

Loppuraportin liite 4

Konversiosuunnitelma

Opintohallinnon tietojärjestelmän modernisointi
Esiselvitys

Johtoryhmä

Pekka Kähköpuro, IT-johtaja, Aalto, (pj)
Satu Kekäläinen, asiakkuuspäällikkö, Aalto
Mikko Markkola, osastopäällikkö, TaY
Kati Kettunen, palvelujohtaja, HY
Merja Eklín, tietohallinnon kehittämispäällikkö, HY
Anneli Lappalainen, opinto- ja
opiskelijapalveluiden vastaava, Aalto
Ilkka Siissalo, tietohallintojohtaja, HY
Susanna Wolkoff, kehittämispäällikkö, HY (siht.)

Projektiryhmä

Tuomas Naakka, projektipäällikkö, HY
Tuomas Hulkkonen, projektisuunnittelija, Aalto
Mari Riihiaho, projektisuunnittelija, Aalto
Sami Hautakangas, tietojärjestelmäpäällikkö, Tampereen yliopisto
Timo Kauramäki, tietotekniikkapäällikkö, HY
Susanna Wolkoff, kehittämispäällikkö, HY

LOPPURAPORTIN LIITTEET

Liite 1: prosessit

Liitteessä kuvataan opintosektorin ylimmän tason prosessit ja annetaan esimerkkejä joidenkin prosessien tavoiteloista. Liite sisältää lisäksi kuvauksen opintosektorin prosessien reunaehdoista ja tiedossa olevista muutoksista, jotka pitää huomioida tavoiteloja määriteltäessä. Liitteessä on myös kuvattu suomalaisen korkeakoulukentän opintosektorin käsite- ja tietomallityön tilanne.

Liite 2: markkinakartoitus

Liitteessä kuvataan esiselvityksessä tehty markkinakartoitus, jossa tunnistettiin ja tutkittiin markkinoilla olevia opintohallinnon järjestelmiä sekä tutustuttiin valmisjärjestelmiä käyttävien yliopistojen ja niiden yhteistyöorganisaatioiden hankintoihin ja toimintaan. Liite sisältää kaksi osaa:

- a) kuvauksen markkinakartoituksesta ja tuloksista
- b) taulukon tunnistetuista valmisjärjestelmistä

Liite 3: järjestelmäarkkitehtuuri

Liite sisältää

- Toiminnallisia kehittämissideoita ja niiden teknisiä ratkaisuehdotuksia
- Yliopistojen tietojärjestelmäpalvelukartat tavoitetilassa 2015-2016. OTM-projektin rajaus on kuvattu sillä tarkkuudella, mikä oli esiselvitysvaiheessa mahdollista. Toiminnallisuuksia ei ole selitetty, mutta ne käyvät ilmi prosessikuvauksista (erillisessä liitteessä). Opetuksen ja opintojen suunnittelun, lukuvuosi-ilmoittautumisen sekä koulutustarjonnan kuvaamisen osalta tarkka rajaus ja toteutusvaihe päätetään riippuen valitusta ratkaisuvaihtoehdosta, rahoituksesta, yliopistojen tietojärjestelmätilanteesta sekä ulkopuolisista tekijöistä (etenkin kansallinen haku- ja valintajärjestelmähanke Kotve/KSHJ huomioitava).
- Yliopistojen nykyisten opintohallinnon järjestelmien liitännät ja liitännät tavoitetilassa
- Tavoitetilakuvaus avoimen yliopiston toiminnoista opintohallinnon järjestelmässä

Liite 4: konversiosuunnitelma

Liite sisältää alustavan kuvauksen siitä miten konversio nykyisestä opintohallinnon järjestelmästä Oodista uuteen järjestelmään tehtäisiin työmääräarvioineen. Suunnitelma on tehty suurimmaksi osaksi Helsingin yliopiston Oodin tietojen näkökulmasta. Aalto-yliopisto on pitemmällä Oodin tietojen kuvaamisessa korkeakoulujen yhteisen käsitelmän XDW:n mukaiseksi, ja lisäksi Aalto on liittänyt vuonna 2010 kolmen korkeakoulun Oodit yhteen tietokantaan, joten Aallon työmäärä voi HY:n työmäärää pienempi. Tampereen yliopisto on päättänyt tehdä konversio ns. SOA eli palvelupohjaisena toteutuksena, joten TaY:n nykyisen opintohallinnon järjestelmän Opsun konversiota ei ole kuvattu.

Liite 5: ratkaisuvaihtoehtoihin liittyvää materiaalia

- ratkaisuvaihtoehtojen soveltuvuus yliopistojen kokonaisarkkitehtuuriperiaatteisiin
- organisoituminen ja eteneminen

Esiselvityksessä on verrattu eri ratkaisuvaihtoehtojen ja nykytilan soveltuvuutta HY:n ja Aallon kokonaisarkkitehtuuriperiaatteisiin. Lisäksi on luonnosteltu projektin organisoitumisen vaihtoehtoja. Mahdollisen keskitetyn ohjelmaa hallinnoivan organisaation rakenteen kuvaaminen ja siihen liittyvät selvitykset tehdään myöhemmässä vaiheessa.

