

**Perustietovarantojen teknisten rajapintojen nykytilan kartoitus ja
kehitysehdotukset**

Sisällys

Perustietovarantojen teknisten rajapintojen nykytilan kartoitus ja kehitysehdotukset	1
1. Yleistä.....	2
2. Nykytilan yhteenveto	3
2.1 Perustietovarantojen väliset rajapinnat	3
2.2 Perustietovarantojen tietopalvelurajapinnat	4
2.3 Lisäarvopalvelun tuottajien rajapinnat	4
3. Nykytilan havainnot	5
3.1 Rekisterien julkiset tekniset kuvaukset ovat puutteellisia.....	5
3.2 Yhteiset koodistot ja sanastot puuttuvat	5
3.3 Keskeisin perustietovaranto on Väestötietojärjestelmä	5
3.4 Tiedonvaihto perustuu määrämuotoisten tiedostojen siirtoihin	5
3.5 Tiedonsiirron tietoturvamennettelyt ovat usein vanhanaikaisia.....	6
3.6 Päätekäyttöiset yksittäiskysely-yhteydet on korvattu web-liittymällä	6
3.7 Keskeiset tulevaisuuden tekniikat ovat HTTP(S), HTML, Web Services, XML	6
3.8 Sovellus-sovellusyhteydet ovat usein sovellus- ja asiakaskohtaisia sovituksia	7
3.9 Valmisohjelmissa ei ole rajapintoja rekisteritiedon hyödyntämiseen	7
3.10 Kaupallinen hyödyntäminen on edistänyt teknistä kehittämistä	7
3.11 Viranomaiskäytössä käytetään lisäarvopalveluntuottajan tietopalvelua.....	7
4. Ehdotukset	8
4.1 Aiempia tietoluovutukseen liittyviä linjauksia.....	8
4.2 Perustietovarantojen keskeisen sisällön ja tietopalvelurajapintojen dokumentaatio helposti saataville	9
4.3 Yhteiset koodistot ja sanastot	9
4.4 Sovellus-sovelluskyselyille yhtenäiset Web Services-rajapinnat.....	10
4.5 Tietoluovutusten painopiste eräsiirrosta sovellus-sovellusyhteyksiin	10
4.6 Siirrettyjen tietojen aikaleimojen määrittely ja yhtenäinen toteuttaminen	10
4.7 Tietoturvallisuusmenettelyjen kehittäminen	11
4.8 Tietopalvelun tekninen erottaminen perusjärjestelmästä	11

1. Yleistä

Tämä selvitys on osa valtionhallinnon arkkitehtuurin kehittäminen-hankkeen Perustietovarannot-projektia. Teknisiä rajapintoja koskevan selvityksen tavoitteena on muodostaa yleiskuva julkisen hallinnon sisällä tapahtuvaan tietojen vaihdon teknisistä ratkaisuista, arvioida niitä suhteessa valtionhallinnon tietotekniikkaa koskeviin linjauksiin ja laatia arvion pohjalta kehittämisehdotuksia.

Tarkastellut perustietovarannot ovat:

- Väestötietojärjestelmä
 - Henkilöt
 - Rakennukset
- Kiinteistötietojärjestelmä
 - Kiinteistörekisteri
 - Lainhuuto- ja kiinnitysrekisteri
- Yritys- ja yhteisötiedot
 - Kaupparekisteri
 - Yhdistysrekisteri
 - Arvonlisävero ja työnantajarekisterit
 - Yritys- ja yhteisötietojärjestelmä
 - Yritys- ja toimipaikkarekisteri
- Ajoneuvoliikennerekisteri
- Kansallinen tie- ja katutietojärjestelmä Digiroad
- Paikkatiedot
- Verotustiedot
- Kansaneläkelaitoksen etuustiedot

Väestötietojärjestelmä, Kiinteistötietojärjestelmä sekä useista eri rekistereistä koostuvat yritys- ja yhteisötiedot ovat *perusrekistereitä*. Ne sisältävät tietoa yhteiskunnan perusyksiköistä – ihmisistä, kiinteistöistä, rakennuksista ja erilaisista rekisteröidyistä yhteisöistä. Muut tarkasteluun otetut tietovarannot sisältävät mm. sellaista viranomaistoiminnassa syntyvää tietoa, jota hyödynnetään laajasti yhteiskunnassa.

