

Jaakko Ketelimäki

Ulkoasiainministeriön PX-Web-tilastojulkaisun käyttöönotto ja automatisointi

Kehitysyhteistyön julkaisupalvelu

Opinnäytetyö

Syksy 2016

SeAMK Tekniikka

Tietotekniikan koulutusohjelma

SeAMK 

SEINÄJOEN AMMATTIKORKEAKOULU
SEINÄJOKI UNIVERSITY OF APPLIED SCIENCES

SEINÄJOEN AMMATTIKORKEAKOULU

Opinnäytetyön tiivistelmä

Koulutusyksikkö: Tekniikan yksikkö

Tutkinto-ohjelma: Tietotekniikan koulutusohjelma

Suuntautumisvaihtoehto: Tietoverkkotekniikka

Tekijä: Jaakko Ketelimäki

Työn nimi: Ulkoasiainministeriön PX-Web-tilastojulkaisun käyttöönotto ja automatisointi

Ohjaaja: Markku Lahti

Vuosi: 2016

Sivumäärä: 35

Opinnäytetyössä esitellään PX-Web-tilastojulkaisun käyttöönotto Suomen Ulkoasiainministeriön tilastojulkaisupalveluna ja sen automatisoimisesta. Vuonna 2013 Suomen Ulkoasiainministeriössä otettiin käyttöön PX-Web-tilastojulkaisu osana URKKI-tietovarastoprojektia. Tilastojulkaisu saatiin onnistuneesti toimimaan automaattisesti.

Opinnäytetyö alkaa perehtymällä yleisesti tietovarastointiin ja sen yhteydessä toimiin järjestelmiin ja prosesseihin. Tämän jälkeen opinnäytetyössä perehdytään PX-Web-palveluun ja sen toimintaan saattamisesta automaattisesti Ulkoasiainministeriön tietovarastoa hyödyntäen. Automatisointi toteutettiin tuottamalla data palveluun ETL-prosessilla Ulkoasiainministeriön URKKI-tietovarastosta. Lopuksi opinnäytetyössä vertaillaan automaattista ja manuaalista tilastojulkaisua sekä kerrotaan yhteenveto käyttöönotosta.

Avainsanat: tilastojulkaisu, PX-Web, tietovarastointi, Ulkoasiainministeriö, automatisointi

SEINÄJOKI UNIVERSITY OF APPLIED SCIENCES

Thesis abstract

Faculty: School of Technology

Degree programme: Information Technology

Specialization: Networking Technologies

Author: Jaakko Ketelimäki

Title of thesis: Implementation and automatization of the PX-Web statistics publishing service of the Finnish Ministry of Foreign Affairs

Supervisor: Markku Lahti

Year: 2016

Number of pages: 35

The thesis introduces the implementation of PX-Web as the statistics publishing service of the Finnish Ministry of Foreign Affairs and the automatization of the service. In 2013 PX-Web statistics publishing service was implemented as a part of the foreign ministry's URKKI-data warehouse project. The statistics publishing service was successfully made to work automatically.

The thesis begins by presenting data warehousing in general, and the systems and processes related to it. Automatization was executed by producing data for the service with the ETL-process from the foreign ministry's URKKI-data warehouse. After that the thesis showcases the implementation of the PX-Web service and the automatization of the service by utilizing the data warehouse of the Finnish Ministry of Foreign Affairs. Finally, the thesis compares the automatic and manual publishing of statistics and summarizes the implementation.

Keywords: statistics publishing, PX-Web, Ministry of Foreign affairs, automatization

SISÄLTÖ

Opinnäytetyön tiivistelmä.....	1
Thesis abstract.....	2
SISÄLTÖ.....	3
Kuva- ja kuvioluettelo	5
Käytetyt termit ja lyhenteet	6
1 JOHDANTO	8
1.1 Työn tausta	8
1.2 Työn tavoite	8
1.3 Työn rakenne	9
1.4 Toimeksiantaja.....	9
2 TIETOVARASTOINTI.....	11
2.1 Tietoresurssi	11
2.2 Operatiiviset järjestelmät.....	11
2.3 Tiedon varastointi tietovarastoon	12
2.4 ETL-prosessi.....	13
2.5 Tietovarastoinnin historiaa	14
3 URKKI-TIETOVARASTO	16
3.1 Tietovaraston lähdejärjestelmät	16
3.2 Tietovarasto ja ETL-prosessit	17
3.3 Raportointi ja Business Intelligence	18
3.4 Käyttöliittymät.....	18
4 IBM INFOSPHERE DATASTAGE	19
4.1 Päätoiminnot.....	19
4.2 Datan muuntaminen.....	20
4.3 Vaiheittaiset tehtävät.....	21
5 PX-WEB.....	22
5.1 Järjestelmän kuvaus	22
5.2 Sovellusvirtaus.....	24
5.3 PX-tiedostojen kuvaus	25

6	AUTOMAATTINEN PX-TIEDOSTOJEN PÄIVITYSPROSESSI JA PALVELUN KÄYTTÖ	27
7	AUTOMAATTISEN JA MANUAALISEN PX-WEB- TILASTOJULKAISUN VERTAILU	32
7.1	Manuaalinen tilastojulkaisu	32
7.2	Automaattinen tilastojulkaisu	32
7.3	Ulkoasiainministeriön automaattinen tilastojulkaiseminen	33
8	YHTEENVETO.....	34
	LÄHTEET	35

Kuva- ja kuvioluettelo

Kuva 1: PX-Web-palvelun aloitussivu	28
Kuva 2: PX-Web-palvelun PX-tiedostojen valintasivu	29
Kuva 3: PX-Web-palvelun PX-tiedostojen muuttujien valintasivu.....	30
Kuva 4: PX-Web-palvelun PX-tiedoston raporttisivu	31
Kuvio 1: Esimerkki tietovarastointiprosessista	12
Kuvio 2: URKKI-järjestelmäarkkitehtuuri 27.2.2015	16
Kuvio 3: PX-Web – järjestelmäkuvaus	23
Kuvio 4: PX-Web - sovellusvirtaus	24
Kuvio 5: PX-tiedostojen automaattinen luominen ja päivitys.....	25

Käytetyt termit ja lyhenteet

BI	Business Intelligence, liiketoimintatiedon hallinta.
Data mart	Data mart on tietovaraston siirtorajapinta, jota tietovarasto käyttää datan siirtämiseen käytettäväksi palveluihin.
Data Vault	Data Vault on tietokanta, jonne kerätään operatiivisista järjestelmistä muokkaamaton raakadata, josta tietovarasto kerää tarkkaa dataa.
DDoS	Distributed Denial of Service, hajautettu palvelunestohyökkäys.
DW	Data Warehousing, tietovarastointi.
CSW-tiedosto	Comma-separated values. Tiedostomuoto, joka sisältää yksinkertaista taulukkomuotoista tietoa tekstimuodossa.
ETL	Extract, Transfer ja Load. Lyhenteellä tarkoitetaan datavara- rastojen yhteydessä tapahtuvaa prosessia datan käsitte- lyä, siirtoa ja tallentamista varten.
ERP	Enterprise Resource Planning, operatiivinen järjestelmä eli toiminnanohjausjärjestelmä.
Excel	Microsoft Excel-tilukko-ohjelman käyttämä tiedosto- muoto.
Maksatukset	Maksatuksilla tarkoitetaan kehitysyhteistyöhön käytettyjä varoja.
Metadatan	Metadatan on dataa, joka ohjaa dataa. Metadatan rakentuu avainsanoista, joiden mukaan määritellään datan käyttäy- tymistä.
Myönnöt	Kehitysyhteistyöhön myönnetyt varat, joiden käyttö voi ja- kautua useamman vuoden ajalle.

Prompti	Kehote tai kysely.
PX-tiedosto	PX-Web-tilastojulkaisussa käytettävä tiedostotyyppi.
PX-Web	Tilastojulkaisuun suunniteltu web-palvelu.
Skripti	Automaattinen ohjelma, jolle voidaan määrittää ajastettu toiminnallisuus.
Sovellusvirtaus	Järjestys/ohjeet, jossa esitetään sovelluksen käytön kulku.
Tilastojulkaisu	Tilastollisen tiedon tai datan julkaisu.
UM	Suomen Ulkoasiainministeriö, ulkoministeriö.

1 JOHDANTO

1.1 Työn tausta

IATI (International Aid Transparency Initiative) on 2008 alkanut maailmanlaajuinen kehitysyhteistyön kampanja, jonka tarkoituksena on kehittää kehitysyhteistyön laatua tekemällä valtioiden ja organisaatioiden kehitysyhteistyöstä läpinäkyvämpää. IATI pisteyttää vuosittain kehitysyhteistyön maksajat erilaisilla läpinäkyvyysmittareilla ja julkaisee, kuinka eri valtiot ja organisaatiot ovat sijoittuneet pisteytyksessä ja vertailee kehitysyhteistyön laatua. (IATI, 2016a; IATI, 2016b.)

Suomella on ollut heikko menestys IATI:n vertailuissa, koska Suomen valtion maksamista kehitysyhteistyövaroista ei ole ollut sähköistä, reaaliaikaista raportointia eikä julkaisupalvelua käytössä aikaisemmin.