Sisällys

1 Tavoitteet	2
2 Oodi-tietokannan tunnuslukuja ja kuntoarvio	2
3 Konversion laatutavoitteita ja havaittuja ongelmakohtia	4
4 Konversion toteutusvaihtoehtoja	6
5 Konversion työvaiheet ja aikataulu karkealla tasolla	9
6 Kustannus- ja työmääräarvio	10

1 Tavoitteet

Tässä liitteessä kuvataan Oodi-tietokannan konversion toteutusvaihtoehtoja. Selvitys antaa myös alustavan kuntoarvion Oodi-tietokannasta ja alustavan kuvan konversion työvaiheista ja työmääristä. Tampereen yliopisto on päättänyt toteuttaa konversiot SOA-pohjaisesti, joten konversiota Opsusta opintohallinnon tietojärjestelmän modernisoinnissa (OTM) toteutettavaan järjestelmään ei ole kuvattu.

Selvityksen taustatietoina olivat Oodi-tietokannan keskeiset tunnusluvut ja tietokannan kuvaukset. Selvitystä varten haastateltiin Helsingin yliopiston opetus- ja opintopalveluiden ja tietotekniikkakeskuksen asiantuntijoita. Lisäksi apuna käytettiin Pauli Saarisen laatimaa Oodi-järjestelmän käyttöönottoprojektin loppuraporttia (2000) ja Mikael Gueckin laatimaa Oodi-selvitysraporttia (Oodi-teknologiakonsultointi 2007).

Konversioon liittyviä termejä ja määritelmiä

integraatio

Integraatiolla tarkoitetaan tässä sovellusten tai järjestelmien liittämistä toisiinsa.

konversio

Konversiolla tarkoitetaan tiedon siirtämistä talletusmuodosta toiseen esim. vanhan tietokannan tietojen siirtoa uudempaan kantaan tai tietokantojen yhdistämistä.

2 Oodi-tietokannan tunnuslukuja ja kuntoarvio

Oodi-tietokannan keskeisiä tunnuslukuja

Selvityksessä keskityttiin Helsingin yliopiston Oodi-tuotantotietokannan kuvauksiin ja sen tunnuslukuihin.

Konversion kannalta keskeisiä tunnuslukuja ovat: lähtötiedon tietokannan tyyppi ja versio, tietokannan fyysinen koko, konvertoitavan datan määrä, tietokantataulujen määrä ja tietokantataulujen datarivien määrät.

Esiselvityksessä käytetyt tunnusluvut ovat HY:n Oodi-asennuksen tuotantoympäristöstä. Tunnuslukujen arvot poimittiin huhtikuussa 2012.

- Oracle-tietokanta, versio 10.2.0.4.9
- Tuotantokannan koko on 37,89 gb
- Tietokantatauluja on yli 1800 kpl, joista huomattava osa on vanhoja tilapäis- tai työtauluja
- Tietokannan kokonaisrivimäärä on hieman yli 110 miljoonaa riviä
- Kun tuotantokannasta rajataan pois erilaiset työ-, loki- ja aputaulut saadaan noin 450 - 500 konvertoitavaa dataa sisältävä taulujen joukko. Mikael Gueckin Helsingin yliopistolle vuonna 2007 tekemässä Oodi teknologiakonsultoinnissa mainitaan taulujen lukumääräksi 580. CSC:n

Confluence Wikissä on dokumentoitu noin 430 tietokantataulua (<https://confluence.csc.fi>). Arviot ovat siis välillä 430 – 580 tietokantataulua.

- 50 suurimman tietokantaulun rivimäärä on yhteensä yli 95 miljoonaa riviä.
- Konversion kannalta keskeisissä tietokantataulussa on yli miljoonan rivin tauluja 11 kpl.
- Konvertoitavaa dataa on arviolta yhteensä noin 65 – 70 miljoonaa riviä.

Johtopäätöksiä:

- ⇒ Vaikka konvertoitava datamäärä on suuri, ei konversion läpimenoaika ja suorituskyky nykytekniikalla kuitenkaan ole merkittävä riski. Konversion tulosten huolellinen tarkastaminen sen sijaan on todennäköisesti hyvin työlästä, joten läpimenoaikojen rajat on suunniteltava huolellisesti tarvittavan tuotantokatkon takia.
- ⇒ Konversion suorituskyvyn varmistamiseksi tarvitaan kuitenkin kokeneita ammattilaisia sitä suunnittelemaan, määrittelemään ja toteuttamaan.
- ⇒ Konversion tulosten tarkastamiseen joudutaan ehkä hankkimaan tai teettämään joitain apuohjelmia ja työkaluja.
- ⇒ Oleellista konversiossa on Oodi-tietokannan rakenteiden ja sisällön hyvä tuntemus – aina suunnittelusta toteutukseen, testaukseen ja tietojen tarkastamisvaiheisiin saakka.