Valittujen perustietovarantojen teknisten rajapintojen nykytilan selvittämiseksi haastateltiin jokaisen tietovarannot osalta nimetty tekninen yhteyshenkilö. Muodostunutta kuvaa täydennettiin samojen tietovarantojen hallinnollisia järjestelyjä koskeneiden haastattelujen tuloksilla. Lisäksi käytettiin julkisista lähteistä saatavia tietoja. Näitä lähteitä ovat mm. rekisteripoolin Perusrekisterit-kirja sekä rekisterinpitäjien ja lisäarvopalvelujen tuottajien Internet-sivut. Haastattelurungon ja arvion laatimisessa käytettiin ohjaavina dokumentteina valtiovarainministeriön suosituksia, erityisesti XML-strategia, rajapintasuositus ja tietoturvaan liittyvät suositukset.

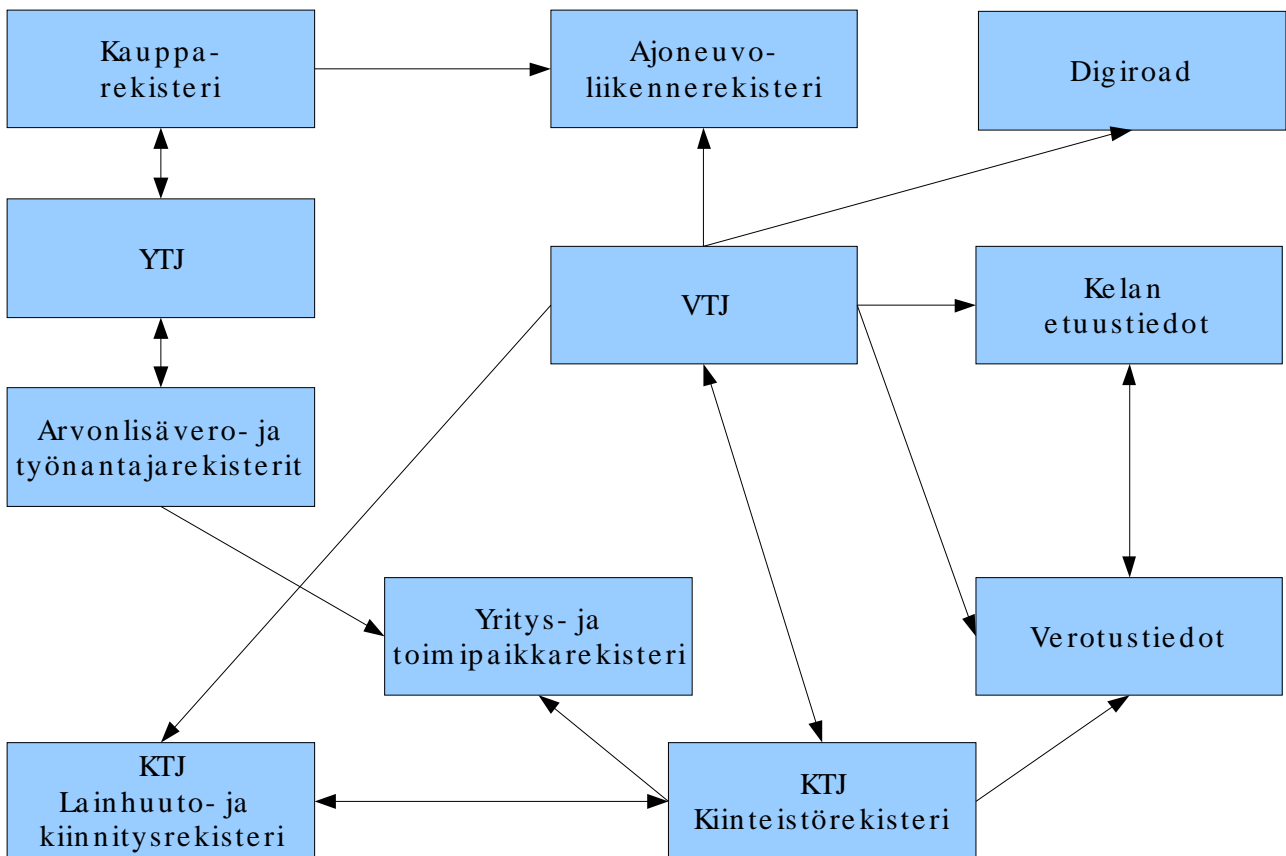
2. Nykytilan yhteenveto

Haastattelujen tulokset on koottu taulukoihin, jotka esitetään liitteissä 1 - 3. Yhteenvetotaulukoissa esitetään perustietovarantojen väliset rajapinnat, perustietovarannosta vastaavan viranomaisen tarjoamat tietopalvelurajapinnat ja lisäarvopalvelun tuottajien tarjoamat rajapinnat perustietovarannoittain.

2.1 Perustietovarantojen väliset rajapinnat

Liitteen 1 yhteenvetotaulukkoon on ristiintaulukoitu perustietovarannot, niin että vaakarivillä esitetään rekisterit tietoa luovuttavassa roolissa ja pystyrivillä vastaanottavassa roolissa. Taulukkoon on kirjattu käytetty tiedonvaihtotapa.

Perustietovarantojen välisiä tietosiirtoja havainnollistaa seuraava kaavio.



Kaaviossa on esitetty tarkasteltujen varantojen väliset tiedonvaihdot. Kuviossa nuoli kuvaa tiedonsiirron suuntaa. Molempiin suuntiin osoittava nuoli tarkoittaa, että tietoa siirretään molempiin suuntiin kyseisten järjestelmien välillä. Erikseen on kuvaamatta paikkatiedot, jotka suhteuttavat useiden eri tietovarantojen tiedot ympäröivään maailmaan.

2.2 Perustietovarantojen tietopalvelurajapinnat