Vuonna 2012 kehitysministeri Heidi Hautalan aloitteesta Ulkoministeriössä alkoi projekti, jonka tarkoituksena oli kehitysyhteistyön käytettävien varojen sähköinen raportointi ja julkaiseminen julkisena palveluna Ulkoministeriön verkkosivuilla. Julkaisuun päätettiin käyttää PX-Web-tilastojulkaisupalvelua. Palvelu julkaistiin 2013 lopussa.

Vuonna 2015 Ulkoasiainministeriön tietotekniikan yksikössä päätettiin laajentaa käyttöä Ulkoministeriössä. Tavoitteena oli saada palvelu julkaisemaan dynaamisesti Ulkoministeriön taloushallinta- ja asianhallintajärjestelmissä jo olevaa ja sinne tuotettavaa dataa.

1.2 Työn tavoite

Opinnäytetyön tavoitteena on selvittää, miten PX-Web-verkkotilastopalvelu toimii osana Ulkoasiainministeriön tietovaraston automaattista raportointia ja kuinka se on toteutettu. Opinnäytetyössä selvitetään myös, miten PX-Web toimii IBM InfoSphere DataStage-ETL-ohjelmiston avulla automaattisesti päivittyvänä tilastojulkaisuna. Tavoitteena on myös pohtia eroa tavallisen, manuaalisesti tehtävän ja automaattisen tilastojulkaisun eroja, haasteita, etuja ja haittoja.

1.3 Työn rakenne

Luku 2 alkaa tietovarastoinnin esittelyllä. Luvussa esitellään tietovaraston toimintaperiaate ja tietovaraston historiaa.

Luvussa 3 esitellään Suomen Ulkoasiainministeriön URKKI-tietovarasto. Samassa luvussa esitellään tietovaraston lähdejärjestelmät, joista tulevaa dataa PX-Webissä käytetään ja julkaistaan.

Luvussa 4 kerrotaan IBM InfoSphere DataStage-ETL-ohjelmistosta ja sen toimintaperiaatteista.

Luvussa 5 esitellään PX-Web ja sen toimintaperiaatteet. Samassa luvussa esitellään myös tiedostotyypit, jota PX-Web-palvelu käyttää.

Luvussa 6 esitellään prosessi PX-tiedostojen automaattisesta päivittämisestä sekä esimerkki PX-Web-tilastopalvelun toiminnasta.

Luvussa 7 vertaillaan automaattista ja käsin tehtävää tilastojulkaisua ja tuodaan esille molempien tapojen edut ja haitat. Luvussa käsitellään Ulkoasiainministeriön automaattisen tilastojulkaisun valintaa manuaalisesti tehtävän sijasta.

Luku 8 on yhteenveto työstä ja se sisältää myös pohdintaa automaattisen ja käsin tehtävän tilastojulkaisun eroavaisuuksia.

1.4 Toimeksiantaja

Suomen Ulkoasiainministeriö vastaa virallisista Suomen valtion ulkoasioista. Sen vastuualueita on muun muassa Suomen edustaminen ulkomailla ja muiden maiden edustustot Suomessa, matkustus- ja vientiluvat, matkustus- ja vientivalvonta sekä kehitysyhteistyö. Toimeksiannosta Ulkoasiainministeriössä vastasivat tiedon- ja palvelunhallintayksikkö sekä kehitysyhteistyön hallinto- ja oikeusyksikkö. (Suomen Ulkoasiainministeriö, 2016.)

Sisällöstä tietovarastossa ja PX-Web-palvelussa vastaa kehitysyhteistyön hallinto- ja oikeusyksikkö, teknisestä toteutuksesta ja hallinnoinnista tiedon- ja palvelunhallintayksikkö.

Tiedon- ja palvelunhallintayksikön tehtäviin kuuluu ulkoasiainhallinnon tietoarkkitehtuurikokonaisuuden hallinnointi ja tietoteknologiaa hyödyntävien projektien ja hankkeiden hallinnointi sekä toteuttaminen.

Kehitysyhteistyön hallinto- ja oikeusyksikön tehtäviin kuuluu:

- Kehitysyhteistyön toiminta- ja taloussuunnittelu.
- Kehitysyhteistyövarojen taloushallinto, sisältäen määräraha- ja valtuusseurannan sekä määrärahapäätösten valmistelun; kehitysyhteistyön taloushallinnon kehittäminen.
- Kehitysyhteistyön tietojärjestelmien ylläpito ja kehittäminen kehityspoliittisen osaston toimivaltaan liittyvin osin.
- Kehityspoliitiikan ja -yhteistyön suunnitteluun ja toteutukseen liittyvä oikeudellinen neuvonta. Osaston tehtäviin kuuluu kehitysyhteistyöhön liittyvä säädösvalmistelu. Hankintatoimen koordinointi sekä ministeriön hankintastrategian mukaisten suoraankintojen koskevien lausuntojen antaminen kehitysyhteistyömäärärahoilla hankintojen osalta.
- Kehitysyhteistyön tilastointi, raportointi ja kehitysyhteistyön läpinäkyvyyden kehittäminen.
- Määräysten ja ohjeiden valmistelu sekä neuvonta yksikön toimivallassa olevien asioiden osalta. (Suomen Ulkoasiainministeriö, 2016.)

2 TIETOVARASTOINTI

Tietovarastointi (Data Warehousing, DW) ja liiketoimintatiedon hallinta (Business Intelligence, BI) ovat nopeasti kehittyviä tietotekniikan alueita yritysten ja julkishallinnon organisaatioiden tarpeista saada käyttöönsä tietojen analysointi- ja raportointivälineitä. Perusjärjestelmät eivät pysty palvelemaan riittävän hyvin vaativia tietojen analysointi- ja raportointitarpeita suurien tietomassojen hallinnassa, jotta niistä saataisiin uutta informaatiota päätöksenteon tueksi. Tietojen jalostaminen omaan, erilliseen tietovarastokantaansa on tarpeen, jotta tietoja voidaan hyödyntää moninlaisilla raportointi- ja analysointiratkaisuilla. (Hovi, Hervonen & Koistininen, 2016, XI.)

2.1 Tietoresurssi

Yrityksissä ja julkisyhteisöissä on monenlaista pääomaa eli resursseja, kuten henkilöstö, kiinteistöt ja tuotantovälineet. Myös organisaatioiden keräämät tiedot ovat arvokas resurssi. Tietojen saaminen talteen on vaatinut investointeja tietojärjestelmiin, laitteisiin ja henkilöiden koulutukseen, lisäksi on tarvittu suuri määrä työtä tietojen keräämiseen.

Tietoa kerätään ja tallennetaan operatiivisiin järjestelmiin. Näitä ovat erilaiset ERP- (Enterprise Resource Planning) tai www-sovellukset, joihin kerääntyvät tiedot joka päivittäisen toiminnan myötä. Operatiivisissa järjestelmissä tiedot ovat usein hajallaan, tietorakenteet ovat hankalia eikä tietoja ole yleensä kuvattu riittävästi, jolloin tiedot eivät ole helposti saatavilla analysointia ja raportointia varten. (Hovi, Hervonen & Koistininen, 2009, 11.)

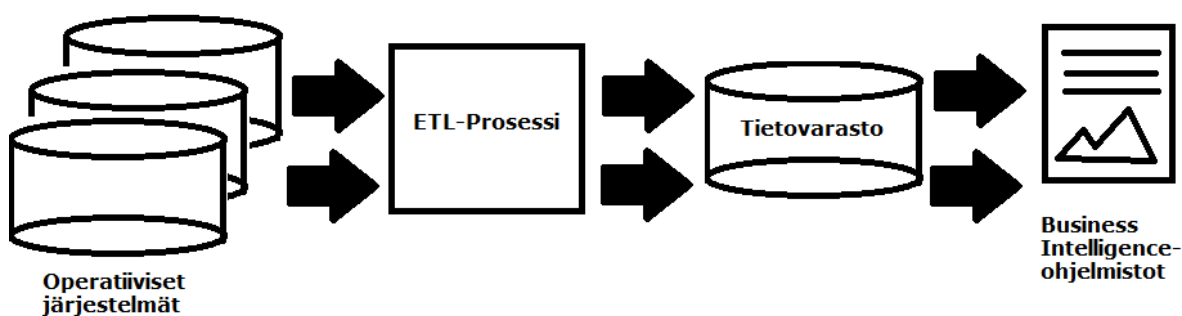
2.2 Operatiiviset järjestelmät

Operatiiviset järjestelmät eli ERP-järjestelmät ovat yksityisen tai julkisen organisaation hallinto-ohjelmistoja. Näillä ohjelmistoilla organisaatio voi kerätä, varastoida, hallinnoida ja tulkita dataa monilla eri osa-alueilla. Sellaisia ovat esimerkiksi tuotesuunnittelu, valmistus, markkinointi, myynti, varaston hallinta ja talous. (Almajali, Masa'deh & Tarhini, 2016.)

Operatiiviset järjestelmät eivät sovi tietojen raportointiin ja analysointiin. Tietojen hyödyntäminen analyyseissä ja raporteissa edellyttää tietokantaa, joka on erityisesti tätä varten suunniteltu ja rakennettu. (Hovi, Hervonen & Koistinen, 2009, 14-15.)