Esimerkkilistaus eräistä keskeisistä Oodi-tuotantotietokannan tauluista (20 kpl) ja niiden rivimääristä:

TABLE_NAME	NUM_ROWS
OPINARVSAN	13 068 336
OPINTO	8 071 623
OPINOHEITAH	7 081 417
ORGOSUU	4 507 226
KIINNITYS	3 758 488
LKILM	2 559 394
OPINTAP	2 440 203
VASTAUS	2 363 449
OPINOIKKOHD	2 050 818
OPINKOHTTEKS	1 779 252
OPISHYOD	1 128 892
WEBPALAUTEOPETTAPORG	884 469
ILMNAYTKIELL	837 284
KAYTLOKI	765 607
RAPHALLVAKHISTORIA	736 620
TODPYYN	719 326
OPINOIK	718 326
HLOKUNNROOL	686 147
SUUNSUOR	649 789
OPINTAPOSA	640 860
Yhteensä	55 447 526

Tietokannan kuntoarvio

Keskeinen tekijä konversion onnistumiseen ja käytettyihin työmääriin on tietokannan kunnolla. Lähtötietojen kuntoarviolla haetaan vastausta siihen, onko alkuperäinen tieto tarpeeksi laadukasta siirrettäväksi seuraavaan elinkaaren vaiheeseen ja miten se sellaiseksi tunnustetaan ja tehdään. Kunkin yliopiston tietokannan kuntoarviossa pitää kiinnittää huomiota mm. seuraaviin tekijöihin:

- tietojen oikeellisuus ja eheys
- tietojen kattavuus
- lähtötiedon kannan tyyppi

Tarkkaa Oodin lähtötietojen kuntoarviota ei esiselvityksessä tehty. Esiselvityksessä tehtiin ainoastaan alustava läpikäynti Oodi-tietokannan tauluista, jotta saadaan hahmotettua konvertoitavan tiedon määrää ja mahdollisia ongelmakohtia. Luotettavaa kuvaa Oodi-tietokannan tietojen kattavuudesta, oikeel-

lisuudesta ja eheydestä ei lyhyessä esiselvityksessä pysty antamaan. Joitakin ongelmakohtia pystyi jo nyt havainnoimaan, joista myöhemmin kohdassa ”Havaittuja ongelmakohtia”.

Tietojen eheyttä, kattavuutta ja oikeellisuutta on tutkittava mieluiten ohjelmallisesti. Huonokuntoisen tiedon kunnostamiseen alkuperäisessä lähteessä, tietojen korjaamiseen ja tietojen hylkäämiseen on otettava kantaa mahdollisimman aikaisessa vaiheessa, koska hylkäämistä lukuun ottamatta kaikki muut vaihtoehdot ovat varsin työläitä (JHS suositus 172).

3 Konversion laatutavoitteita ja havaittuja ongelmakohtia

Konversion keskeisten laatutavoitteiden määrittelyssä huomioidaan seuraavat tekijät:

- tiedon muuttumattomuus lähteen ja kohteen välillä
- tiedon jäljitettävyyden konversioissa, kun tietoa siirretään useista lähteistä
- uuden tietokantaratkaisun asettamat vaatimukset
- suurien tietomassojen konversion tehokkuus eli mm. läpimenoaikojen arviointi
- läpimenoaikojen rajat tuotantokatkon takia.

Mikäli konversio on tarkoitus kilpailuttaa, kuvausten laatu ja ajantasaisuus sekä realistinen arvio tietosisällön laadusta ovat tarpeen toimittajakandidaateille mahdollisimman oikean työmääräarvion laatimisessa.

Konversion tarkempi suunnitelma ja kuvaus kannattaa tehdä siinä vaiheessa, kun mallinnukset tulevasta toteutusvaihtoehdosta ja tietokannasta ovat käytettävissä. Tarkempaa konversiosuunnitelmaa varten on käytävä läpi yksityiskohtaisesti Oodi-tietokannan dokumentaatio, mm. tietoturva-arkkitehtuuri, tietoarkkitehtuuri ja kuvaukset tietokannoista. Mikäli jonkin osa-alueen kuvauksia puuttuu, voidaan ne tehdä esimerkiksi takaisin mallinnuksella eli reverse engineering -tekniikalla sopivaa apuohjelmaa käyttäen (JHS suositus 172). Oodi-tietokannan taulujen määrä tuo vaikeuksia reverse engineering -tekniikan käyttämiseen. Oletettavasti tästä syystä Oodi-tietokantatauluista ei ole saatavana kattavaa graafista kuvausta.

