

VYP - Virtuaaliyliopiston portaaliryhmä

www.virtuaaliyliopisto.fi

Kurssitiedon siirtopilotti Raportti

Omistaja	VYP
Tyyppi	Raportti
Tiedoston nimi	siirtopilotit_raportti.tex
Turvaluokitus	julkinen
kohderyhmä	SVYKY/VYP/kurssitietotyöryhmä/yhteistyöyliopistot

Versiohistoria

Versio	Päiväys	Editoija	Keskeiset muutokset
0.9	27.1.2004	mtoivone	dokumentti kesken
1.0	28.1.2004	mtoivone	ensimmäinen versio
1.1	2.2.2004	mtoivone	korjaukset dokumettiin

Sisältö

1	Johdanto	2
2	Pilotin tavoite	2
3	Pilotin taustat	2
4	Pilotin käynnistyminen	3
5	Pilotin eteneminen - aikataulut	4
5.1	Siirtodokumentin toteutus	4
5.2	Kenttien merkitykset ja siirtoprotokollat	4
5.3	Toteutus	4
5.4	Testaus	5
6	Pilotin tekninen toteutus	6
7	Pilotin toimivuus	7
7.1	Jyväskylän yliopisto	7
7.2	Svenska handelshögskolan	7
8	Kommentit kurssitiedon standardointityöryhmälle	8

1 Johdanto

SVY:n portaalin Kurssitarjonta on toiminut kesästä 2001 alkaen verkko-opetusta tuottavien yliopisto-opettajien ja kursseja suorittavien opiskelijoiden julkaisu- ja tiedohakualustana. Kurssitarjonnan graafiseen ilmeeseen ja käytettävyyteen on panostettu useissa kehitysversioissa. Viimeisin 10/2003 version päivitys loi palvelulle lopullisen ilmeen ja uudisti samalla palvelun teknistä toteutustapaa.

Yksi keskeisimmistä palvelun toimintaa haittaavista seikoista on ollut sopivien tiedonsiirtorajapintojen puute. Tarvittavien rajapintojen puuttuessa kurssien syöttö Kurssitarjonnan tietokantaan on toteutettu opettajille tarjotun syöttölomakkeen kautta. Kurssien syöttö on kuitenkin ollut täysin verkko-kurssien tuottajien aktiivisuuden varassa. Samojen kurssitietojen syöttö kahteen paikkaan koettiin ongelmalliseksi ja kurssien lukumäärä jäikin vain noin 80 kurssin tasolle. Palvelua on pyritty mainostamaan erilaisissa SVY:n järjestämissä tilaisuuksissa, jolloin kurssien lukumäärä on hetkeksi noussut, mutta pudonnut nopeasti tämän jälkeen normaalille tasolle. Opiskelijoiden keskuudessa palvelu on todettu mielenkiintoiseksi ja kurssien hakuja on tehty noin 10 000 kuukaudessa.

2 Pilotin tavoite

Kurssitiedon siirron pilotoinnin tavoitteena oli tutkia menetelmiä kurssitiedon siirron automatisoimiseksi yliopistojen opintotietojärjestelmistä SVY:n portaalin Kurssitarjontaan. Työn perustana käytettiin TIEKE:n standardointityöryhmän kehittämää kurssitiedon metatietomäärittystä [1] sekä CSC:n laatimaa implementaatiota määrittelytyöstä (XML-skeema) [2].

3 Pilotin taustat

SVY:n portaalin tekniikasta vastanneisiin henkilöihin otettiin yhteyttä ensimmäistä kertaa jo Oulun yliopiston Campus Futurus -hankkeen toimesta 11.11.2002. Oulun yliopiston Maija Väänänen lähestyi tarpeella siirtää portaalin Kurssitarjonnasta Oulun yliopiston kurssit näkymään Campus Futurus-palvelun verkkosivulla.