2.3 Tiedon varastointi tietovarastoon

Kuviossa 1 on esimerkki tietovarastointiprosessista.



Kuvio 1: Esimerkki tietovarastointiprosessista
(Hovi, Hervonen & Koistinen, 2009, 14)

Kuvio 1 kuvaa tietojen jalostusketjua perusjärjestelmien tietokannoista tietovarastoon. Tietoja muokataan raakatietomuodosta jalostettuun, raportointi- ja kyselykäyttöön sopivaan muotoon ja ladataan tietovarastoon määräajoin, yleensä kerran päivässä. Tietovaraston tiedot ovat vain lukukäytössä. Niitä ei päivitetä suoraan tietovarastoon, koska silloin operatiiviset tietokannat ja tietovarasto eivät olisi ajan tasalla keskenään. Tietojen muutokset tulevatkin aina operatiivisista järjestelmistä. Osa operatiivisista järjestelmistä voi olla yrityksen ulkopuolella, jolloin tietovarastoon ladataan myös ulkoista tietoa yhdistettäväksi omiin tietoihin. (Hovi, Hervonen & Koistinen, 2009, 14.)

Edellä kuvattua prosessia kutsutaan tietovarastoinniksi (Data Warehousing). Seuraavassa on kuvattu kaavioon liittyviä vaiheita:

- Tiedot kertyvät eri operatiivisten perusjärjestelmien tietokantoihin, jossa niitä myös pidetään ajan tasalla. ETL-vaiheessa (Extract – Transform – Load) pe-

rusjärjestelmien tiedot luetaan ja muokataan tietovarastokannan edellyttämään muotoon, mikä edellyttää yhdenmukaistamista ja yhdistämistä. Lopuksi tiedot ladataan tietovarastoon.

- Tietovarasto (Data Warehouse) on tietojen helppoa ja nopeaa hakua varten suunniteltu tietokanta. Tietovarastoon tallennetaan monen vuoden tiedot, jolloin tiedot mahdollistavat trendianalyytit. Tietovaraston tiedot määritellään ja kuvataan metadatalle.
- Tietovaraston tietoja kysellään, analysoidaan ja raportoidaan erilaisilla BI-välineillä. Tietovarastosta ajetaan valmisraportteja, parametroitavia raportteja ja uusia kyselyjä. (Hovi, Hervonen & Koistinen, 2016, 14-15.)

Koodit. ETL-vaiheessa eri operatiivisten järjestelmän koodeja on yhdenmukaistettava ja saatettava ymmärrettävämpään muotoon, esimerkiksi kun jossakin järjestelmässä sukupuolikoodit ovat "M" ja "N" ja toisessa "1" ja "0". Silloin sovitaan, että tietovarastossa käytetään koodeja "M" ja "N" ja kaikkien muiden lähdejärjestelmien koodit yhdenmukaistetaan tähän. (Hovi, Hervonen & Koistinen, 2016, 14-15.)

Tiedon jalostaminen. Tietoa jalostetaan raakatiedosta raportointia ja kyselyjä paremmin palvelemaan muotoon. Tietovaraston ETL-prosessissa otetaan raakadatatista haluttu data, jolle suoritetaan BI-välineille tarpeelliset muutokset ja laskutoimitukset valmiiksi. ETL-prosessin aikana tehdään usein myös oikeellisuustarkistuksia, jotta tietovarastoon ei siirry virheellistä tietoa. Operatiivisissa järjestelmissä olevat tiedot saadaan koottua selkeään, helppolukuiseen muotoon käyttäjien saataville erilaisilla BI-alueen kysely- ja raportointivälineillä. (Hovi, Hervonen & Koistinen, 2016, 14-15.)

2.4 ETL-prosessi

ETL (Extract, Transform, Load) on prosessi tietovarastoinnissa. Se on vastuussa datan ottamisesta lähdejärjestelmistä ja sen sijoittamisesta tietovarastoon. ETL-prosessissa on seuraavat toiminnot:

- Datan kerääminen operatiivisista järjestelmistä, datan yhdistäminen yhdeksi tietovaraston muotoon, jolloin se on valmis muokkausprosessiin.
- Datan muokkaaminen, mikä sisältää sääntöjen sisällyttämisen, siivoamisen, suodattamisen, datan yhdistämisen ja erottamisen, datan paikkojen vaihtamisen tarvittaessa ja datan tarkistussäännöt.
- Datan lataaminen tietovarastoon tai raportointityökaluihin. (Datawarehouse4u, 2009.)

2.5 Tietovarastoinnin historiaa

Tietotekniikan historian alkuajoista asti on yhtenä tietotekniikan eduista esitelty tietojen raportoinnin hyötyjä.

Johdon järjestelmät. Tietotekniikan kehittyessä erilaisia johdon MIS-järjestelmiä (Management Information Systems) eli päätöksenteon tukijärjestelmiä sekä EIS-järjestelmiä (Executive Information Systems) alettiin kehittämään. Usein ongelmana oli, mitä tietoa ja millä tavalla sitä esitettiin organisaation johdolle. Näissä järjestelmissä ei ollut mahdollista porautua tarkemmalle tasolle dataan. Ongelmana oli myös, että johdon järjestelmissä tiedot eivät olleet kaikkien saatavilla. (Hovi, Hervonen & Koistininen, 2016, 10-11.)

Infokannat. 1980-luvulla syntyi informaatiotietokannat eli infokannat erona operatiivisille tietokannoille. Isoissa yrityksissä otettiin käyttöön raportointiin erikoistuneita palveluyksiköitä, joita kutsuttiin nimellä Infocenter. Niitä varten luotiin omia tietokantoja, joihin tarpeen mukaan kopioitiin erilaisia tietoja operatiivisista järjestelmistä. Infokantojen ongelma oli, ettei niitä kehitetty kokonaisuuksina. (Hovi, Hervonen & Koistininen, 2016, 10-11.)

Taulukkolaskenta. Tietotekniikan kehittyessä tultiin vaiheeseen, jossa henkilökohtaiset tietokoneet ja taulukkolaskentaohjelmat yleistyivät. Tämä mahdollisti yleisesti tietojen tutkimisen ja analysoinnin omatoimisesti. Ongelmana oli, että tiedot saattoivat olla väärin, ja eri henkilöillä saattoi olla samasta tiedosta eri versio, mikä vaikeutti

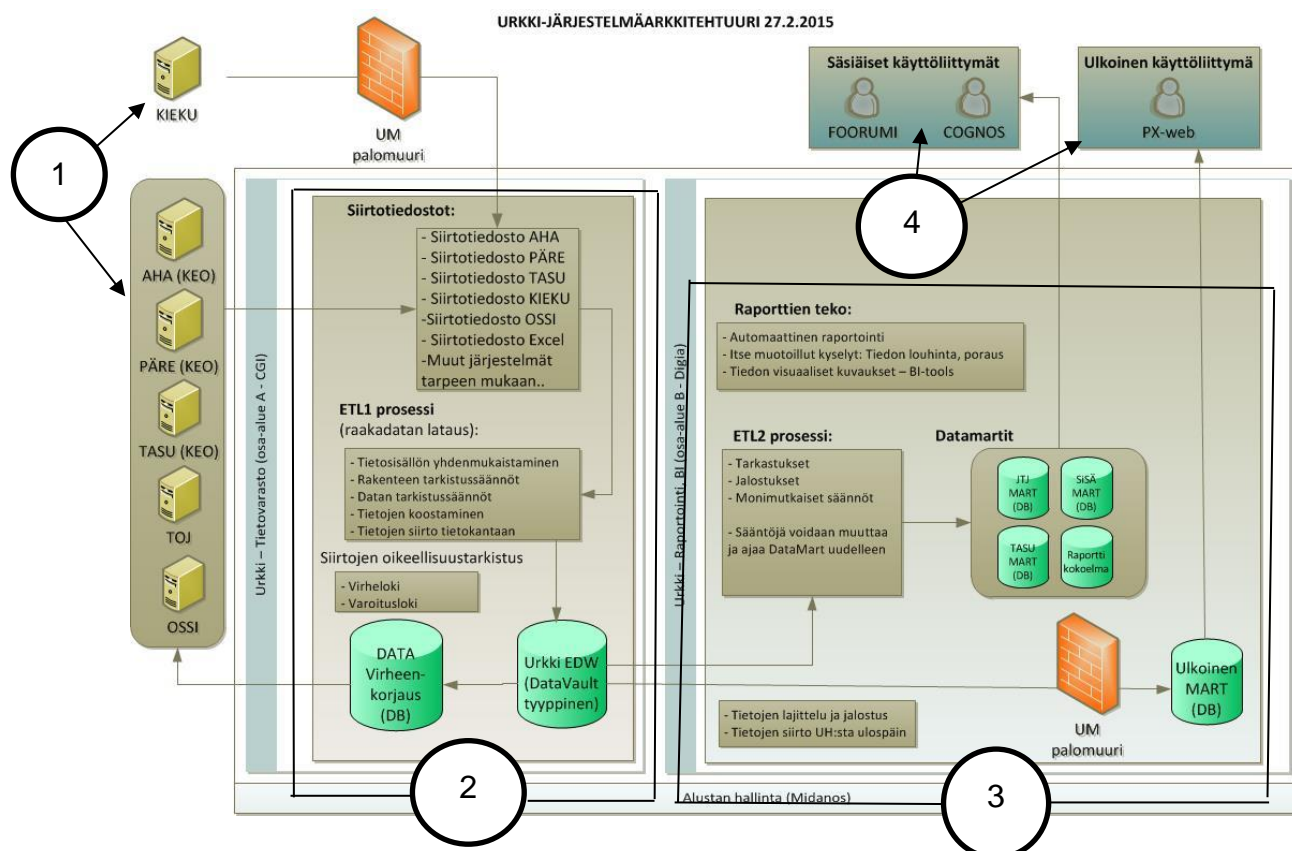
päätöksentekoa. Lisäksi tiedot saattoivat olla puutteellisia ja vanhentuneita, eikä tietojen määrittäjiä tai kuvauksia ollut. (Hovi, Hervonen & Koistinen, 2016, 10-11.)

