keskuskirjasto
Pirjo Rajakiili, HY Terveystieteiden keskuskirjasto
Niranjan Rajani, Svenska handelshögskolan
Annikki Roos, Kansanterveyslaitoksen kirjasto
Nina-Mari Salminen, Maa- ja elintarviketalouden tutkimuskeskus
Maria Schröder, Svenska handelshögskolans bibliotek
Liisa Sipola, Suomen Pankki, Siirtymätalouksien tutkimuslaitos
Jani Stenvall, Helsingin yliopiston kirjasto
Iris Tahvanainen, EVTEK
Tuula Tervashonka, Lapin yliopisto
Olli Vilanka, Svenska handelshögskolan
Rita Voigt, Teknillisen korkeakoulun kirjasto
Tiina Äärilä, Helsingin yliopisto, Hallintovirasto, Tietopalvelut

1. Kysely

Tilaisuuden aluksi Ari Muhonen kysyi kunkin osallistujan organisaation institutional repository (IR) -tilanteesta. Useimmilla organisaatioilla oli jo materiaalia (raportteja, väitöskirjoja ym.) vapaasti saatavilla verkossa, mutta vain yhdellä organisaatiolla (VTT) on OAI-yhteensopiva IR. TKK:n kirjasto saa omansa valmiiksi näinä aikoina.

Kysely myös paljasti vastaajien erilaiset käsitykset ja näkemykset siitä, mikä on IR. Siksi olisi tarpeen aloittaa yhteistyö Suomessa luomalla yhtenäisen käsitteistö. Jo institutional repository -termillä ei ole kunnollista suomenkielistä vastinetta. Tähän on tarjottu termiä ”julkaisuarkisto”, mutta se ei anna aivan oikeaa kuvaa IR:stä.

2. Tekniikka

Timo Hellgren VTT:ltä ja Rita Voigt Teknillisen korkeakoulun kirjastosta esittelivät lyhyesti, miten heidän organisaatioidensa IR:t on luotu. Molemmat ovat tehneet tarvittavan softan, Hellgren alusta asti käyttäen Perl-kieltä ja Voigt hyödyntäen Oldenburgin php-kielistä open source-skriptiä. Molemmissa tapauksissa varsinainen tietokanta käyttää Trip-ohjelmistoa.

Hellgren nosti esille kysymyksen, olisiko tarpeen luoda Suomeen oma OAI-yhteensopiva hakumoottori.

Kimmo Koskinen kertoi lyhyesti Viikin tiedekirjaston, Hankenin ja TKK:n kirjaston yhteisprojektista, jossa on tarkoitus tutkia Dspace-ohjelmaa sekä mahdollisesti muita vapaasti saatavilla olevia IR-softia. Projekti on vasta alussa ja se kestää koko vuoden 2004.

Koskinen kysyi Waaijersilta Hollannin ARNO-softasta ja sen saamisesta testattavaksi. Waaijers totesi, että se ei ole vielä vapaasti saatavilla, koska se on huonosti dokumentoitu. DARE-raportista on kyllä saatavilla tietoa ohjelmistosta. Raportti on saatavilla osoitteessa

www.surf.nl/en/publicaties/index2.php?oid=21

Lisää tietoa eri IR-ohjelmistoista on saatavilla muun muassa osoitteissa

www.openarchives.org/tools/tools.html

www.soros.org/openaccess/software/

Esa-Pekka Keskitalo totesi, että HYKin juuri hankkima DOMS-ohjelmisto on OA-yhteensopiva, joten myös sitä voidaan käyttää IR:n luomiseen. Keskustelussa sen huonona puolena pidettiin kuitenkin sitä, että kyseessä on kaupallinen ohjelmisto. Lisäksi DOMS on vasta hankittu, joten nopeille etenijöille sen käyttöönotto tapahtuu liian myöhään.