Osiassa selvitettiin tarkasteltujen perustietovarantojen tietojen suoran hyödyntämisen mahdollisuuksia kartoittamalla ne tekniset käyttötavat, joilla rekisterien sisältö on tiedonhyödyntäjän käytettävissä. Perustietovarantojen tietopalvelurajapintojen ja protokollien yhteenveto on liitteenä 2 olevassa taulukossa.

2.3 Lisäarvopalvelun tuottajien rajapinnat

Lisäarvopalvelulla tarkoitetaan palvelua, jossa kolmannen osapuolen toimija saa tietoa perustietovarannoista ja jalostaa sitä tai sen käyttömahdollisuuksia edelleen. Liitteenä 3 olevassa taulukossa on esitetty tarkasteltujen tietovarantojen keskeiset lisäarvopalvelun tarjoajat ja niiden tarjoamat perustietovarantojen tietopalvelurajapinnat.

3. Nykytilan havainnot

Alla on esitetty keskeiset haastattelujen ja muun työn perusteella muodostuneet havainnot, joiden perusteella on laadittu kehitysehdotukset. Haastateltavat eivät nähneet teknisiä esteitä ylitsepääsemättöminä, vaan useimpien näkemyksen mukaan kehittämisen esteet ovat hallinnollisia. Eri haastatteluissa esiintuodut kehitysnäkymät noudattavat varsin hyvin samoja teknisiä linjauksia.

3.1 *Rekisterien julkiset tekniset kuvaukset ovat puutteellisia*

Vain harvasta tarkasteltavan tietovarannon hyödynnettävissä olevista tiedoista ja teknisistä rajapinnoista on julkisesti saatavilla ajantasainen ja kattava kuvaus. Tietovarannon hyödyntämisen suunnittelu on hankalaa, koska tietoa varannon rajapinnoista on vaikeaa saada käyttöön. Tämä johtaa helposti järjestelmäkohtaisiin rajapintoihin ja päällekkäisten tietovarantojen perustamiseen.

3.2 *Yhteiset koodistot ja sanastot puuttuvat*

Monet järjestelmät käyttävät samoja peruskoodistoja. Koodistoja on vaikea saada käyttöön. Tämä johtaa mm. siihen, että järjestelmiin talletetaan turhaa tietoa ja sitä vaaditaan sähköisissä palveluissa asiakkailta. Yksinkertainen esimerkkinä tästä on yleinen tapa pyytää postinumeron lisäksi postitoimipaikan nimeä, vaikka se löytyy postinumerokoodistosta.