Tietovarasto. Data Warehouse (tietovarasto) esiintyi terminä ensimmäisen kerran vuonna 1988 Devlinin ja Murphyn artikkelissa ”An architecture for business and information system” (Arkkitehtuuri bisnes- ja informaatiojärjestelmälle) IBM Systems Journal -lehdessä. W. H. Inmonia pidetään USA:ssa yleisesti tietovarasto-konseptin isänä. (Hovi, Hervonen & Koistinen, 2016, 10-11.)

Business Intelligence. Myöhemmin alettiin käyttää termiä Business Intelligence (BI, liiketoimintatiedon hallinta) tietovarastoinnin yhteydessä. BI määriteltiin tarkoittamaan käyttäjien työkaluosaamista, tietovarastoinnilla puolestaan tarkoitetaan latausprosessien ja itse tietovarastotietokannan suunnittelua ja toteutusta. Tietovarasto on IT-ammattilaisten aluetta, kun taas BI-työkaluja hyödyntävät liiketoiminnan käyttäjät. (Hovi, Hervonen & Koistinen, 2016, 10-11.)

3 URKKI-TIETOVARASTO

URKKI-tietovarasto on Ulkoasiainministeriön tietovarasto, joka otettiin käyttöön vuonna 2013. Tietovarasto mahdollistaa eri ministeriön osastotoille tehdyt, lähes reaaliaikaiset raportit ja sähköisen arkistoinnin, esimerkiksi kehitys yhteistyön projekteissa ja hankkeissa.



Kuvio 2: URKKI-järjestelmäarkkitehtuuri 27.2.2015 (Ulkoasiainministeriön kuva, 2015)

3.1 Tietovaraston lähdejärjestelmät

Ulkoasiainministeriön URKKI-tietovarasto kerää tietoa kuudesta operatiivisesta lähdejärjestelmästä (kuvion 2 kohta 1), joista 5 on Ulkoasiainministeriön sisäisiä ja yksi on ulkoinen. Sisäisistä järjestelmistä kolme on kehitys yhteistyöyksikön alaisia.

Valtion keskitetty taloushallintajärjestelmä KIEKU. Tästä ulkoisesta järjestelmästä saadaan datasyöte kehitys yhteistyön maksatuksille.

Asianhallintajärjestelmä AHA. Ulkoasiainministeriön asianhallintajärjestelmä käsittelee datasyötteen seuraavista lähteistä: konsulaatit, vientivalvonnan, vientiluvat, kaupan esteet, vientirajoitukset ja kehitysyhteistyön datasyötteen. Tämä järjestelmä tuottaa osan raporttien sisällöstä, tekijöistä ja toimijoista.

Päätösrekisteri PÄRE. Tämä järjestelmä sisältää kaikki valtion kehitysyhteistyön päätökset ja sieltä saadaan datasyöte kehitysyhteistyön myönnoille.

Taloussuunnittelu TASU. Tämä on taloussuunnitteluosaston järjestelmä, johon syötetään tietoa kehitysyhteistyöhön suunnitelluista maksatuksista.

Osoite- ja sidosryhmärekisteri OSSI. Tämä järjestelmä sisältää Ulkoasiainministeriön Suomessa sijaitsevien työntekijöiden ja lähetystöjen paikalta palkattujen työntekijöiden yhteystiedot, heidän tehtävänsä ja osastonsa. Lisäksi järjestelmä sisältää ulkomaisten lähetystöjen paikalta palkattujen henkilöiden osoitteet, yhteystiedot ja tehtäväkuvaukset sekä kaikkien Ulkoasiainministeriön yhteistyökumppaneiden tiedot. Tämä järjestelmä tuottaa osan raporttien sisällöstä, tekijöistä ja toimijoista.

3.2 Tietovarasto ja ETL-prosessit

URKKI-tietovarastossa on kaksi ETL-prosessia. Ensimmäisessä ETL-prosessissa lähdejärjestelmistä kerätään mahdollisimman paljon muuttamatonta raakadataa Data Vault-tyyppiseen tietokantaan. Datalle suoritetaan tietosisällön yhdenmukaistaminen, rakenteen ja datan tarkastaminen tarkastussääntöjen mukaisesti, koostaminen ja siirto tietovarastoon. Tämän jälkeen datalle tehdään virhetarkastus. Jos datassa havaitaan virheitä, tarvittavat korjaukset tehdään lähdejärjestelmiin. Datan siirto tapahtuu automaattisesti kerran päivässä (kuvion 2 kohta 2).

Toisessa ETL-prosessissa kerätään haluttu data BI-työkaluille, joissa varsinaiset raportit tehdään ja varastoidaan. Tässä vaiheessa data on jalostettua.

Ulkoasiainministeriön URKKI-tietovarastossa ETL-prosessien suorittamiseen käytetään IBM InfoSphere DataStage – ETL -ohjelmistoa.

3.3 Raportointi ja Business Intelligence

Toisen ETL-prosessin jälkeen kerätty tieto siirtyy datamartteihin automaattista raportointia varten (kuvion 2 kohta 3). Tämä prosessi sisältää tarkemmat säännöt ja tarkastuksen kerätylle, jalostetulle datalle. Myös manuaalinen raporttien luonti on mahdollista.

3.4 Käyttöliittymät

Datamarteista jalostettua tietoa siirretään käytettäväksi sisäisissä käyttöliittymissä Ulkoasiainministeriön intrassa käytössä oleviin palveluihin ja ulkoiseen käyttöliittymään, joka on PX-Web-tilastopalvelu (kuvion 2 kohta 4).

4 IBM INFOSPHERE DATASTAGE

IBM InfoSphere DataStage on IBM:n kehittämä ETL-työkalu ja osa IBM:n Information Platform Solutions- ja InfoSphere-palveluja. (IBM, 2016).

DataStage mahdollistaa organisaatioiden suunnitella datavirtoja, jotka keräävät dataa yhdestä tai useammasta operatiivisesta järjestelmästä, muokkaa sen arvokkaammaksi ja toimittaa sen yhteen tai useampaan tietokantaan tai sovellukseen. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 3-4).

4.1 Päätoiminnot

DataStage suorittaa muuntamisen ja siirron operatiivisista järjestelmistä kohdejärjestelmiin yhtenä eränä ja reaaliaikaisesti. Datalähteisiin voi sisältyä indeksoituja tiedostoja, peräkkäisiä tiedostoja, relaatiotietokantoja, arkistoja, ulkoisia datalähteitä, yrityssovelluksia ja viestijonoja. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 20-21.)

DataStage hallinnoi dataa, joka saapuu ja otetaan vastaan ajoittain tai säännöllisesti. DataStage kykenee suorittamaan rinnakkaisia prosessoitteja useiden multiproessorialustojen välillä, mahdollistaen DataStagen skaalautuvuuden kasvavissa datamäärissä ja reaaliaikaisuuden tarpeissa. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 20-21.)

Yhdistämällä ohjelman käytön IBM Information Server -palvelun kanssa DataStage yksinkertaistaa hallinnointidatan kehittämisen näyttämällä, minne ja kuinka data on tallennettu eri lähdejärjestelmissä. DataStage kykenee myös yhdistämään hajallaan olevaa dataa yhteen, luotettavaan rekisteriin, siivoamaan ja standardoimaan informaatiota, poistamaan duplikaatteja ja luomaan linkit rekistereiden välille eri järjestelmissä. Tämä hallinnointidata voidaan ladata operatiivisiin järjestelmiin, tietovarastoihin ja hallinnointiohjelmistoihin. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 20-21.)

DataStage toimii seuraavilla ydinominaisuuksilla:

- Yhdistettävyys keskitettyihin järjestelmiin, ylläpidettäviin ei-jatkokehitettäviin järjestelmiin ja organisaatioiden hallintojärjestelmiin.
- Funktiokirjasto, joka sisältää datan tarkistussäännöt ja muunnossäännöt.
- Maksimaalinen suoritusteho käyttäen rinnakkaista, korkeatehoista prosessointiarkkitehtuuria.
- Ominaisuudet kehittämiseen, jakamiseen, ylläpitoon ja saatavuuteen. Järjestelmä kykenee vaihtamaan metadatan käyttöä analyysien tekemisen ja ylläpitämisen välillä.
- Toimii kerta-ajoina, reaaliaikaisena kuin Web-palveluna. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 20-21.)