Kurssitiedon metadata-hanke oli tässä vaiheessa kesken ja mitään tuloksia ei oltu julkaistu. CSC näki, että kurssitietojen syöttö manuaalisena portaalin Kurssi-

tarjonnan tietokantaan ja edelleen sieltä takaisin yliopistojen omiin järjestelmiin ei ollut teknisesti järkevää. Keskusteltiin tahosta, joka olisi oikea määrittelemään opistotietojärjestelmiin liittyviä asioita. Asia päätettiin ottaa uudelleen esiin vuoden 2003 puolella.

Campus Futurus -hankkeesta otettiin jälleen yhteyttä 21.3.2003. Edelleen käytiin keskustelua määrittelytyön problematiikasta. Kansallinen määrittely kurssitiedolle puuttui edelleen. Lisäksi keskustelua SVY:n portaalin kurssitarjonnan ja Oodi- opintotietojärjestelmien yhteensopimattomuudesta. Välivaiheen ratkaisuna Campus Futurus-hankkeelle toimitettiin SVY:n portaalin omaan Kurssitarjontaan pohjautuva kuvaus kurssien rakenteesta ja toteutettiin portaalin kurssitietokannasta XML-siirtodokumenttia luova PHP-skripti. Tämän jälkeen SVY:n portaalin Kurssitarjontaan tallennetut Oulun yliopiston kurssitiedot oli haettavissa HTTP:n yli CF:n järjestelmiin ja työ katsottiin CSC:n osalta suoritetuksi.

4 Pilotin käynnistyminen

Sysäyksen pilotin käynnistymiseen antoi keväällä 2003 TIEKE:n koordinoiman standardointityöryhmän tekemä ehdotus kurssitiedon XML-skeemasta (Harri Salminen, CSC).

5 Pilotin eteneminen - aikataulut

5.1 Siirtodokumentin toteutus

1.8. - 31.8.2003: TIEKE:n standardointityöryhmän ja CSC:n Harri Salmisen työn tuloksena syntyi elokuussa 2003 kurssitiedon XML-skeema. Yhteistyökumppaneita etsittiin ja kiinnostuksena pilotointiin ilmoitti Jyväskylän yliopisto Antti Auerin johdolla. Skeeman version v0.2:n pohjalta laadittiin siirtodokumentti, joka toimitettiin Korppi-opintotietojärjestelmän ylläpidosta vastanneille henkilöille (Pauli Kujala, Vesa Lappalainen).

5.2 Kenttien merkitykset ja siirtoprotokollat

1.9. - 30.9.2003: Keskustelua Harri Salmisen kanssa siirtodokumentin kenttien merkityksestä. Päivämäärien ja kellonaikojan kirjaaminen herätti Jyväskylän suunnalla ihmetystä. (kts. muutosehdotukset).

Alustavaa keskustelua kurssien tiedonsiirtoprotokollasta. Päädyttiin toteuttamaan siirto Jyväskylän kanssa eräajotyypisesti suoraan HTTP:n yli XML-siirtodokumenttina. Järjestelmässä siirto ajatettaisiin tapahtuvaksi joka yö. Samalla käytiin keskustelua myös hakutoiminnoista ja hakuun liitettävistä parametreista.

Kurssimäärittelyn eri kenttien merkitys oli tietyin osin ongelmallista selvittää. Päätettiin toteuttaa siirto vain sellaisten kenttien kohdalla, jossa voidaan taata tietojen semantiikka eri järjestelmien välillä.

Svenska handelshögskolan ilmoitti kiinnostuksensa tulla mukaan pilottiin. Hankenille toimitettiin vastaava kurssikuvaus kuin Jyväskylälle.

5.3 Toteutus

1.10 - 31.10.2003: Kurssitarjonnan sen hetkinen kehitysversio siirrettiin tuotantoon ja siirtopilotin tekninen toteutus voitiin aloittaa kehityskoneella. Ensimmäiset testikurssit siirtyivät tietokantaan puoliautomaattisesti. 17.10.2003 toimitettiin Jyväskylän yliopistosta CSC:lle muutosehdotuksia kurssitietomäärittelykseen. Nämä toimitettiin edelleen Harri Salmiselle ja kurssitietotyöryhmän käsiteltäväksi. Kurssitietomäärittelykseen ehdotettiin muutoksia, mutta pilotti päätettiin viedä läpi v0.2 määritysten pohjalta.