3. Sisältö

Organisaation omat opinnäytteet ja raportit ovat luonnollista materiaalia IR:ssä. Aineiston pitäisi kuitenkin olla kansainvälisesti kiinnostavaa. Keskustelussa nostettiin myös esiin kaiken organisaation henkilökunnan julkaiseman aineiston sijoittaminen IR:ään ainakin metadatatason.

Näin IR kokoaa tiedon kaikesta organisaation julkaisutoiminnasta yhteen paikkaan, jolloin se toimii organisaation näyteikkunana ulospäin.

Monella suomalaisella yliopistolla on jo oma julkaisurekisterinsä. Niistä voisi tehdä OAI-yhteensopivia. Leo Waaijers kertoi esimerkkinä Wageningenin yliopiston IR:stä, jossa on noin 40 000 tietuetta. Näistä 6000 - 7000 tietuetta sisältää myös kokotekstin. Tämä IR ei siis sisällä pelkkiä kokotekstejä, vaan myös pelkkiä metadatatietueita.

Keskustelussa mietittiin, mikä versio dokumentista pitää panna IR:ään. Usein lehtiartikkeleita editoidaan vielä julkaisuvaiheessa, ainakin typografia voi muuttua kirjoittajan versiosta.

Kuka sitten omistaa elektroniset dokumentit? Mihin ne sijoitetaan? Jos ne annetaan kirjaston hoteisiin, niitä ei ainakaan muutettaisi. Hyvä puoli voisi olla myös se, että dokumenttien linkkiosoitteet voitaisiin pitää pysyvinä. Näin dokumentit olisivat käytettävissä myös pitemmän ajan kuluttua niiden luomisesta. Tämä kaikki kuitenkin toisi lisää työtä kirjastolle. Samoin metadatan syöttö jäisi kirjaston kontolle. Kirjoittajille sitä työtä ei voisi antaa, he voisivat olla vastahakoisia sen tekemiseen, ja lopputulos olisi hyvin hajanainen.

4. Tekijänoikeudet

Tekijänoikeudet asettavat rajoituksia siihen, mitä aineistoa IR:ään voidaan sijoittaa. Ulkomaiset vieraamme totesivatkin, että parasta olisi, jos kirjoittajat jo etukäteen varmistaisivat tekijänoikeuksien säilymisen itsellään. Lundissa on tekeillä mallisopimus kirjoittajan ja kustantajan välille, jossa tämä on otettu huomioon. Guedonin mukaan myös Sparcilla on mallisopimuksia. Rabow toi esille mahdollisuuden ongelmiin, jos artikkelilla on useampi kirjoittaja. Mitä tapahtuu jos kaikki kirjoittajat eivät säilytäkään tekijänoikeutta artikkeliin itsellään? Ja jos kirjoittajia on useita, lupa artikkelin käyttöön pitää saada kaikilta heiltä. Tämä aiheuttaa omat ongelmansa artikkelin käytölle.

Monet väittelijät tekevät nykyisin niin sanottuja nippuväitöskirjoja. Näistä yliopisto voi väittelijän kanssa tehtävän sopimuksen perusteella julkaista vapaasti yhteenvedon. Sen sijaan erillisten artikkelien asettaminen verkkoon vaatii luvan tekijänoikeuden haltijalta, joka yleensä on kustantaja. Teknisessä korkeakoulussa luvat pyydetään ja lähes aina ne saadaankin. Näin he voivat asettaa väitöskirjat kokonaisina verkossa käytettäväksi. Myös Waaijers totesi, että heillä ei ole koskaan ollut ongelmia kustantajien kanssa lupien saamiseksi elektronisiin eripapainoksiin. Sen sijaan kirjoittajat ovat joskus olleet nihkeitä. Waaijers tosin totesi myös, että kysyttäessä Elsevier ensin antoi luvan artikkelien käyttöön, mutta veti sen sitten vähän ajan kuluttua pois.

5. Kirjaston rooli

Kirjaston roolia open access -toiminnassa mietittiin kahdelta kannalta. Ensin kysyttiin, onko IR kirjaston tehtävä. Tähän vastattiin nopeasti vastakysymyksellä, mikä muu yksikkö sen sitten tekisi. Atk-keskuksella ei ole siihen tarvittavaa tietämystä, vaikka he teknisen puolen osaisivatkin hoitaa.