4.2 Datan muuntaminen

Datan muuntaminen ja liikuttaminen on prosessi, jossa lähdedata valitaan, muutetaan ja kartoitetaan siihen muotoon, mihin kohdejärjestelmä sen tarvitsee. Prosessi manipuloi dataa, jotta se vastaisi yritys-, määrittely- ja eheyssääntöjä sekä muuta dataa kohdejärjestelmässä. Muuntaminen voi esiintyä seuraavissa muodoissa:

- Aggregoiminen – datan yhdistämistä tai summaamista yhdeksi arvoksi.
- Perusmuunnos – datan varmentaminen oikeaan datatyyppiin.
- Puhdistaminen – epäjohdonmukaisuuksien ja virheiden korjaamista lähdedatassa.
- Johtaminen (derivointi) – datan muuntamista monesta lähteestä käyttäen sääntöjä tai algoritmia.

- Rikastaminen – datan yhdistämistä sisäisistä tai ulkoisista lähteistä tuoden lisää merkitystä datalle.
- Normalisointi – tarpeettoman datan ja mahdollisen monistuneen datan vähentämistä.
- Yhdistäminen – datan yhdistämistä monesta lähteestä käyttäen rinnakkaisia etsimis-, liittämis-, ja yhdistämisoperaatiota.
- Kääntäminen – rekistereiden muuntamista syötteestä useisiin eri rekistereihin tietovarastossa.
- Lajitleminen – toisiinsa liittyvien rekistereiden ryhmittelyä ja datan järjestämistä datan arvojen perusteella. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 21-22.)

4.3 Vaiheittaiset tehtävät

DataStagen tehtävät koostuvat useista yksittäisistä vaiheista, jotka on liitetty yhteen ja jotka kuvaavat datavirtaa lähteestä tai kohteeseen.

Vaihe sisältää yleensä vähintään yhden datasyötteen ja yhden ulostulon ja osa vaiheet pystyvät käsittelemään enemmän kuin yhden syötteen ja ulostulon. Jokaisella vaiheella on esi-asetetut ja muokattavat ominaisuudet, jotka kertovat vaiheelle, kuinka toimia tai prosessoida dataa. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 22-24.)

Vaiheet ja niiden yhteen liittämiset voidaan ryhmitellä jaetussa säiliössä. Näitä voidaan käyttää jaetusta säiliöstä eri rinnakkaistehtävissä. Tehtävä voidaan myös määritellä suljettuun säiliöön, mutta sitä voidaan käyttää vain tehtävässä mihin se on määritelty. (Alur, Takahashi, Toratani & Vasconcelos, 2008, 22-24.)

5 PX-WEB

PX-Web on tilastojulkaisupalvelu, joka toimii käyttöliittymänä moniulotteisiin taulukotietokantoihin. Tilastokeskuksen StatFin-tietokanta ja monet muut tietokannat on toteutettu PX-Webillä. (Tilastokeskus, 2016).

PX-Web-tietokanta koostuu PC-Axis-tiedostomuodossa olevista taulukoista (tilastokuutioista). PX-Web-palvelut toimivat kaikilla standardiselaimilla, ja niissä on laaja valikoima tiedostoformaatteja taulukoiden siirtämiseksi omiin sovelluksiin ja tarpeisiin. (Tilastokeskus, 2016.)

Automaattisen päivityksen ongelma. Käyttöönottamisessa huomattiin, että PX-Web-tilastojulkaisupalvelu toimii manuaalisesti tilastojulkaisualustana ja ettei järjestelmää ole suunniteltu käytettäväksi automaattisesti tai osana automaattista kokonaisuutta. Palvelussa käytettävien tiedostojen datan päivittyminen perustuu manuaalisiin toimenpiteisiin.

Ratkaisu ongelmaan. Tästä syystä Ulkoasiainministeriössä käytössä olevaan PX-Webiin on kehitetty ominaisuus, jossa tiedostot tuotetaan kerran päivässä ulkoisen ETL-ohjelmiston avulla (IBM InfoSphere DataStage) ja siirretään PX-Web-palvelimelle. Näin saadaan toteutettua automaattisesti päivittyvä data lähdejärjestelmästä ilman mahdollisia tietoturvariskejä lähdejärjestelmille, samalla tämä palvelu ei tuota kuormittumista lähdejärjestelmissä. PX-Web ei lähetä kyselyitä tai pyyntöjä tietokantapalvelimille, sillä suoraa tietokantayhteyttä lähdejärjestelmiin ei ole. Tästä syystä myös palvelunestohyökkäykset (DDoS) kohdistuvat vain PX-Web-palvelimelle, mutta ei lähdejärjestelmille.

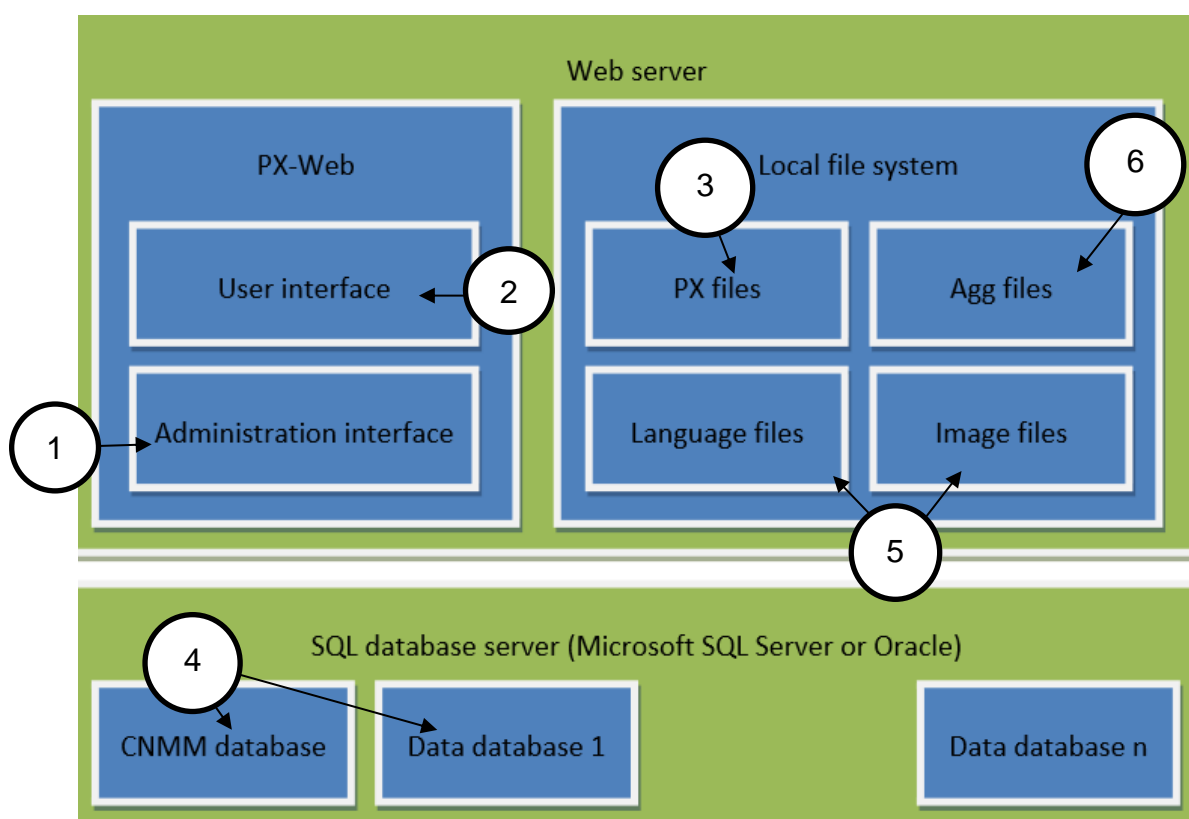
5.1 Järjestelmän kuvaus

PX-Web-palvelussa on kaksi erillistä käyttöliittymää:

- **Hallinto** – Hallintokäyttöliittymää käyttävät järjestelmän ylläpitäjät hallinoidakseen ja ylläpitääkseen PX-Web-palvelua (kuvion 3 kohta 1).

- **Käyttäjä** – Käyttäjien osa sovellusta, joka on näkyvillä loppukäyttäjille ja joka on varsinainen julkaisusovellus (Kuvion 3 kohta 2). (Statistiska centralbyrån, 2013, 4.)

Julkaistava tilastodata voidaan tallentaa PX-tiedostona paikallisesti Web-palvelimelle (kuvion 3 kohta 3) tai ulkoiselle SQL-palvelimelle (kuvion 3 kohta 4), joka käyttää Common Nordic Meta-mallia (CNMM). PX-Web kykenee myös yhdistämään kahta eri tyyppiä olevaa datalähdettä. SQL-palvelimen nimi ja tyyppi määritellään SqlDb.config-tiedostossa. Palvelimelle tallennetaan myös palvelussa käytettävät kieli- ja kuvatiedostot (kuvion 3 kohta 5) sekä aggregointitiedostot (kuvion 3 kohta 6). (Statistiska centralbyrån, 2013, 4.)

