

# CERN Workshop on Innovations in Scholarly Communication (OAI4)

19.-22.10.05

CERNissä pidetty neljäs Open Archives Initiative –konferenssi jakautui kolmeen päivään, joiden sisältö oli yleispiirteissään seuraavanlainen: 1) OAI-PMH:n soveltaminen käytännössä; 2) Open access ja 3) julkaisut ja muut tutkimusaineistot.

OAI-PMH, pääosassa Los Alamos

Ensimmäisenä päivänä, torstaina, näkyvimmin esillä oli Los Alamosin tutkimuskeskuksen kirjaston merkittävä tutkimus- ja kehitystyö, joka etenee laajalla rintamalla. **Herbert Van de Sompel** itse ei tuonut esiin mitään järjestyttäviä uutuuksia – hän on käsitellyt samoja asioita muissakin yhteyksissä tänä vuonna, ja kirjoittanut aiheista lehtiin – mutta harvoin näkee yhtä innostavaa esiintyjää. Pelkän metadatan haravoimisen lisäksi on alettu haravoida aineistoja, jotka pakataan mukaan käärimällä ne XML-kuoreen. Järjestelmä toimii käytännössä Los Alamosissa, johon kootaan American Physical Society tuottamaa digitaalista aineistoa (ks. [artikkelia](#)). Vastaavien järjestelmien laajamittainen toteuttaminen tulee vaatimaan hyviä tietoliikenneyhteyksiä ja suuria tallennuskapasiteetteja; lisäksi DRM-aspektien on oltava kunnossa.

Van de Sompel puhui myös kehitystyöstä, jota tehdään digitaalisten varantojen yhteensopivuuden lisäämiseksi (aDORe-arkkitehtuuri). Tässä näkyy jo seuraava vaihe avointen digitaalisten varantojen kehityksessä, eli että puhutaan muustakin tieteellisestä aineistosta kuin julkaisuista. OAI-PMH:n alkuperäisestä tehtävästä, eli metadatan haravoinnista, voi mainita, että kehitystyössä on nyt paneuduttu metadatan valikoivaan haravointiin ajankohtaleiman (datestamp) perusteella.

Los Alamosin ryhmästä kannattaa mainita myös **Johan Bollenin** esitys. Hänen tutkimuksessaan on pyritty kehittämään impact factoria luotettavampia julkaisujen vaikuttavuuden mittaustapaa. Se tehdään analysoimalla julkaisujen käyttäjätietoja, jotka kootaan OAI-PMH-protokollan avulla. Los Alamosissa ei tosin puhuta julkaisuista, vaan tieteellisen kommunikaation yksiköistä (ks. [artikkelia](#)). Uskon että tämäntyyppinen julkaisujen vaikuttavuuden mittaaminen yleistyy muutaman vuoden sisällä.



Kuvassa CERNin nykyaikaisimmat rakennukset. Muuten kampus muistuttaa italialaista yliopistoa 1970-luvulla. Hienoimmat laitteet ovat tietysti maan alla.

**Stu Weibelin** esitys pysyvistä identifioijista oli pohdiskeleva ja ajatuksia herättävä. Hän puhui paljon URIen (Universal Resource Identifier) kehittämiseen ja hallinnoimiseen liittyvistä byrokraattisista ongelmista. Yhtenä osittaisratkaisuna tilanteeseen toimisi Weibelin ja Van de Sompelin kehittämä [INFO-URI](#).

Open access - julkaisut ja tutkimusaineistot

Perjantaina puhuttiin avoimesta julkaisutoiminnasta vähemmän teknisellä tasolla. **Alma Swanin** erinomaiseen esitykseen open accessin viimeaikaisista vaiheista kannattaa tutustua. Näkymät näyttävät sen perusteella hyvin valoisilta. Oli miten oli, minusta esityksessä erityisen tärkeä on kalvo 59, joka perustuu Swanin ja **Sheridan Brownin** tämänvuotiseen tutkimukseen (ks. esittely [aikaisemmassa Verkkarissa](#)). Kysymykseen "Miksi tutkijat julkaisevat tutkimuksiaan" vastaajien yleisin motiivi oli tulosten välittäminen tiedeyhteisölle. Perjantain ohjelmasta voi mainita myös äskettäin alkaneen brittiläisen *Versions*-projektin, jossa pyritään ratkaisemaan netissä pyöriviin julkaisun eri versioihin liittyviä ongelmia.