Tietokannan ja sen tietosisällön osaaminen

Yliopistoilla, CSC:llä ja Oodin toimittaja Logicalla on vahvaa teknistä ja Oodin sisältöpuoleen liittyvää osaamista. Samoin Oodi-tietokannan ja siihen liittyvien rajapintojen dokumentaatiota on käytettävissä (CSC:n Confulence Wiki). Koska asiantuntijoita on niin substanssi- kuin toteutuspuolelta käytettävissä ja dokumentaatiota on saatavilla, helpottaa se jatkossa konversion tarkempaa suunnittelua, määrittelyä ja toteutusta.

Havaittuja ongelmakohtia

Tietokantataulujen määrä tuotantokannassa on todella suuri (> 1800 kpl). Kaikki taulut on käytävä huolellisesti läpi, jotta löydetään todellinen konvertoitava data ja sen sijaintitaulut.

Tauluissa on runsaasti attribuutteja (sarakkeita), joiden sisältö ei välttämättä avaudu muille kuin Oodi-tietokantaan ennestään tunteville tai dokumenttien kautta. Tämä on varsin tyypillinen tilanne, joihin konversioiden suunnittelijat ja toteuttajat joutuvat. Muutamia esimerkkejä:

- opinto taulu: GENOPIN, ECTSLAAJ, PAINKERRKA, SUUNSUORLK
- opinoikkohd taulu: KOHTTASNRO
- opinoik taulu: KOTOPPLAIT, LKMAKSKOK
- opintap taulu: TOIVOPET, VIETILMTAUL, LK

Mikael Gueck sivuaa Oodi-selvitysraportissaan asiaa seuraavalla tekstillä: ”Oodin tietokannassa on yli 580 tietokantataulua. Taulujen nimistä ei käy ilmi niiden tarkoitus, eikä substanssitason faktojen suhde tietokantatauluihin ja niiden faktoihin käy ilmi mistään näkemästäni dokumentista. Tietokantataulut on dokumentoitava sillä tarkkuudella, että jokaisen taulun esittämän faktan merkitys sovellukselle ja

vielä merkittävämmiin sovelluksen takana olevalle substanssille, on käytävä ilmi. Lisäksi jokaisen taulun suhde toisiin tauluihin, ja suhteiden merkityksen niille faktoille joita suhteet kuvaavat, on oltava selvä” (Mikael Gueck, Oodi teknologiakonsultointi 2007).

Lisäksi eräissä tietokantataulujen sarakkeissa data saattaa olla erilaista kuin mitä sarakkeen nimi antaisi ymmärtää. Tässä on myös yliopistokohtaisia variaatioita.

Tietokannan tiedonkäsittelymekanismit pitää myös tarkastaa ja huomioida konversion sääntötietoja määriteltäessä. Esimerkiksi tietokannan PL/SQL proseduurit, taulujen triggerit (herättimet), tietokannan sekvenssi-rakenteet yms. on käytävä huolellisesti läpi. Työtä vaikeuttaa se, että käsittelysääntöjä ja sääntötietoja on ilmeisesti sekaisin PL/SQL, Java ja Uniface-koodissa. Kaiken kaikkiaan konversiosääntöjen laatiminen on hyvin työläs ja aikaa vievä vaihe.

Johtopäätös:ärkevintä olisi teettää konversio tietokannan tietojen ulosluvun osalta järjestelmän toimittajalla, Logically, jolla on valmiina tietokannan ja sen sisällön tuntemusta ja osaamista. Sopimusteknisesti toimittaja pitää sitouttaa ja motivoida konversion tekemiseen. Oodi-tietokannan dokumentaatiota joudutaan mahdollisesti myös tarkentamaan ennen konversion toteutusta tietyiltä osilta, kuten Gueck selvitysraportissaan toteaa. Oodi-tietokannan tietojen latauksen uuteen tietokantaratkaisuun voi puolestaan teettää uuden ”OTM-järjestelmän” toimittajalla. Järjestelmätoimittajan lisäksi konversiossa tarvitaan huomattava määrä yliopistojen omaa työtä sekä Oodi-konsortion kehitysyksikön asiantuntemusta.

Muutaman esimerkinomainen Oodi-tietokantataulun kuvaus:

opinto	Tyhjä	arvo	Tyyppi
Nimi			
OPINTO	NOT NULL		NUMBER (12)
HLO	NOT NULL		NUMBER (12)
OPINKOHD	NOT NULL		NUMBER (12)
OPINOIK			NUMBER (12)
ENNPAAAT			NUMBER (12)
ORGANISAATIO			NUMBER (12)
OPINSTAT	NOT NULL		NUMBER (2)
LAAJUUS	NOT NULL		NUMBER (5, 2)
GENOPIN	NOT NULL		NUMBER (1)
KIELI	NOT NULL		NUMBER (4)
LASKSUOROTT	NOT NULL		NUMBER (1)
OPINKOHTYYYP	NOT NULL		NUMBER (2)
KAYTMUUT	NOT NULL		NUMBER (1)
TAPPVM	NOT NULL		DATE
SUORTYYP	NOT NULL		NUMBER (2)
SUORPVM			DATE
KORPVM			DATE
KIRJPVM			DATE
HYVPVM			DATE
PISTMAAR			NUMBER (5, 2)
LASKPISTMAAR			NUMBER (5, 2)
ECTSLAAJ			NUMBER (5, 2)
ILMTAULPVM			DATE
ENNALKPVM			DATE
ENNPAAATPVM			DATE
SUORULKMAA			NUMBER (1)
TULSUOROTT			NUMBER (1)
VIIMVOIMPVM			DATE
SUORMUU			NUMBER (1)
PROJEKTI			VARCHAR2 (10)
AVAAJA	NOT NULL		NUMBER (12)
AVAUPVM	NOT NULL		DATE
MUUTTAJA	NOT NULL		NUMBER (12)
MUUTPVM	NOT NULL		DATE
OPISAIKOPIN			NUMBER (12)
PAINKERRKA	NOT NULL		NUMBER (5, 2)
ERILPAAT			NUMBER (1)
SUUNSUORLK			NUMBER (3)
OPISPALHIST			NUMBER (1)
LAAJOP	NOT NULL		NUMBER (5, 2)