Koodistojen muutosten hallintaan ei ole käyttäjäorganisaatioille yhtenäisiä ohjeita eikä käytäntöjä. Ajankohtaisena esimerkkinä tästä on kuntakoodien käsittely tietojärjestelmissä kuntaliitosten yhteydessä.

Keskeisen tietosisällön kohdalla yhteiset sanastot puuttuvat ja käsitteistö on paikoin horjuvaa.

3.3 *Keskeisin perustietovaranto on Väestötietojärjestelmä*

Väestötietojärjestelmässä (VTJ) olevat henkilö- ja rakennustiedot ovat eniten käytetty tietovaranto. Suuri joukko viranomaisjärjestelmiä saa nämä tiedot, erityisesti henkilötiedot VTJ:stä. VTJ-tiedot ovat myös merkittävin tietojoukko muiden tarkasteltujen perustietovarantojen primääritietoja täydentävästä tiedosta.

3.4 *Tiedonvaihto perustuu määrämuotoisten tiedostojen siirtoihin*

Tietovarantojen hyödyntäminen toisessa järjestelmässä (myös perusvarantojen välillä) perustuu tyypillisesti eräsiirtoon. Eräsiirrosta muuttuneet tiedot siirretään yleensä FTP:llä suljetussa verkossa määrämuotoisina tekstitiedostoina hyödyntävään järjestelmään. Järjestelmän tietokanta alustetaan siirtämällä ensin yleensä koko aineisto, minkä jälkeen määrävälein tapahtuvilla muutosten siirroilla pyritään pitämään rekisterin kopio ajan tasalla.

Useat hyödyntävät järjestelmät on rakennettu niin, että toisesta perusrekisteristä saatua tietoa voidaan muuttaa ("päivittää") järjestelmän omassa tietokannassa, jos "parempaa" tietoa satutaan saamaan esimerkiksi suoraan asiakkaalta. Toimintatapa aiheuttaa ainakin sen, että eri viranomaisilla on samasta asiasta eri tietoa.

Väestörekisterikeskus pitää yllä tietoa toiselle toimittamiensa tietojen poimintakriteereistä ja tiedoista tämä perusteella tehdään VTJ:n tiedoista muutostiedostot asiakaskohtaisesti. Hyvää siinä

on se, että jokaiselle toimitetaan vain omaan tietojoukkoon kohdistuvat muutokset ja siirrettävä tietojoukko pysyy tarkoituksenmukaisena. Haastatteluissa ei selvinnyt, onko menettely tietojen hyödyntäjien kannalta toimiva.

Suuria harvoin siirrettäviä tietojoukkoja siirretään myös erillisillä massamuistivälineillä (CD, DVD, siirrettävät kiintolevyt).

Siirtotiedoston muoto on tyypillisesti määrämuotoinen tekstitiedosto. Käytännössä usein siirretään yhä "reikäkortin kuvia". Kehitys on kulkemassa pikku hiljaa kohti XML-muotoa. Skeemoja tiedon kuvaamiseen ei välttämättä ole valmiina tai edes tekeillä, jolloin kehitellään omia. Omien skeemojen kehittäminen voidaan nähdä ongelmana, koska ne eivät mahdollisesti ole yhteensopivia "tulevien" standardien kanssa. Tämä on kuitenkin usein ainoa tapa edetä eikä siirtyminen tulevaisuudessa "standardiskeemoihin" ole välttämättä suuri ongelma. Suurin työ on esittää tiedot ylipäänsä käyttökelpoisessa rakenteisessa muodossa.

3.5 Tiedonsiirron tietoturvamenettelyt ovat usein vanhanaikaisia

Eräsiirroissa siirtoteitä ei yleensä salata eikä siirtotiedostoja salata tai allekirjoiteta digitaalisesti, koska siirtojen on vallitsevan käytännön mukaan määrä tapahtua suljetussa verkossa tai suljettujen verkkojen välillä. Siirtotiedostojen alkuperän ja eheyden varmistamiseen ei juuri ollut käytössä selkeitä mekanismeja. Alkuperän varmistaminen perustuu käytännössä aina tunnistamalla palvelimen IP-osoite ja tarkastamalla siirtäjän käyttäjätunnus/salasana.