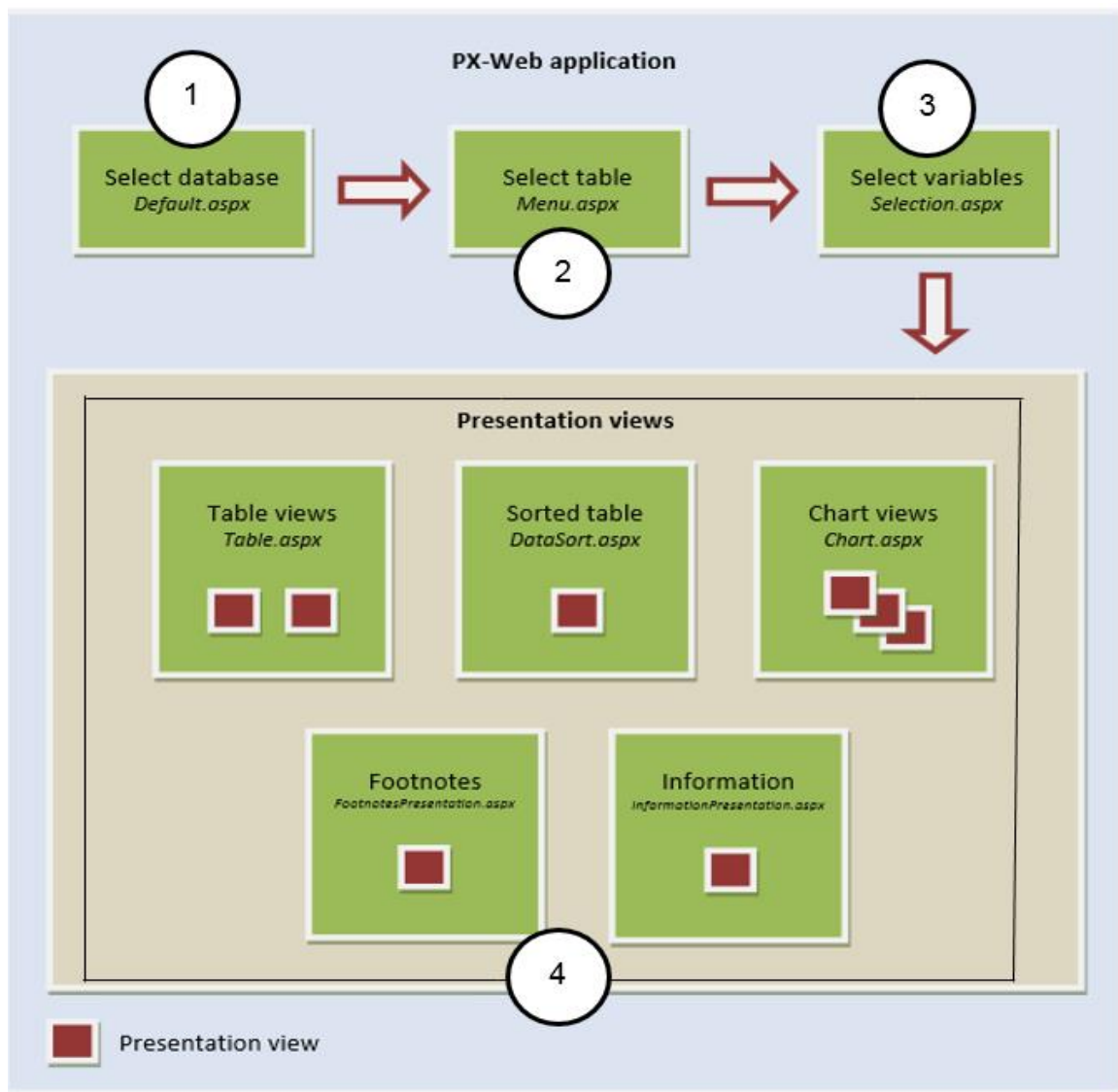


Kuvio 3: PX-Web – järjestelmäkuvaus
(Statistiska centralbyrån, 2013, 3)

Ulkoasiainministeriön käytössä ei ole suoraa tietokantayhteyttä, vaan Web-palvelimelle tallennetaan paikallisesti PX-tiedostot, joita palvelu käyttää. Tiedostojen siirto toteutetaan ajastetulla siirrolla lähdepalvelimilta ja päivitysskriptillä.

5.2 Sovellusvirtaus

Käyttäjä valitsee palvelussa haluamansa tietokannan (kuvion 4 kohta 1), jonka jälkeen käyttäjän pitää valita haluamansa tiedosto (kuvion 4 kohta 2). Tämä avaa kyselyn palvelussa, josta käyttäjä valitsee haluamansa muuttujat (kuvion 4 kohta 3) jolloin palvelu muodostaa niistä taulukon. Muodostettua taulukkoa voidaan esittää erilaisilla esitystavoilla, sekä sille voidaan suorittaa erilaisia toimintoja. Esimerkiksi taulukosta voidaan muodostaa kaavioita, tallentaa käyttäjän valitsemaan tiedostomuotoon, muokata taulukon sarakkeiden ja pylväiden paikkoja sekä katsoa tiedostosta annettua informaatiota. (kuvion 4 kohta 4). (Statistiska centralbyrån. 2013.)

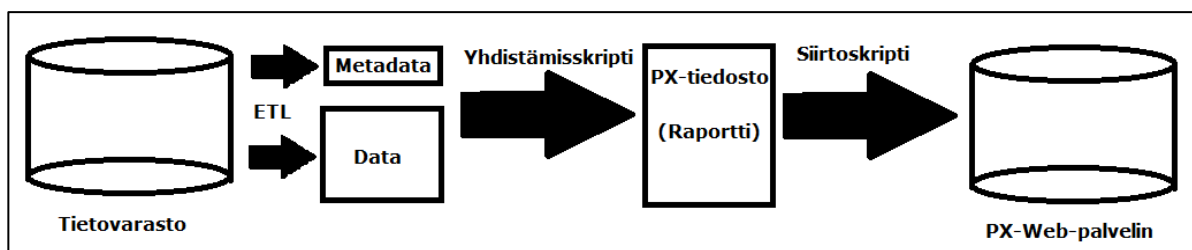


Kuvio 4: PX-Web - sovellusvirtaus (Statistiska centralbyrån, 2013, 5).

5.3 PX-tiedostojen kuvaus

PX-tiedostot (tai PC-Axis-tiedostot, tietokuutiot) koostuvat metadata-osiosta ja varsinaisesta data-osiosta. Metadata-osio sisältää dataa ohjaavat avainsanat ja niiden attribuutit. Metadata on tallennettu tietovarastoon, josta ETL-prosessi valitsee metadatan mukaisesti datasisällön. Metadata-osiota seuraa tiedostossa varsinainen datasisältö. Jos datasisältö on virheellisesti muotoiltu, niin tiedosto ei toimi. (Likidis, 2013.)

Kuviossa 5 on esitetty automaattinen PX-tiedostojen luominen ja päivitys.



Kuvio 5: PX-tiedostojen automaattinen luominen ja päivitys

Ajastettu päivitys. Tiedostojen sisältödata tuotetaan ajastetusti IBM:n InfoSphere DataStage -ohjelmistolla ja liitetään yhteen yhdistämiskriptillä metadatan mukaisesti. Tämän jälkeen muodostuneet tiedostot viedään ajastetusti toiselle palvelimelle. Siellä automaattinen päivitysskripti päivittää siirtyneen datan palveluun käytettäväksi halutuin aikavälein.

Aggregointitiedostot. Palvelussa voidaan lisäksi käyttää aggregointitiedostoja eli summaus- ja ryhmittelytiedostoja. Niillä voidaan yhtä aikaa määrittää datan ryhmittelyä sekä yhteenlaskua. Esimerkiksi kehitysyhteistyössä voidaan ryhmitellä maananosien mukaan ja laskea maanosassa olevien kehitysapua saavien valtioiden yhteensä saama avun määrä.

Yhdistämiskripti. Yhdistämiskripti lukee ajastetusti palvelimelta päivitettävien tiedostojen nimien perusteella sinne ETL-prosessin lataamat tiedostot CSV-tiedostomuodossa. Ensimmäinen tiedosto on tietovarastosta ladattu metadata-osio, toinen on datasisältö. Skripti siirtää datasisällön metadata-osion kanssa samaan CSV-tiedostoon ja muuntaa sen PX-tiedostoksi.

Siirtoskripti. Siirtoskripti siirtää yhdistämiskriptin luoman PX-tiedoston datamarkista paikalliselle PX-Web-palvelimelle ajastetusti.

6 AUTOMAATTINEN PX-TIEDOSTOJEN PÄIVITYSPROSESSI JA PALVELUN KÄYTTÖ

Ulkoasiainministeriön tilastopalvelun data päivittyy automaattisesti kerran päivässä. Data tulee lähdejärjestelmästä, josta data siirtyy ensin URKKI-tietovarastoon, jonka jälkeen yhdistämiskripti yhdistää uuden data-osion metadata-osion kanssa. Sieltä päivittyneet PX-tiedostot siirtyvät datamartin kautta PX-Web-palvelimelle. Palvelimella ajastettu päivitysskripti lataa päivitettyt tiedostot palveluun käytettäviksi.