ALKPERLAAJ	NOT NULL NUMBER(1)
opinoik	
Nimi	Tyhjä arvo Tyyppi
OPINOIK	NOT NULL NUMBER(12)
HLO	NOT NULL NUMBER(12)
KOTOPPLAIT	NUMBER(12)
MYONTAJA	NOT NULL NUMBER(12)
PERUUTTAJA	NUMBER(12)
MYONPVM	NOT NULL DATE
ALKPVM	NOT NULL DATE
PAATPVM	NOT NULL DATE
OPINOIKKATT	NOT NULL NUMBER(2)
KIINTIO	NUMBER(4)
PERUSTE	NOT NULL NUMBER(2)
VALTAP	NUMBER(2)
VAIHOHJ	VARCHAR2(10)
OPINALOIPVM	DATE
PERUSTPVM	DATE
PERPVM	DATE
OPINLUON	NUMBER(2)
OPINTARK	NOT NULL NUMBER(5)
OPINOIKASEM	NUMBER(2)
VUOSKURS	VARCHAR2(6)
MAKSMAAR1	NUMBER(12,2)
MAKSVAL1	VARCHAR2(3)
MAKSMAAR2	NUMBER(12,2)
MAKSVAL2	VARCHAR2(3)
AVAAJA	NOT NULL NUMBER(12)
AVAUPVM	NOT NULL DATE
MUUTTAJA	NOT NULL NUMBER(12)
MUUTPVM	NOT NULL DATE
OPISORG	NOT NULL NUMBER(12)
VALMISTUNUT	NUMBER(1)
OIKPERKOOD	NUMBER(2)
LASKPAATPVM	DATE
LKMAKSOK	NUMBER(1)
MAKSTILKOUL	NUMBER(1)

4 Konversioon toteutusvaihtoehtoja

1) Kertakonversio – ”Big Bang”

- Perinteinen lähestymistapa
- Vastaanottavan pään tietokantaratkaisun ja tietomallin oltava valmis
- Tässä toteutustavassa voitaisiin hyödyntää esim. korkeakoulujen tietomallia (=XDW-malli) siirtorajapintana niiltä osin kuin XDW-määrittelyt on toteutettu
- XDW-mallista puuttuvien kenttien sisältö voidaan viedä erillisen konversiorajapinnan kautta tai vaihtoehtoisesti suoraan uuteen kohdekantaan
- Siirrettyihin tietoihin ei tarvitse enää palata eikä vanhan tietokannan tietoja tai konversiosääntöjä enää tarvitse enää ylläpitää.

Johtopäätöksiä:

- ⇒ Konversiossa kannattaa vakavasti harkita XDW-rajapinnan hyödyntäminen konversioon väli-rastona. Tästä olisi paljon etuja; uuden tietokantaratkaisun malli saataisiin sovitettua samalla kohtuullisen vaivattomasti korkeakoulujen tietomalliin eli XDW-malliin niiltä osin kuin XDW-määrittelyt on toteutettu.
- ⇒ Toinen vaihtoehto on tehdä konversiosierrot mahdollisimman yksinkertaisesti ja suoraviivaisesti vanhan ja uuden ratkaisun välillä. Sovitus korkeakoulujen tietomalliin (XDW- malli) tehtäisiin myöhemmin erikseen.

Konversiorajapinnan käyttöä harkittaessa, kannattaa huomioida, että Aalto-yliopistolla on DW tietovarasto käytössä tietojen tietue-asteelle asti. Aallon tietovarasto perustuu XDW-malliin. Siellä on valmiina ja testattuna opiskelijat, kaikki opinto-oikeuksiin liittyvät tiedot, opintosuoritukset, suoritettut tutkinnot. Työn alla ovat HOPS-tiedot, tutkintorakenteet ja opetustapahtumat. Aallon DW-työ ostetaan kokonaan

ulkupuolelta. Aallolla on tällä hetkellä oodin teknisessä ylläpidossa kaksi henkilöä, jotka tekevät lisäksi myös muita kuin Oodiin liittyviä tehtäviä.