Tietoturvamielessä salauksen yleinen puuttuminen ei ole välttämättä ongelma. Siirrettävä tieto on usein julkista ja siirrot tapahtuvat suljetussa verkossa. Salausmekanismien puuttuminen voi kuitenkin muodostua esteeksi joustavalle siirtotien valinnalle ja voi jopa mahdollistaa tietojen tahallisen väärentämisen. Kehitys on jo pitkään kulkenut suuntaan, että organisaatioiden välisiä yhteyksiä ei enää muodosteta suljetussa verkossa vaan verkkojen välinen liikennöinti tapahtuu Internetin yli. Usein olisi helpompaa sopia yhteisestä siirrettävän aineiston salaustavasta ja -avaimista kuin yhteisen VPN-yhteyden perustamisesta.

Tiedon voimassaoloa ei voi aina tarkistaa. Riskinä on esimerkiksi, että päivitys jää jostain syystä lataamatta ja sitä ei huomata. Tiedoista ei voi aina päätellä, ovatko ne ajan tasalla vai ei. Aikaleimat puuttuvat tai niiden toteutus on sovelluskohtaista, mikä vaikeuttaa niiden hyödyntämistä tietoja vastaanottavissa järjestelmissä. Kunnollinen ja yhtenäinen aikaleimojen toteuttaminen mahdollistaisi tiedon luontevan väliaikaisen tallentamisen kohdejärjestelmään.

3.6 Päätekäyttöiset yksittäiskysely-yhteydet on korvattu web-liittymällä

Erityisesti 3270-pääteyhteyksiä on korvattu web-yhteyksillä. Tämä on ollut luonteva kehitys ja toteuttaminen ovat olleet kohtalaisen helppoa, koska sekä web-liittymä että 3270-päätesovellukset ovat molemmat sivuorientoituneita.

Myös muuten Internet-tekniikoita on otettu ja otetaan koko ajan käyttöön eikä se välttämättä vaadi koko järjestelmän uusimista. Keskeisimmät kehityspolut ovat olleet XML:n käyttöönotto määrämuotoisen tiedon esittämisessä, HTTP/HTTPS-protokollan käyttöönotto kuljetusprotokollana, SSL:n käyttö yhteyksien salauksessa ja Web Services-tyyppisen rajapinta-ajattelun omaksuminen.

3.7 Keskeiset tulevaisuuden tekniikat ovat HTTP(S), HTML, Web Services, XML

Haastatteluissa saatujen järjestelmien kehityssuunnitelmia koskevien kommenttien perusteella valitut tekniikat ja arkkitehtuurit noudattavat varsin yhdenmukaista linjaa. Uudistettavissa järjestelmissä trendinä on toteuttaa tiedonvaihtoa XML-muotoisina sanomina käyttäen Web

Services-rajapintaa. Rajapintaa hyödynnetään omien liittymien toteuttamisessa ja sama rajapinta tarjotaan ainakin teknisesti myös muille sovelluksille.

Toteutettavaksi valitussa arkkitehtuurissa operatiivinen järjestelmä ja tietopalvelujärjestelmä erotetaan tyypillisesti toisistaan. Erottamisella tavoitellaan operatiivisen järjestelmän suojaamista, tietopalvelujärjestelmän skaalautuvuutta ja redundanssia. Nykyisten järjestelmien jäykkä skaalautuvuus on ollut eräiden haastateltujen käsityksen mukaan esteenä laajojen reaaliaikaisten sovellus-sovellus kyselyjen käyttöönotolle.

Osa tiedoista on sellaista, ettei niiden suora kyselykäyttö ole tarkoituksenmukaista ainakaan ilman niitä jalostavaa tai integroivaa välikerrosta tai syystä, että siirrettävät tietomassat ovat erittäin suuria. Näissä tapauksissa eräsiirrot jäävät käyttöön rinnalle mutta on oletettavaa, että niiden määrä muuten tulee vähenemään huomattavasti.