Lauantaina kuultiin hyvin kiinnostavia esityksiä siitä, miten tutkimusaineistoja on alettu ottaa mukaan tieteellisen kommunikaation prosesseihin. **Peter Murray-Rust** Cambridgestä käsitteli sitä, miten tämä on toteutettu kemian alalla. Hän aloitti esityksen toteamalla, että kemiaa julkaistaan edelleen samalla tavalla kuin 1860-luvulla! Lisäksi suuri osa informaatiosta jää kokonaan julkaisematta. Murray-Rust esitteli [tutkimusryhmänsä](#) kehittämiä välineitä, joilla tutkimuksen kohteita voidaan havainnollisemmin ottaa mukaan julkaisuihin. Koska alan informaatiota julkaistaan paljon tekstimuodossa, hän esitteli myös työkaluja, joita on kehittänyt etsimään tekstimuotoisissa kemiallisten kaavojen esityksissä olevia virheitä.

Seminaarin lopuksi kuultiin rahoittajan näkemys avoimeen tieteelliseen julkaisutoimintaan. **Robert Terry** edustaa merkittävää brittiläistä lääketieteen rahoittajaa Wellcome Trustia, jonka esittämä rahoittajan vaatimuslista on varsin kiinnostava; toivon näkeväni vastaavanlaisia joskus myös Suomessa. Terryn mukaan tutkimusrahoittajan vaatimuslista on seuraavanlainen:

- julkaisun välitön saatavuus
- pitkäkestoinen digitaalinen arkisto
- hakumahdollisuudet, tietyn alan julkaisut yhden väylän kautta
- olemassaolevan tutkimuskäytännön noudattaminen
- integrointi tutkimusaineistojen tietokantoihin
- strategia: parantaa arviointia ja vaikuttavuutta
- kestävä ja laadukas vertaisarviointijärjestelmä

Wellcome Trust on alkanut hyväksyä tutkimuksen sivukuluiksi open access – kirjoittajamaksut. (Lieneekö tämä sitten aiheuttanut sen, että BioMedCentralin

kirjoittajamaksut ovat jyrkästi nousseet.)

Kokonaisuutena OAI4 oli todella inspiroiva ja virkistävä sekä esitysten että sosiaalisen ohjelman puolesta. Se vakuutti ainakin minut siitä, että tieteellisen tiedon sujuvamman saatavuuden edistämiseksi OAI-PMH on keskeinen väline. Varsinaisia julkaisuarkistojen menestystarinoita ei esitelty, vaikka joissakin yliopistoissa julkaisuarkiston kriittinen massa onkin selvästi kasvanut. Erityisen onnistuneelta tuntuu Alankomaiden DARE-hanke, jossa julkaisuarkistojen aineiston kerääminen on aloitettu menestyneimmistä tutkijoista, ja muut seuraavat. Sovelluksiksi ovat monin paikoin vakiintuneet [EPrints](#) ja [DSpace](#), ja kaupalliset ohjelmistonvalmistajat loistivat poissaolollaan ExLibriksen sponsoroinnista huolimatta. Mainitut kaksi sovellusta tuntuvat tarjoavan varsin hyviä ja käyttäjälähtöisiä avoimen julkaisuarkiston palveluita, ja kehitystyö on vauhdikasta. Itse en ole koskaan ollut yhtä kiinnostavassa konferenssissa.

- Esitykset löytyvät [konferenssin sivuilta](#).
- [OAI4-uutisia](#) esitteli **Rita Voigt** [Metadatapäivässä](#) 10.11.
- **Herbert Van de Sompelin** [kotisivu](#), jossa artikkeleita.

*Kalle Korhonen  
kirjastopalvelujen koordinointiyksikkö  
puh. 191 21724  
kalle.korhonen at helsinki.fi*



« [EDELLINEN ETUSIVU SEURAAVA](#) »