2) Vaiheittainen käyttöönotto, vanhan Oodi-järjestelmän ajaminen uuden rinnalla

- Tietojen ryhmitysesimerkki vaiheittaisessa käyttöönotossa:
 - ✓ Opiskelijat, opinnot, opinto-oikeudet. Osiot voidaan toteuttaa eri tekniikoilla ja tarvittaessa erilaisiin kohdetietokantoihin.
 - ✓ Kommunikointi vanhan Oodi-järjestelmän ja uuden OTM-järjestelmän välillä toteutetaan esim. SOAP (Simple Object Access Protocol) ja WSDL (Web Service Description Language) tekniikoilla.

Johtopäätöksiä:

- ⇒ Vaiheittainen käyttöönotto edellyttää paljon uusia ja huolellisesti testattuja rajapintaratkaisuja Oodin ja uuden OTM-järjestelmän välille
- ⇒ Oodin / OpeOodi SOAP protokollan mukaisten web services rajapintojen toteutus oli aikanaan hyvin työlästä; alkuperäiset työmääräarviot ylittyivät ja toteutuksen testaus oli hankalaa. Rajapintojen testaaminen on ylipäätään varsin työlästä ja hankalaa.
- ⇒ Edellyttäisi, että analysoitaisiin huolellisesti nykyisten palvelurajapintojen toiminnallisuus ja käyttökelpoisuus (OpeOodin web services rajapintojen dokumentit, <https://confluence.csc.fi>)
- ⇒ Sisältää paljon riskejä: sovittava selkeästi esimerkiksi mitä päivitetään uudessa OTM-järjestelmässä, mitä vanhan Oodin kautta, tietojen yhdenmukaisuuden varmistaminen – on varmasti ongelmallista ja hankalaa.
- ⇒ Miten toteutetaan tilastointi ja muu raportointi kahden järjestelmän käytössä?
- ⇒ Integrointi sitoisi paljon kehitys- ja testausresursseja sekä rahaa uuden järjestelmän kehittämisestä, jotta vanhan Oodin kanssa kommunikointi onnistuisi
- ⇒ Realistisia työmääriä on erittäin vaikea arvioida, sen takia tässä esiselvityksessä niitä ei annettu.

Muita havaintoja

Muistiinpanoja eräistä kiinnostavista vaihtoehtoista ja mahdollisuuksista, joita konversioesiselvityksessä tuli esille.

Muita vaihtoehtoja perinteiselle relaatiotietokannalle ovat esimerkiksi:

- Neo4j, OrientDB, Kyoto Cabinet, MongoDB, CouchDB
- Näistä uusista tietokantaratkaisuista tosin ei Suomessa ole paljoa kokemuksia
- Maailmalla näitä on käytössä eräissä isoissa ratkaisuihin, mm. sosiaalinen media, pankkiiriliikkeet ja sähköinen kaupankäynti
- Näiden vaihtoehtojen edut ja haitat kannattaa selvittää tarkemmin, jos sellainen vaihtoehto on tarjolla jossain valmisratkaisussa (esiselvityksen ratkaisuvaihtoehdot B ja C) tai päädytään omaan in house -kehittämiseen (esiselvityksen ratkaisuvaihtoehto A)
- Konversion kannalta tuo jonkin verran uuden opiskelua perinteisen ja tutun relaatiomallin sijaan
- Nykyinen relaatiomalli ei ole ideaalinen opintohallinnon tietojen tallennukseen, koska puurakenteista tietoa on paljon. Opintohierarkiat, tutkintorakennepuut yms. voitaisiin toteuttaa graafitietomalleina, jolloin niiden käsittely olisi huomattavasti nykyistä relaatiomallia helpompaa.

Arkiston käyttöönotto – nyt siihen on hyvä tilaisuus

- Konvertoidaan Oodista vain aktiivitetot, esim. uudet ja aktiiviset opiskelijat ja opinto-oikeudet siirretään konversiossa uuteen tietokantaan
- Vanha Oodi jätetään arkistoksi tai konvertoidaan vanha, passiivinen aineisto erilliseen arkistokantaan. Samoin kannattaa miettiä jätetäänkö esim. sähköiset opinto-oppaat ja tutkintora-

kenteet arkistotietokantaan. Ne voidaan ehkä kätevimmin luoda uuteen OTM-järjestelmään ja tietokantaan manuaalisesti uudistamistyön yhteydessä.

- Virkailijakäyttöä varten vanha Oodi tietoineen olisi arkistona johon olisi pienehkö lisenssimäärä.
- Tässä pitäisi todennäköisesti samalla rakentaa jokin kätevä tapa aktivoida tätä passiivista aineistoa tarvittaessa aktiiviseen tietokantaan (esimerkiksi jatko-opiskelijat ym.). Aktivointitoiminnallisuuden rakentaminen on lisäkustannus.
- Kahden järjestelmän rinnakkaiskäyttö edellyttää myös konversiosäätöjen ylläpitoa, mikä tuo lisätyötä ja lisäkustannuksia.
- Arkistoratkaisun avulla saadaan käsiteltävän datan määrä kohtuullisen pieneksi.