3.8 Sovellus-sovellusyhteydet ovat usein sovellus- ja asiakaskohtaisia sovituksia

Rekistereihin on harvoin määritelty yleistä sovellus-sovellusrajapintaa. Rajapintoja tehdään hyödyntäjän tilauksen mukaisesti asiakaskohtaisesti. Tämä johtaa siihen että rajapintoja syntyy monia erilaisia ja niitä on teknisesti vaikea hallita. Asiakaskohtaiset rajapinnat ovat osoittautuneet erittäin vaikeiksi järjestelmien muutostilanteissa, koska muutokset joudutaan toteuttamaan ja ottamaan käyttöön samanaikaisesti.

Puuttuva valmis rajapinta voi helposti johtaa päätökseen, että hyödyntävään järjestelmään on lopulta helpompi toteuttaa oma tietovarasto.

3.9 Valmisohjelmissa ei ole rajapintoja rekisteritiedon hyödyntämiseen

Julkisessa hallinnossa käytetään paljon valmisohjelmistoja. Näissä ei yleensä ole valmiutta hyödyntää muualta saatavia perustietoja. Toimittajilla ei ole halua toteuttaa valmiuksia, jos muutoksille ei ole tiedossa ostajia. On oletettavaa, että jos perusrekistereihin olisi selkeät hyvin määritellyt tietopalvelurajapinnat sovellus-sovelluskyselyille, niitä hyödyntävien erilaisten valmisohjelmistojen valmistajille olisi kannattavaa toteuttaa niitä hyödyntävät kyselyrajapinnat. Vakioidut rajapinnat ja tietosisältö johtaisivat todennäköisesti tiedon käsittelyn laadun paranemiseen vastaanottavassa järjestelmässä esim. aikaleimojen johdonmukaisena hyödyntämisenä.

3.10 Kaupallinen hyödyntäminen on edistänyt teknistä kehittämistä

Lisäarvopalvelujen tuottajia varten on eräisiin rekistereihin toteutettu sovellus-sovellusyhteyksiä. Toteutetut rajapinnat ovat osoittautuneet toimiviksi ja niitä on alettu hyödyntää myös viranomaiskäytössä. Kaupallisten toimijoiden esittämällä toiveilla on ollut enemmän vaikutusta kuin valtionhallinnon tietojärjestelmiä koskevilla suosituksilla.

3.11 Viranomaiskäytössä käytetään lisäarvopalveluntuottajan tietopalvelua

Viranomaiskäytössä on usein päädytty lisäarvopalveluntuottajan tietopalvelun käyttöön tietovarannon suoran käytön sijaan. Tämä voi johtua paitsi siitä, ettei varantoon edes ole käytettävissä suoraa tietopalvelurajapintaa niin ennen kaikkea siitä, että lisäarvopalveluntuottajan tietopalvelurajapinta koetaan ylivoimaisesti toimivimmaksi käytettäväksi.

4. Ehdotukset

Teknisiä ehdotuksia laadittaessa on pyritty noudattamaan käytännönläheistä lähestymistapaa. Ehdotuksien tarkoitus ei ole luoda kokonaisvaltaista ratkaisua vaan niillä pyritään edistämään toimeksiannon mukaisesti perustietovarantojen hyödyntämistä.

Esitetyt ehdotukset on laadittu perustietovarantojen hyödyntäjien näkökulmasta. Ehdotuksia ei voi asettaa prioriteettijärjestykseen vaan ne muodostavat pitkälti toisistaan riippuvan kokonaisuuden. Ehdotukset ovat monilta osin soveltamiskelpoisia myös perustietovarantojen välisten rajapintojen yhteydessä. Tämä työ on kuitenkin hyvä erottaa omaksi kokonaisuudekseen, josta vastaa perusrekisteriasioiden neuvottelukunta.

4.1 Aiempia tietoluovutukseen liittyviä linjauksia

Työn tueksi on käyty läpi asiaan liittyviä aiempia linjauksia. Niistä tärkeimpiä ja tätä työtä tukevia linjauksia on lueteltu alla.

Valtiovarainministeriön VM 31/01/2001 suositus “Valtion tietotekniikan rajapintasuosituksia“.