5.4 Testaus

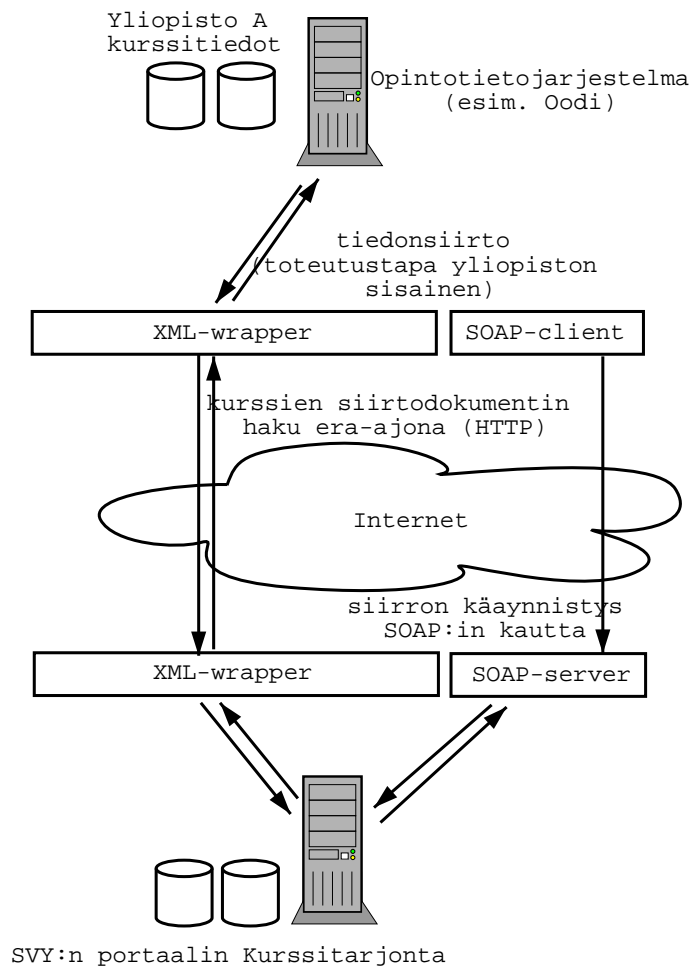
1.11.-30.11.2003: SOAP-rajapinnat testikäytössä sisäisesti sekä Hankenin kanssa. Hanken raportoi teknisistä ongelmista yhteydenotossa. Annettiin teknistä tukea mm. testaamalla Hankenin hakuskriptejä portaalin palvelimella (toimivat ok). Annettiin Hankenin käyttöön vastaava versio portaalin palvelimella käytössä olevasta NuSOAP-kirjastosta, mutta ongelmat jatkuivat. Pilottia jatkettiin Jyväskylän yliopiston kanssa.

1.12.2003-15.1.2004: Kytettiin kurssisiirto toimintaan ja hakemaan Korpin tuotantopalvelimen dataa. Törmättiin välittömästi ongelmiin suurien siirtotiedostojen muodostuksessa ja siirrossa muodostuvan latenssin kanssa. Lisäksi käytetty ActiveLink XML-kirjasto ei kyennyt käsittelemään suuria XML-tiedostoja. Joululoman ajan palvelu oli pois käytöstä ja siirrossa turvauduttiin erityisjärjestelyihin, missä siirtodokumentti ensin siirrettiin palvelimelle, jaettiin pienempiin osiin ja vasta tämän jälkeen ajettiin Kurssitarjonnan kurssikantaan. Kurssitarjonnan kurssikantaan siirrettyjen kurssien lukumäärä on noin vajaa 1000 kurssia.