Toisekseen kysyttiin, mitä virkaa kirjastolla yleensäkin on open access -aikakaudella. Open access ei vaadi kokoelmapolitiikkaa eikä aineiston välittäjiä. Yleinen toteamus oli, että kirjastolla on ja pysyy monia tehtäviä. Kirjaston rooli informaatiolukutaidon opettajana korostuu verkkotiedon määrän kasvaessa, samoin kirjastoa tarvitaan edelleen tiedon järjestäjänä. Ja onhan kirjastoilla

edelleen historialliset kokoelmansa. Kirjasto säilyy tieto-organisaationa (knowledge organization role). Guedonin mukaan kustantajat voivat tehdä kirjastot tarpeettomiksi, mutta open access tarvitsee niitä.

Leo Waaijers esitti mielenkiintoisen pohdinnan kirjaston asemasta tiedon tuotantoketjussa. Ohessa ovat Waaijersin aiheeseen liittyvät kalvot (kenniscyclus.ppt).

6. Yhteistyö

Workshopin viime metreillä muutaman osallistujan jo lähdettyä mietittiin, millaista yhteistyötä Suomessa IR:n parissa voitaisiin tehdä. FinnOA-työryhmä on perustamassa vuodelle 2004 alatyöryhmiä, joista yhden alana on IR. Ari Muhonen kysyi halukkuutta lähteä mukaan alatyöryhmään, mutta innokkuutta tällaiseen toimintaan ei vielä tästä ryhmästä löytynyt. Sen sijaan ehdotettiin, että IR:stä pidettäisiin erillinen seminaari vuoden 2004 syksyllä. Tätä pidettiin hyvänä ideana.

Yhteydenpitoon IR:stä kiinnostuneiden välillä ehdotettiin omaa sähköpostilistaa. Tämä kuitenkin todettiin liian raskaaksi. Oman listan sijaan voidaan hyvin käyttää FinnOA:n yhteistä sähköpostilistaa. Siihen voi ilmoittautua FinnOA:n kotisivujen (www.lib.helsinki.fi/FinnOA/) kautta.

Yhteenvedo

Institutional repository on vielä varsin uusi asia useimmille workshopissa mukana olleille organisaatioille. IR oli käsitteenä epäselvä, siksi kotimainen yhteistyö pitäisi aloittaa sen määrittelyllä. Tässä yhteydessä pitäisi myös miettiä, millä tavalla määritellään IR:n sisältö. Waaijersin mukaan IR on nimenomaan metadatatietokanta, mutta tiedon vapaan saatavuuden kannalta juuri kokotekstit ovat IR:n suola.

Yksi iso ongelmakokonaisuus on IR:n tekninen rakentaminen. Useimmilla organisaatioilla on julkaisurekisteri, jonka pohjalle IR voidaan pystyttää. Sopivan softan löytämiseen ja muokkaamiseen toivottiin kansallista yhteistyötä. Pääkaupunkiseudulla on jo alkanut projekti, joka tutkii DSpacea ja ehkä myös muita hyviä ohjelmakandidaatteja. HYK on hankkinut DOMS-järjestelmän, joka myös olisi mahdollinen alusta. Tämän vuoden kuluessa saadaan varmasti kaikkia hyödyttäviä vastauksia tekniikkakysymyksiin.

Kansallinen FinnOA-työryhmä seuraa IR:n kehitystä. Varsinaista alatyöryhmää ei vielä tässä vaiheessa kannata perustaa. On parempi ensin odottaa meneillään olevien hankkeiden tuloksia ja yleisen tietoisuuden lisääntymistä. Tavoitteena voisi kuitenkin olla ensi syksynä pidettävä IR-seminaari. IR:än liittyvää keskustelua voidaan käydä FinnOA:n sähköpostilistalla.