Datasisältö. PX-tiedostojen datasisältö tulee neljästä lähdejärjestelmästä: Asianhallintajärjestelmästä AHA, Osoite- ja sidosryhmärekisteristä OSSI, Valtion keskitetystä taloushallintajärjestelmästä KIEKU ja Päätösrekisteristä PÄRE (järjestelmät esiteltä luvussa 3.1). AHA ja OSSI tuottavat dataan sisällön tekijöistä ja toimijoista, kuten kehitysapua saavista maista, kehitysavun toimialoista, kehitysavun eri rahoituskanavista ja kehitysyhteistyöhankkeista. PÄRE tuottaa sisällön PX-tiedostoissa myöntöihin ja KIEKU tuottaa datan maksatuksille.

Tiedonhakuäännöt. Kehitysyhteistyön PX-tiedostojen automaattista tuottamista varten on luotu seitsemän erilaista pohjaa, jossa on tiedonhakuäännöt. Näiden pohjalta ETL-prosessi tuottaa tiedostot halutussa muodossa ja oikealla datalla. Nämä pohjat PX-tiedostoille ovat:

- Kehitysyhteistyön myönnöt ja maksatukset
- Kehitysyhteistyön myönnöt ja maksatukset kohdemaan mukaan
- Kehitysyhteistyön myönnöt ja maksatukset kohdemaan ja rahoituskanavan mukaan
- Kehitysyhteistyön myönnöt ja maksatukset kohdemaan ja toimialan mukaan
- Kehitysyhteistyön myönnöt ja maksatukset kohdemaan ja hankkeen mukaan
- Kehitysyhteistyön myönnöt ja maksatukset osuutena bruttokansantulosta
- Kehitysyhteistyön myönnöt ja maksatukset, hankkeet

Seuraavassa kappaleessa esitetään esimerkki palvelussa olevasta automaattisesti päivittyvistä PX-tiedostosta.

Esimerkki: Kehitysyhteistyön myönnöt ja maksatukset kohdemaan mukaan. Palvelun käyttö alkaa aloitussivulta valitsemalla hiirellä ”Suomen kehitysyhteistyö”-teksti. (Kuvan 1 kohta 1). Tästä siirrytään valitsemaan tarkasteltavia PX-tiedostoja.

FORMIN.FINLAND.FI Ulkoministeriön tilastopalvelu suomi Käyttöohje

Suomen kehitysyhteistyö

Suomen kehitysyhteistyön tilastot kokoavat tietoja Suomen julkisesta kehitysyhteistyöstä. Suomen ulkoministeriön ylläpitämä tietokanta sisältää tilastoja vuosilta 2006–2017.

Tilastotietokannassa on tietoja kehitysyhteistyön toteutuneista maksatuksista ja myönnoistä. Toteutuneilla maksatuksilla tarkoitetaan kehitysyhteistyöhön tietyinä vuonna käytettyjä varoja. Myönnoillä tarkoitetaan kehitysyhteistyöhön vuoden aikana myönnettyjä varoja, joiden käyttö voi jakautua useamman vuoden ajalle.

Tietokannan sisältämät tiedot perustuvat Suomen virallisen kehitysyhteistyön raportointiin Taloudellisen yhteistyön ja kehityksen järjestölle (OECD). Voit ladata tilastotaulukot ja -kuviot tietokannasta haluamassasi muodossa.

Valitse tietokanta:
Suomen kehitysyhteistyö

Ulkoministeriö **Palaute: tilastot.um@formin.fi**

Kuva 1: PX-Web-palvelun aloitussivu

Seuraavaksi valitaan tarkasteltava PX-tiedosto. Esimerkkiä varten on valittu palvelusta tiedosto nimeltä ”Kehitysyhteistyön myönnöt ja maksatukset kohdemaan mukaan”. Painamalla hiirellä tiedoston nimeä saadaan valintasivu auki (kuvan 2 kohta 1).



>> Suomen kehitysyhteistyö

1 Valitse taulukko 2 Valitse muuttajat 3 Näytä taulukko

Suomen kehitysyhteistyö

Myönnöt ja maksatukset

- Kehitysyhteistyön maksatukset [Päivitetty 27.1.2016]
- Kehitysyhteistyön maksatukset osuutena bruttokansantulosta [Päivitetty 27.1.2016]
- Kehitysyhteistyön myönnöt ja maksatukset kohdemaan mukaan [Päivitetty 27.1.2016]
- Kehitysyhteistyön myönnöt ja maksatukset kohdemaan ja toimialan mukaan [Päivitetty 27.1.2016]

Ulkoministeriö Palaute: tilastot.um@formin.fi

Kuva 2: PX-Web-palvelun PX-tiedostojen valintasivu

Valinnan jälkeen palvelussa avautuu kyselysivu. Kyselysivulla annetaan hakumuuttujiin halutut arvot (kuvan 3 kohta 1). Tässä esimerkissä valittavina arvoina ovat maat, myönnöt, maksatukset ja vuosi. Jokaisesta muuttujasta voi valita kaikki arvot, mutta jokaisesta muuttujasta on valittava vähintään yksi arvo. Arvot voi valita valintaruuduista. Painamalla hiirellä valintaruudun yläreunan ensimmäistä nappia vasemmalta voi valita kaikki muuttujan arvot (kuvan 3 kohta 2). Tämän jälkeen tuleva toinen nappi vasemmalta poistaa kaikki valinnat ja näiden oikeanpuoleiset nuolinäppäimet siirtyvät arvojen ylimpään tai alimpaan valittavaan arvoon. Valintaikkunoiden alareunassa on tekstikenttä, jolla voi hakea tarkemmin muuttujan valintoja. Valintaikkunoiden alapuolella on alavetovalikko, josta voi valita esitysnäkymän valitulle datalle (kuvan 3 kohta 4). Alavetovalikon vieressä olevalla ”Jatka”-napilla palvelu siirtyy näyttämään valituista arvoista dataa (kuvan 3 kohta 5). Oletusarvona palvelu näyttää datan taulukkonäkymässä.



[Suomen kehitys yhteistyö >>](#)
[Myönnöt ja maksatukset >>](#)
[Kehitys yhteistyön myönnöt ja maksatukset kohdemaan mukaan](#)

1 Valitse taulukko **2** Valitse muuttajat **3** Näytä taulukko

Kehitys yhteistyön myönnöt ja maksatukset kohdemaan mukaan

Valitse muuttajat **Tietoja taulukosta**

Merkitse valintasi ja valitse esitysmuoto (taulukko ruutuun tai tiedostomuoto). Valintaohje
 *-merkityille muuttujille tarvitaan ainakin yksi arvo

Kohdema * Väli-Amerikka	Myönnöt ja maksatukset * Maksatukset	Vuosi * 2011
Yhteensä 22 Valittu 1	Yhteensä 2 Valittu 1	Yhteensä 11 Valittu 1
VALI-AMERIikka Antigua ja Barbuda Barbados Belize Costa Rica Dominica	Maksatukset Myönnöt	2011 2012 2013 2014 2015 2016
Etsi <input type="text"/> <input type="button" value="▶"/>	Etsi <input type="text"/> <input type="button" value="▶"/>	Etsi <input type="text"/> <input type="button" value="▶"/>
<input type="checkbox"/> Alkaen riviltä	<input type="checkbox"/> Alkaen riviltä	<input type="checkbox"/> Alkaen riviltä

Valittuja taulukkosoluja: 1 (enimmäisraja on 100 000)
 Selailunäkymä on rajoitettu 3 000 riviin ja 1 000 sarakkeeseen

Ulkoministeriö Palautte: tilastot.um@formin.fi

Kuva 3: PX-Web-palvelun PX-tiedostojen muuttujien valintasivu

Seuraavaksi aukeaa raporttisivu. Raportissa näkyvät valittujen muuttujien data. Datat yläpuolella on kolme eri alavetovalikkoa. Ensimmäisestä voi muuttaa taulukon näkymää ja suorittaa erilaisia sisäänrakennettuja laskutoimituksia (kuvan 4 kohta 1). Toisesta voi ladata ja tallentaa tiedoston erilaisissa tiedostomuodoissa, esimerkiksi CSV- tai Excel-tiedostona (kuvan 4 kohta 2). Kolmannesta voi valita erilaisia kuvioita ja kaavioita, joilla esittää dataa (kuvan 4 kohta 3).

Taulukossa oletusarvoisesti vasemmassa reunassa pystysarakkeessa näkyy (kuvan 4 kohta 4) maamuuttujan arvot, vaakariveillä myönnöt ja maksatukset-muuttuja (kuvan 4 kohta 5) ja sen alapuolella vuosimuuttujan arvot (kuvan 4 kohta 6).