6 Pilotin tekninen toteutus

Pilotti toteutettiin CSC:n puolella SVY:n portaalin kurssitarjonnan¹ yhteyteen tehtyinä muutostöinä. Lisäksi pilotin yhteydessä tarvittiin muutostöitä saattamaan yliopistojen opintotietojärjestelmät tuottamaan metatietomäärittelyn mukaista kurssitietoa.

CSC:n toteutus perustui XML-pohjaisten tiedonsiirtorajapintojen käyttöön ja näiden toteutuksessa hyödynnettiin PHP:n NuSOAP²- ja ActiveLink XML³- ohjelmakirjastoja.



Kuva 1: Kurssitiedon siirron arkkitehtuurikuvaus

¹SVY:n portaalin Kurssitarjonta: <http://www.vy.fi/kurssitarjonta>

²NuSOAP: <http://dietrich.ganx4.com/index.php?category=NuSOAP>

³ActiveLink XML: <http://www.active-link.com/intranet/software.php>

Kuvassa 1 esitetään järjestelmän arkkitehtuurikuvaus. Perusajatuksena on se, että yliopistojen ja SVY:n portaalin puolella kurssitieto muunnetaan sellaiseen muotoon, että se vastaa rakenteeltaan Tiede:n kurssitietomäärittelyä. Tämän jälkeen XML-pohjainen siirtodokumentti on siirrettävissä eri järjestelmien välillä HTTP:n yli. Siirto voidaan käynnistää joko ajastuksin tai SOAP-kutsurajapinnan kautta yliopistoista.

7 Pilotin toimivuus

7.1 Jyväskylän yliopisto

Pilotti on toiminnassa Jyväskylän yliopiston Korppi-opintotietojärjestelmän ja SVY:n Kurssitarjonnan kehitysversion välillä. Kurssitiedosta siirretään vain ne kentät, jotka ovat semantiikaltaan yhteensopivia järjestelmien välillä.

7.2 Svenska handelshögskolan

Hankenin kanssa ajettiin testejä SOAP-kutsurajapinnan testaamiseksi, mutta teknisten vaikeuksien vuoksi siirrot eivät onnistuneet. Syyksi arveltiin versioeroavaisuuksia NuSOAP-kirjaston kanssa. Hankenin käyttämiä hakuskriptejä testattiin CSC:n palvelinympäristössä ja ne toimivat virheettömästi. Projektiryhmän toimesta Hankenin käyttöön toimitettiin SVY:n portaalissa käytetty versio NuSOAP-kirjastosta, mutta ongelmaa jatkui ja sitä ei ehditty pilotin puitteissa selvittämään. SOAP-kutsurajapinnan käyttöä siirron käynnistämiseen jouduttiin testaamaan vain CSC:n omilla palvelinlaitteistoilla.

8 Kommentit kurssitiedon standardointityöryhmälle

Tähän osioon on kerätty kurssitiedon standardointityöryhmälle kerättyjä kommentteja. Kommentit ei ole järjestetty mihinkään erityiseen tärkeysjärjestykseen.

1. Kurssitiedon metatietomäärittelyn tarkempi dokumentointi.

Kurssitietomäärittelyssä esiintyneiden kenttien merkityksen dokumentointi ei nykyisessä versiossa (v0.2) ole riittävä järjestelmän kattavaa implementointia ajatellen. Semantiikkakysymykset nousevat esiin useiden kurssitietomäärittelyn kenttien kohdalla. Tietyllä kentällä saattaa olla eri merkitys kahden eri opintojärjestelmän välillä.

