

ISO TC 46

3.6.2008

Juha Hakala
juha.hakala (at) helsinki.fi

Kuten jo otsikosta ilmenee, tämä blogiteksti kertoo tietopalvelualan standardoinnista. Tässä vaiheessa lukija todennäköisesti huokaisee: onpa kuivakka aihe. Mutta ei syytä huoleen; tylsimmilläänkin standardityö on hyödyllistä, ja jos niikseen tulee, siinä voi olla kaikki hyvän draaman ainekset. Tuore esimerkki tästä on OOXML-standardin käsittely ISO:ssa. Siitä prosessista ei todellakaan jännitystä ja suuria tunteita puuttunut – onneksi lopputulos on jotakuinkin niin hyvä kuin näillä eväillä voi odottaa. Valmiin standardin (joka on haettavissa osoitteesta <http://www.ecma-international.org/publications/standards/Ecma-376>. (<http://www.ecma-international.org/publications/standards/Ecma-376.htm>)) lisäksi sovittiin siitä, miten OOXML:ää ja sen kanssa osin päällekkäistä ODF-standardia pyritään yhtenäistämään.

Tässä blogissa ei kuitenkaan ruodita OOXML-valmistelua, vaan kerron toukokuun 19.-23 päivä pidetystä ISO:n tekninen komitea 46:n kokouksesta.

Meitä suomalaisia oli paikalla kolme: Jaana Kilkki, jonka alaa on Archives and records management (TC 46:n alakomitea 11), Markku Laitinen (alacomitea 8, Quality – Statistics and performance evaluation) sekä allekirjoittanut. Erityisesti alakomitea 11 on viime vuodet lisännyt vauhtia kaiken aikaa. Toimintaan osallistuvien maiden ja asiantuntijoiden määrä on noussut, samoin kehitteillä olevien standardien määrä. Jaana on ansiokkaasti osallistunut useiden SC 11 –standardien kehittämiseen ja myös huolehtinut siitä, että nämä standardit julkistetaan Suomessa SFS-versioina.

SC 11:n toiminta antaa positiivisen esimerkin siitä, että vaikka standardointi on lähes sataprosenttisesti vapaaehtoistyötä, se voi saada aikaan tulosta niin että heikompaakin menoa hirvittää. Asiakirjahallinnon perusta on nyt määritelty ISO-standardien, mikä luo alalle samantapaista pohjaa kuin luettelointisäännöt aikanaan kirjastoille.

Standardointi ei ole itsetarkoitus, vaan se on teknistä kehitystyötä, jossa tavoitteena on käytännön hyöty. Me elämme maailmassa joka on osin standardien määrittelemä; ne takaavat sen, että tietokoneemme voivat kommunikoida keskenään Internetissä tai että voimme kytkeä kodinkoneet pistorasiaan. Yleisesti ottaen monimutkaisuuden lisääntyminen lisää standardoinnin tarvetta ja monimutkaistaa standardeja. Ensimmäinen merkkivalikoimastandardi American Standard Code for Information Interchange (ASCII; katso <http://en.wikipedia.org/wiki/ASCII> (<http://en.wikipedia.org/wiki/ASCII>)) kehitettiin 1960-1963 ja sisälsi 128 merkkiä, joista 93 oli tulostettavia. UNICODE-standardin (<http://en.wikipedia.org/wiki/Unicode> (<http://en.wikipedia.org/wiki/Unicode>)) uusimmassa versiossa 5.0 on noin 100.000 merkkiä. Onneksi suunta voi kulkea joskus myös monimutkaisuudesta yksinkertaisuuteen; tästä on alla pari esimerkkiä.

Vakiinnuttuaan standardeista tulee itsestäänselvyyksiä, joita emme edes ajattele. Luetteloiijat ottavat annettuna sen, että luettelointisääntöjen, MARC-formaatin ja Z39.50-tiedonhakustandardin ansiosta voimme kopioida julkaisujen bibliografisia tietoja periaatteesta kaikkialta. Mutta tekniikan kehitys voi ajaa standardiemme ja järjestelmiemme ohi. Nämä muutostrendit pitää tunnistaa, ja sovellustemme tekninen perusta modernisoida ajoissa, jotta emme maalaisi itseämme nurkkaan.

TC 46 on reagoinut muutamilla alueilla muutospaineisiin kiitettävän nopeasti. ISBN ja ISSN on uusittu, ja nyky muodossaan ne palvelevat meitä muutamia vuosikymmeniä, ennen kuin seuraava revisio on tarpeen.

Mutta käsissämme on myös muutama kuuma peruna, joiden osalta tulevaisuuden suunta on vielä auki.