>> Suomen kehitys yhteistyö >> Myönnöt ja maksatukset >> Kehitys yhteistyön myönnöt ja maksatukset kohtamaan mukaan

1 Valitse taulukko 2 Valitse muuttujat 3 Näytä taulukko

Näytä taulukko Tietoja taulukosta

Muuta ja laske Tallenna nimellä Taulukkonäkymä

Tallenna hakusi

Kehitys yhteistyön myönnöt ja maksatukset kohtamaan mukaan

	Maksatukset				Myönnöt			
	2010	2011	2012	2013	2010	2011	2012	2013
VÄLI-AMERIikka	29 933 356	31 464 032	24 751 720	19 661 199	33 817 047	33 858 909	35 962 259	3 807 960
Antigua ja Barbuda	22 500	7 000	19 161	17 500	31 000	7 000	27 940	18 000
Barbados	3 495	0	0	0	3 900	0	0	0
Belize	5 167	15 000	0	26 000	20 588	15 000	0	26 000
Costa Rica	141 600	126 345	80 199	29 778	0	36 271	0	4 128
Dominica	15 000	1 000	18 500	1 100	16 000	1 000	20 000	0
Dominikaaninen Tasavalta	166 038	208 804	128 335	188 620	2 000	397 188	11 629	4 645
El Salvador	105 107	480 689	85 078	97 807	13 236	400 000	279 650	0
Grenada	0	19 000	0	0	0	25 000	0	0
Guatemala	745 049	1 973 419	1 315 137	1 006 642	1 831 888	2 206 642	512 578	65 000
Haiti	9 401 454	5 924 837	1 833 335	3 368 281	8 984 880	5 742 599	8 580 397	187 432
Honduras	953 276	6 138 389	3 949 350	2 453 992	2 182 016	9 113 466	16 848 662	45 000
Jamaika	12 290	57 709	0	7 500	70 000	57 709	0	7 500
Kuuba	5 709	0	0	18 925	0	18 925	0	0
Meksiko	735 986	603 089	421 365	598 571	621 645	565 754	718 962	616 821
Nicaragua	12 870 443	11 307 285	11 020 111	3 011 769	11 872 090	11 040 401	3 326 318	2 473 508

Kuva 4: PX-Web-palvelun PX-tiedoston raporttisivu

7 AUTOMAATTISEN JA MANUAALISEN PX-WEB-TILASTOJULKAISUN VERTAILU

Automaattisella ja manuaalisesti tehtävällä PX-Web-tilastojulkaisulla on molemmilla omat etunsa, hyötynsä ja haittansa. PX-Web-tilastopalvelu, joka on kehitetty pohjoismaiden tilastokeskusten yhteistyönä, on suunniteltu toimimaan manuaalisesti tehtävien tilastodatojen julkaisua varten. Ulkoasiainministeriön tarpeesta PX-Web saatiin toimimaan automaattisesti osana URKKI-tietovarastoa.

7.1 Manuaalinen tilastojulkaisu

Manuaalisessa tilastojulkaisussa datan kerääminen, tarkistaminen ja julkaisu tehdään käsin.

Edut. Etu manuaalisessa tilastojulkaisussa on datan tarkkuus. Koska käsiteltävää dataa on voitu tarkastaa ja varmentaa julkaisusta vastuussa olevan henkilön toimesta ennen tilastointia ja julkaisua, tilastodata on tarkkaa ja mahdollisimman virheetöntä.

Hyödyt. Tilastoihin perustuvat laskutoimitukset ja tilastojen muutosten seuraaminen on sitä luotettavampaa, mitä tarkempaa tilastollinen data on.

Haitat. Manuaalisesti tehtävä tiedon kerääminen, kokoaminen ja tarkistaminen vie paljon aikaa. Tämän takia julkaisutiheys on harvaa ja julkaistava data ei sisällä reaaliaikaista dataa. Reaaliaikaisuuden puute aiheuttaa ristiriidan läpinäkyvyyden kanssa.

7.2 Automaattinen tilastojulkaisu

Automaattisessa tilastojulkistamisessa tilastodata tuotetaan tietojärjestelmän avulla.

Edut. Etuna automaattisessa tilastojulkaisussa on nopeus. Julkaistavaa dataa saadaan lähes reaaliaikaisesti, jolloin nähdään nykyhetkessä tilastojen sisältö ja muutokset.

Hyödyt. Reaaliaikaisen tilastodatan hyöty tulee esille läpinäkyvyydessä. Datan manipuloinnin mahdollisuudet vähenee, kun raportoitu data pystytään julkaisemaan saman päivän sisällä, kun se on raportoitu. Reaaliaikaisista tilastoista voidaan myös havaita muutoksia nopeammin ja niihin voidaan myös reagoida nopeammin.

Haitat. Tilastojulkistamisen nopeudesta johtuen tilastodatan tarkastaminen ja tarkkuus ei ole välttämättä samaa tasoa kuin manuaalisesti tehtävässä, jolloin data saattaa sisältää virheitä. Siksi käytettäessä automaattista tilastopalvelua tulee ottaa huomioon tämä mahdollisuus.

7.3 Ulkoasiainministeriön automaattinen tilastojulkaiseminen

Palvelua käyttöönottaessa vaatimuksena oli, että tilastojulkaisun päivitys tapahtuu päivittäin. IATI:n läpinäkyvyyskampanjaan perustuvassa kehitysapua myöntävien valtioiden ja organisaatioiden vertailussa yhtenä pisteytysperusteena on datan julkaisun nopeus, josta tuli vaatimus palvelun päivittymisnopeudelle.

Keskustelua käyttöönottovaiheessa aiheutti mahdollisuus, että automaattisen tilastojulkaisun datassa on virheellisyttä. Kuitenkin tultiin siihen tulokseen, että mahdollisuus datan virheellisyyteen oli vähäinen ja päivitysnopeuden ollessa tärkeämpi tavoite tultiin tulokseen, että automaattinen julkaisu on manuaalista parempi.

8 YHTEENVETO

Työ on osoittanut, että PX-Web sellaisenaan on toimiva ohjelmisto tilastojen verkkojulkaisemista varten. Tästä osoituksena on, että Suomen Tilastokeskuksen, Riista- ja Kalatalouden Tutkimuslaitoksen, Helsingin kaupungin ja Suomen Tullilaitoksen tilastojulkaisut on toteutettu tällä menetelmällä.

Suomen Ulkoministeriön käyttöön kyseinen menetelmä ei kuitenkaan soveltunut. Tämä johtuu siitä, että manuaalinen tiedostojen tuottaminen on liian työvoimasidonnaista ja aikaa vievää. Tästä johtuen sivujen päivittymisrytmi on Ulkoministeriön tapauksessa liian hidasta.

PX-Webbiä modifioimalla onnistuttiin tuottamaan siitä version, jonka päivittymisnopeus on riittävä. Ulkoministeriössä käytetään tämän opinnäytetyön kirjoittamishetkellä yhden päivän päivittymisnopeutta, mikä oli Ulkoministeriön tietohallinnon vaatimusmäärittelyn mukainen.

Palvelu on käytettävissä osoitteessa:

stats.formin.fi

LÄHTEET

- Almajali, D. Masa'deh, R. & Tarhini, A. 2016. Antecedents of ERP systems implementation success: a study on Jordanian healthcare sector. [Viitattu 08.11.2016] Painos 29. Numero 4, 549–565. Saatavissa: <http://www.emeraldinsight.com/doi/abs/10.1108/JEIM-03-2015-0024?journalCode=jeim>
- Alur, N., Takahashi, C., Toratani, S. & Vasconcelos, D. 2008. IBM InfoSphere DataStage Data Flow and Job Design. [Verkkójulkaisu]. IBM. [Viitattu 9.11.2016]. Saatavissa: <http://www.redbooks.ibm.com/redbooks/pdfs/sg247576.pdf>
- Datawarehouse4u. 2009. ETL process. [Verkkosivu]. Datawarehouse4u. [Viitattu 9.11.2016]. Saatavissa: <http://datawarehouse4u.info/ETL-process.html>
- Hovi, A., Hervonen, H. & Koistinen, H. 2009. Tietovarastot ja Business Intelligence. Jyväskylä: Docendo.
- IBM. 2016. IBM InfoSphere DataStage. [Verkkosivu]. IBM. [Viitattu 11.10.2016]. Saatavissa: <http://www-03.ibm.com/software/products/fi/ibminfodata>
- IATI. 2016a. International Aid Transparency Initiative. [Verkkosivu]. IATI. [Viitattu 01.11.2016]. Saatavissa: <http://www.aidtransparency.net/>
- IATI. 2016b. IATI Standard. [Verkkosivu]. IATI. [Viitattu 01.11.2016]. Saatavissa: <http://iatistandard.org/>
- Likidis, P. 2013. PX-file format specification. [Verkkójulkaisu]. Statistiska centralbyrån. [Viitattu 27.10.2015]. Saatavissa: http://www.scb.se/Upload/PC-Axis/Support/Documents/PX-file_format_specification_2013.pdf
- Statistiska centralbyrån. 2013. PX-Web system description. [Verkkójulkaisu]. Statistiska centralbyrån. [Viitattu 17.2.2016]. Saatavissa: <http://www.scb.se/Upload/PC-Axis/Download/PX-Web/2013/SystemDescription-pxweb.pdf>
- Suomen Ulkoasiainministeriö. 2016. [Verkkosivu]. Suomen Ulkoasiainministeriö. [Viitattu 01.11.2016]. Saatavissa: <http://formin.fi/Public/default.aspx>
- Tilastokeskus. 2016. PX-Web. [Verkkosivu]. Tilastokeskus. [Viitattu 02.11.2016]. Saatavissa: https://tilastokeskus.fi/tup/pcaxis/px_web_ominaisuudet.html