2. Kurssitietomäärittelyn aikaleima

Määrittelyssä käytetty tapa kuvata aikaleima herätti kummastusta. Aikaleimoja käytettiin useassa kohtaa siirtodokumenttia ja se tarpeettomasti kasvatti dokumentin pituutta.

```
<vuosi>2003</vuosi>  
<kuukausi>09</kuukausi>  
<paiva>15</paiva>  
<tunti>09</tunti>  
<minuutti>00</minuutti>
```

Parempi tapa siirtodokumentin pituutta ajatellen voisi olla esimerkiksi:

```
<aikaleima>YYYYMMDDhhmm</aikaleima>
```

3. Joustavan opinto-oikeuden kurssien esitystapa siirtodokumentissa

Yliopiston opintotarjonta voi sisältää kurseja, mitkä on tarkoitettu vain yliopiston omille opiskelijoille tai ne voivat olla avoimia muiden yliopistojen opiskelijoille joustavan opinto-oikeuden turvin. Miten tämä tieto (jookurssi: kyllä/ei) merkitään kurssimäärittelyssä? Jyväskylän Korppi- opintotietojärjestelmään ja SVY:n portaalin kurssitarjontaan jouduttiin tekemään tämän seikan vuoksi muutostyö.

4. Avainsanat

Miten käytännössä voidaan kontrolloida, että käyttäjä antaa tietyn asiasanaston mukaisia avainsanoja (esim. YSA)? Mikä on schemen tyyppi, jos kurssin syöttäjä haluaa määritellä avainsanat itse?

5. Käytetyn koodausjärjestelmän yksikäsitteinen ja pysyvä tunnus

Tämä vaatisi tarkemman selvityksen. Pilotin aikana tämän kentän merkitys ei auennut. Viitataan tällä kurssin koodiin, joka pysyy aina samana vuodesta toiseen vai onko esim. juokseva numero?

6. **Attribuutti *tarkiste***
Miten MD5-tiivisteen on ajateltu toimivan? Nyt se sijoittuu varsinaisen siirtodokumentin sisään, mistä tiiviste pitäisi laskea ja aiheuttaa sekaannusta. Pitäisi varmaan selventää, että tiivisteen laskeminen kohdistuu vain itse kurssidataan ja määrittää alku- ja loppukohdat.
7. **Vaihteleva opintoviikkojen määrä**
Miten merkitään tilanne, mikäli kurssilta on saatavissa vaihteleva määrä opintoviikkoja. Tarvittaisiin minimi- ja maksimiopintoviikkomäärät erikseen.
8. **Kurssin maksullisuus**
Tarvittaisiin kentät tiedolle, jos kurssi on maksullinen sekä tarvittavat maksutiedot.
9. **Kurssin sijoittuminen opintovuoteen**
Tarvittaisiin lisäys määrittelyyn, että mihin kohtaa lukuvuotta kurssi sijoituu. Kevät, kesä, syksy, talvi, muut opintojaksot ja satunnaisesti luennoitavat.
10. **Tarkemmat yhteystiedot**
Tarvitaan määrittelyyn tarkemmat tiedot opetuksen järjestäjistä. Osoite, toimipaikka ja muut yhteystiedot.
11. **Osallistujamäärät**
Tarvitaan määrittelyyn tieto opetusryhmän minimi- ja maksimi osallistujamääristä.
12. **Kurssille pääsy**
Voiko harvemmin järjestettävillä kursseilla jonottaa? Miten se kirjataan määrittelyyn?
13. **Kurssin kuvaus** Tarvittaisiin uusi kenttä. Miten esitetään tekstimuotoinen lyhyt kuvaus (esim. 200 merkkiä) kurssin sisällöstä? Tämä on tarpeellinen hakutoimintoja ja kurssin tietojen listausta ajatellen.
14. **Kurssin toistuvuus** Kentän merkitys ei auennut pilotin aikana. Ehkä tämän toimintatapaa voisi purkaa dokumentoinnissa paremmin auki.

Lähteet

- [1] *TIEKE standardointiryhmä*: Kurssitiedon metatietomäärittäminen
<http://www.tieke.fi> → opetusteknologia → kurssitietomäärittäminen
- [2] *Harri Salminen, CSC*: Kurssitiedon metatietomäärittäminen XML-skeema
http://meta.tv.funet.fi/schemaDocs/kurssitieto_xsd