Tässä blogikirjoituksessa otan esille vain kaksi standardia, tiedonhakuprotokolla ISO 23950:n eli ANSI/NISO Z39.50:n (katso <http://en.wikipedia.org/wiki/Z39.50> (<http://en.wikipedia.org/wiki/Z39.50>) ja <http://www.loc.gov/z3950/agency/> (<http://www.loc.gov/z3950/agency/>)) sekä kaukopalvelustandardi ISO ILL:n. Ne ovat molemmat niin sanotusti moniongelmaisia standardeja, kumpikin tavallaan.

Olen vilpittömästi sitä mieltä, että Z39.50 on, vaikkakaan ei ainoa tiedonhakustandardi, niin kuitenkin paras. Miksi se sitten on kirjastomaailman ulkopuolella täysin tuntematon?

Z39.50:ssä on ainakin pari atk-tekniisesti inhottavaa ratkaisua. Ensiksi, se edellyttää, että laitteistoriippumattomana datan koodaustapana sovelletaan ISO:n OSI-mallin ammoista ASN.1/BER –notaatiota (Abstract Syntax Notation One / Basic Encoding Rules). Standardin syntyäikoihin 80-luvun puolella ASN.1/BER oli parasta mitä rahalla sai, mutta 2000-luvun ratkaisu on XML. Toiseksi, standardi on yhteydellinen, eli asiakas- ja palvelinsovellus luovat keskinäisen istunnon, ja kummankin osapuolen on oltava "kartalla" sen suhteen, mitä tiedon hakija on viimeksi tehnyt. Edellisessä haussa syntyneen tulosjoukon käyttäminen seuraavassa haussa ei onnistu, jos palvelin on ehtinyt unohtaa käyttäjän hakujen välissä.

Jo vuosia sitten Z39.50:n kehittäjät päättivät kuoria standardista vanhan hapatuksen pois. Tuloksena syntyi SRU (Search/Retrieval via URL). Paperilla se on mainio tiedonhakustandardi. ASN.1/BER on korvattu XML:llä ja standardi on yhteydetön. Illuusio siitä, että käyttäjällä on istunto palvelinkoneella, luodaan client-sovelluksessa Nelli-portaalin tapaan. Z39.50:n eksoottiset palvelut joita ani harva kirjastojärjestelmä toteutti, on karsittu SRU:sta pois.

Valitettavasti SRU:stakaan ei ole tullut maailman suosituinta hakuprotokollaa eikä toistaiseksi edes standardia. Päinvastoin, SRU:lla on koko joukko kilpailijoita kuten Amazonin A9, jotka pyrkivät oikealta ja vasemmalta ohi. Mikä meni pieleen?

SRU:n kehittäjät pyrkivät välttämään kirjastoleiman sekä jäykäksi koetun ISO-standardointimekanismin. Jo muutama vuosi sitten päätettiin valita SRU:n standardointiympäristöksi OASIS (Organization for the Advancement of Structured Information Standards, <http://www.oasis-open.org/home/index.php> (<http://www.oasis-open.org/home/index.php>)). Tämä on merkinnyt vuosien viivästystä. Kongressin kirjaston mukaan OASIS-standardointi on vasta nyt alkamassa. Kukaan ei tiedä, miten paljon työ vie aikaa, ja miten paljon muutoksia SRU:hun joudutaan tekemään.

Pyrkimys markkinoida SRU:ta kirjastojen ulkopuolelle voi tulla kalliiksi, mutta eläkäämme toivossa että saamme loppujen lopuksi uuden ja entistä ehomman SRU-standardin. OASIS-versio on tarkoitus hyväksyttävä myös ISO:ssa saman Fast track –prosessin avulla, jota sovellettiin myös Microsoftin ajamaan OOXML-standardiin. Oletan että SRU:n kohdalla prosessi tulee olemaan tuskattomampi kuin OOXML:n tapauksessa.

Z39.50:n tuki ei tule kirjastojärjestelmistä katoamaan lähivuosina. Todennäköinen kehityskulku on se, että SRU-tuki rakennetaan Z39.50-palvelimen päälle asennettavilla apuohjelmilla, jollaisia on jo nyt saatavilla. Eri asia on sitten se, tukevatko tulevat XML-pohjaisia metadataformaatteja tukevat yhteiset käyttöliittymäsovellukset enää Z39.50-standardia, tai onko Z:lle enää käyttöä vaikka se olisikin tuettu.

ISO ILL:n tapauksessa vanhan protokollan katoaminen kuvioista on luultavasti totaalisempi ilmiö, vaikka tässäkin tapauksessa standardi häipyi takavasemmalle "not with a bang but a whimper".