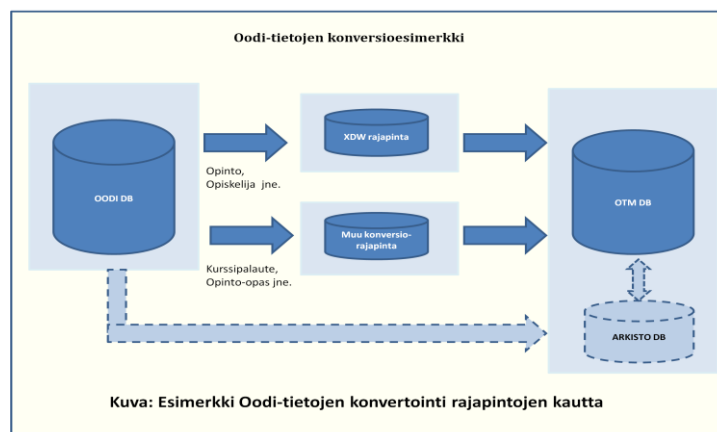
Erillisen arkistotietokannan käyttö ei ole aivan ongelmaton ja suoraviivaista. Historiakannasta on maininta myös M. Gueckin selvityksessä, jossa ansiokkaasti kuvataan mm. raportointiin mahdollisesti syntyviä ongelmia erillisen arkistotietokannan yhteydessä:

”Säilytetään nykyinen Oodi historiallisen tiedon lähteenä, ja luodaan kokonaan uusi oppilastietojärjestelmä Oodin kokemusten pohjalta.”

”Tämä vaihtoehto ottaa huomioon olemassa olevan tiedon määrän, ja rajoittaa uuteen järjestelmään siirrettävän historiallisen tiedon monimutkaisuutta vain niiden prosessien faktoihin, jotka ovat auki sillä hetkellä kun siirrytään vanhasta järjestelmästä uuteen. Tällöin pitää keksiä jokin ratkaisu sellaisten raporttien ajamiseen, jotka käsittelevät faktoja joita on sekä uudessa että vanhassa kannassa. Variantti tästä vaihtoehdosta on se mahdollisuus, että selviää, että vanhassa tietokannassa on niin paljon tarpeetonta ja yksinkertaisilla loogisilla operaatioilla hävitettävää redundanssia, että vanhan 585 tietokantataulun tietokannan olennaiset faktat saadaan siirrettyä täysin eheinä uuteen alle 100 taulun ylläpidettävään tietokantaan.” (Mikael Gueck, Oodi teknologiakonsultointi 2007).

Esimerkkejä konversiosuunnittelusta

Tietojen konvertointi rajapintojen kautta:



Esimerkki tietokantataulujen konversiosuunnittelusta:

OodiDB	XDW	OTMDB	Huomioitavaa
opinto	XDW < taulu 1 >	OTM- taulu 1	Rivejä yli 8 milj.
opinoikkohd	XDW < taulu 2 >	OTM- taulu 2	Rivejä yli 2 milj.
opinoik	XDW < taulu 3 >	OTM- taulu 3	Viite-ehyedet tarkistettava huolellisesti
opintap	XDW < taulu 4 >	OTM- taulu 4	Rivejä yli 2 milj.
opinarvsan	XDW < taulu n >	OTM- taulu n	Rivejä yli 10 milj.

K_KAYTJARJ	Ei konvertoida	Ei konvertoida	Aputaulu, taulun tietoja ei konvertoida
Loki	Ei konvertoida	Ei konvertoida	Lokitaulu, taulun tietoja ei konvertoida

Tietokantataulujen attribuutit konversiosuunnitelmassa:

OodiDB	XDW	OTMDB	Huomioitavaa
opinto	XDW <opinto>	OTM-opinto	
OPINTO	< XDW Schema, kenttä>	OTMDB_Study	Pakollinen arvo
HLO	< XDW Schema, kenttä>	OTMDB_Person	Pakollinen arvo
OPINKOHD	< XDW Schema, kenttä>	OTMDB_Trgetstd	Pakollinen arvo
...jne	...jne	...jne	
LAAJOP	< XDW Schema, kenttä>	OTMDB_	
ALKPERLAAJ	< XDW Schema, kenttä>	OTMDB_	
opinoikkohd	XDW <opinoikkohd>	OTM-opinoikkohde	
OPINOIK			Pakollinen arvo
OPINKOHD			Pakollinen arvo
ALKPVM			Pakollinen arvo
...jne	...jne	...jne	
MUUTTAJA			Trigger
MUUTPVM			SYSDATE arvo

5 Konversion työvaiheet ja aikataulu karkealla tasolla

Esimerkinomainen lista konversion työvaiheista:

0. Konversion projektointi ja projektisuunnitelma
1. Tietokannan dokumentaatio, läpikäynti ja tarkastus
2. Tietokannan kuntoarvio
3. Aikataulusuunnittelun tarkennus
4. Tekninen suunnittelu ja dokumentointi
5. Siirto- ja apuohjelmien rakentaminen
6. Testaus, sekä konversio että apuohjelmat
7. Tietojen tarkastus testiympäristössä
8. Konversion suorituskyvyn virittäminen
9. Konversion suoritus
10. Tietojen tarkastus, mahdollinen korjaus ja "puhdistaminen"

6 Kustannus- ja työmääräarvio

Seuraavassa on hieman taustatietoa ja tunnuslukuja Oodin käyttöönottoon liittyneistä konversion työmääristä. Konversio oli tyypiltään kertakonversio. Lähde Pauli Saarisen laatima "OODI JÄRJESTELMÄN KÄYTTÖÖNOTTOPROJEKTIN LOPPURAPORTTI 25.2.2000".

Eräitä keskeisiä kohtia aikataulusta ja työvaiheista:

- Konversioiden suunnittelu käynnistyi 20.4.1999.
- Konversion pohjana oli konsortion tuottama yhteinen käyttöönottoprojekti. Osa vastaanottavan pääkonversiosta oli suunniteltu ja toteutettu siinä.
- Ensimmäiset versiot konversioiden sääntötiedostoista olivat valmiina syyskuun 1999 alkuun mennessä.
- Syys- lokakuussa aloitettiin konversioiden testaus.
- Marraskuun lopussa todettiin, että opintosuoritusten konversioiden toteutus ja testaus oli vielä pahasti kesken.
- Varsinaiset konversiot tuotantoympäristöön käynnistyivät joulukuun 1999 alussa.
- Tiedot ladattiin ensin konversiotietokantaan ja sieltä tietokantadumpina tulevaan tuotantotietokantaan.

Monien ongelmien jälkeen konversio saatiin tehtyä 28.12.1999 mennessä. Testaus ja tietojen tarkistus siirtyi tuotantokantaan.

Konversioon arvioidut ja käytetyt työmäärät:

- Konversioiden, sääntötietojen ja tietokannan tietosisällön suunnittelun arvioitiin alun perin vievän 20 henkilötyöpäivää; toteutuma oli 96 henkilötyöpäivää.
- Lisäksi testaus ja koekäyttö veivät 182 henkilötyöpäivää, josta merkittävä osa oli konversion ja tietosisällön testausta. Alkuperäisessä suunnitelmassa tähän arvioitiin 15 henkilötyöpäivää.
- Kokonaistyömäärä oli noin 300 henkilötyöpäivää.

Helsingin kauppakorkeakoulun vanhan opintohallinnon järjestelmän (KKJ) tietojen konvertointi ja siivous uuteen Oodi-tietokantaan vei lähes 1000 henkilötyöpäivää (M. Eklin).

- Toimittajalta ostettua työtä oli noin 120 henkilötyöpäivää
- Viisi omaa työntekijää teki konversioon liittyviä tehtäviä noin vuoden ajan, nettotyöaika yhteensä oli noin 800 henkilötyöpäivää
- Suurin työmäärä kertyi tietojen sovittamisesta Oodin-tietokantaan sekä lisätietojen täydentämisen sääntöjen suunnittelusta ja määrittelystä. Lisäksi datan "puhdistaminen" oli hyvin työläs vaihe.

Johtopäätöksiä:

- ⇒ Kaiken kaikkiaan siirrettävän tiedon määrä on nyt aivan toista luokkaa kuin vuoden 1999 kertakonversiossa. Koska tietoja pitää tarkastaa ja korjailla, voi toteuma työmäärien osalta olla hyvinkin lähes 1000 henkilötyöpäivää per Oodi-yliopisto.
- ⇒ Yliopistoille yhteisiä työvaiheita on suhteellisen vähän. Työvaiheista suuritöisimmät ovat: kunkin Oodi-tietokannan kuntoarvio, konvertoitavan datan valinta, konversiosääntöjen määrittely, konversion tietojen tarkastus, mahdollinen korjaus ja "puhdistaminen" – nämä työvaiheet pitää tehdä pääsääntöisesti jokaisessa yliopistossa erikseen.
- ⇒ Ulkoa ostettavaa työtä voisi olla esimerkiksi siirto- ja apuohjelmien rakentaminen, konversioajojen suoritus ja avustaminen tietokannan kuntoarviossa. Huomattava osa konversioon liittyvästä työstä (~80 %) on kuitenkin yliopistojen omaa työtä – tietojen tarkastamista, mahdollisia korjauksia ja "puhdistuksia" ennen ja jälkeen konversion.
- ⇒ Jos ulkoa ostettava työmäärä olisi ~15 % – 25 % eli noin 150 – 250 henkilötyöpäivää, niin sen kustannus olisi noin 100 000 € - 200 000 € toimittajasta riippuen.