ISO ILL jakaa Z39.50:n kanssa tiedon koodaukseen liittyvän ongelman. Lisäksi sillä on vielä omia vakavia vaikeuksia, jotka liittyvät siihen, miten kaukopalvelu on toiminnallisesti muuttunut.

ISO ILL on erittäin monimutkainen protokolla, mitä tulee kaukopalveluyksikköjen ja -sovellusten väliseen viestinsiirtoon. Toiminnan perusfilosofia on peräisin 70-luvulta, suurista tieteellisistä kirjastoista. Tilaukseen ei välttämättä tule suoraa vastausta, vaan vastaanottajakirjastolla on lukuisia eri vaihtoehtoja tilauksen edelleen välittämiseen jne. Tämän monimutkaisuuden hallinta on viime kädessä henkilökunnan

vastuulla, mutta myös sovellusten on pidettävä jokaisen tilauksen osalla tarkasti kirjaa siitä, missä mennään.

ISO ILL:n tapauksessa kävi aikaa myöten ilmeiseksi, että vanhasta protokollasta voitu siirtää uuteen tekniseen ja toimintaympäristöön mitään. 2000-luvun mittaan standardille on hissun kissun kehitetty seuraajaa, jonka vähemmän herttainen nimi on The OpenURL Request Transfer Message (eli lyhesti RTM; <http://oclcpcica.org/?id=1409&ln=uk> (<http://oclcpcica.org/?id=1409&ln=uk>)). Kuten osoitteesta näkee, OCLC:llä on ollut aktiivinen rooli standardin kehittämisessä. Kehittäjäryhmä kuvaa tavoitetta näin:

The OpenURL Request Transfer Message (RTM) will be used to direct requests for loan, copy, access to lookup or digitization of an item. It is intended to be used for directing requests to Resource Delivery Systems and / or Electronic Resolver Systems.

Esimerkit RTM:n erilaisista käyttötavoista ovat valaisevia: niistä näkyy hyvin se, miten kaukopalvelutoiminta on muuttunut kolmessakymmenessä vuodessa:

- *a request placed in Google Scholar by a user associated with the University of Amsterdam that is sent by Google to the national Dutch delivery service (NCC).*
- *a request placed in worldcat.org by a user associated with the University of Sydney that is sent by worldcat.org to the national Australian delivery service (Libraries Australia).*
- *a request placed in Libraries Australia by a user associated with the University of Amsterdam that is sent by Libraries Australia to the national Dutch delivery service (NCC).*

Esimerkeistä voi helposti nähdä senkin, että RTM-määrittelyyn perustuvat sovellukset olettavat kaukopalvelun olevan asiakastoiminteista.

Toistaiseksi emme tiedä, milloin RTM:stä tulee standardi. Välttämättä näin ei käy milloinkaan. ISO TC 46 etenee asiassa varoen. Niinpä kaukopalvelusovelluksen hankintaprojektissa on pidettävä standardien suhteen mieli avoinna: vanhan standardin tuki on toistaiseksi pakollinen, ja sen rinnalle voi toivoa uutta.

Joka tapauksessa tulee viemään vuosia ja runsaasti työtä, ennen kuin SRU ja RTM – tai niiden seuraajat – korvaavat lopullisesti nykyiset standardimme. Ja koko TC 46:n agendalla nämä kaksi protokollaa ovat vain muru: komitean agendalla on lähes sata standardia, ja niiden määrä kasvaa jatkuvasti. Erityisen haasteen asettavat uusien järjestelmien kuten digitaalisen aineiston pitkäaikaissäilytysjärjestelmien edellyttämät standardit, koska niiden kehittämiseen tarvittavaa asiantuntemusta on kovin harvoilla, ja heillä on yleensä muutakin tekemistä kuin standardien rustaaminen.

Luotan joka tapauksessa siihen, että jatkossakin saamme aikaan ne määritykset, joita järjestelmämme edellyttävät. Suomessa tätä urakkaa hoitaa omalta osaltaan opetusministeriön Kansallinen digitaalinen kirjasto –projektin tekninen asiantuntijaryhmä, jonka huomio tulee kohdistumaan erityisesti uusien asiakasliittymäsovellusten eli yhteisten tiedonhakuprosessien sekä pitkäaikaissäilytyksen edellyttämiin standardeihin ja muihin määrityksiin. Tässä työssä tulemme varmasti olemaan tiiviissä yhteistyössä ISO:n ja muiden standardointijärjestöjen sekä kansainvälisten kollegoidemme kanssa.

Pyydämme lähettämään palautteen suoraan kirjoittajalle:
juha.hakala(at)helsinki.fi