”Yhtenä keskeisenä elementtinä suosituksessa oli XML-standardin käyttö. Suosituksen mukaan XML:n käyttöä tulee lisätä ja opetella ja XML-standardin hyödyntämisen ydin on tarkoituksenmukaisten ja yhteiskäyttöisten sanastojen luonti ja käyttö. XML:n käyttö onkin laajentunut nopeasti paitsi valtionhallinnossa myös kunnissa.” (Julkishallinnon XML-strategia, VM 18/2003)

Julkishallinnon XML-strategia (VM 18/2003) ja julkishallinnon XML-strategian toteuttaminen (VM 7/2004)

”Keskitettyjen tietokantojen rakentaminen on toimiva ratkaisu joissakin tapauksissa, mutta sellaisiin on laajalla rintamalla vaikea päästä ja se johtaa monimutkaisiin riippuvuuksiin. Viime aikoina hajautettujen tietokantojen tekniikalla ja erilaisilla middleware -ratkaisuilla on pystytty käsittelemään hyvin hajautettua heterogeenista tietoa. Kuitenkin ratkaisujen rakentaminen on kestänyt kauan ja kustannukset ovat nousseet korkeiksi. XML -teknologia mahdollistaa edellä mainittua ratkaisumallia edullisemmän ja tehokkaamman tavan integroida eri alustoille kehitettyjä järjestelmiä.”

Valtiovarainministeriön työryhmäraportti 3/2005 ”Sanastot siltana saumattomalle tiedonvaihdon”

”Käytännön tarve ohjaa tekemään sanastotyötä pienin askelin siellä missä tarve on ajankohtainen. Tarvelähtöisyyden pitääkin olla ehdottomasti tärkein sanastotyötä ohjaava voima. Sanastotyön organisointi pelkästään yleisten sanastojen luomiseksi on tavallisesti resurssien tuhlaamista.”

Rekisteripoolin työ ja RAKU-pilotti

Perusrekisterijärjestelmän referenssiarkkitehtuuri-dokumentissa otetaan kantaa XML:n rooliin ja XML-muodossa tapahtuva kommunikointiin.

Tämän työn tekemisessä on tutustuttu myös rekisteripoolin RAKU-aineistoon. Hankkeessa on

alkuperäisten määrittelyjen jälkeen toteutetun pilotin (2003) kautta valmistunut palvelun määrittelyt ja työryhmä toteaa, että ”*palvelu voidaan teknisesti toteuttaa ja sen avulla voidaan tehostaa rakennusvalvonnan työtä sekä tarjota rakennusvalvonnan asiakkaille parempaa palvelua. Aikasäästöt voivat olla merkittäviä. RAKU-ratkaisua voidaan myös soveltaen hyödyntää useissa muissa tehtävissä.*”

Kansallinen paikkatietostrategia (MMM 10/2004)

Paikkatietoinfrastruktuurin kehittämistä on käsitelty paikkatietoasiain neuvottelukunnan laatimassa kansallisessa paikkatietostrategissa.

Perusrekisterilainsäädännön tarveselvitys (SM 2006/60)

Tämän työn jo ollessa käynnissä ilmestyi raportti Perusrekisterilainsäädännön tarveselvitys. Siinä yhtenä toimenpide-ehdotuksena on perusrekistereitä koskeva teknisen integraation kehittäminen:

”Luodaan ja rakennetaan valtion IT-strategian Yhteentoimivuuden kehittämissuunnitelmassa tietojen suurille julkishallinnon käyttäjille, erityisesti kunnille yhteisistä toiminnallisista ja teknisistä vaatimuksista lähtevä integraatoratkaisu, joka mahdollistaa perusrekisteritietojen käytön useista perusrekistereistä yhdellä tilauksella. Perusrekisterinpitäjät toteuttavat järjestelmiinsä ValtIT:n johdolla ratkaisun mukaiset rajapinnat ja KuntaIT vastaa ratkaisun mukaisten rajapintojen toteuttamisesta kuntien tietojärjestelmiin.”

4.2 Perustietovarantojen keskeisen sisällön ja tietopalvelurajapintojen dokumentaatio helposti saataville

Perustietovarantojen keskeinen sisältö ja käytettävissä olevat tietopalvelurajapinnat tulee dokumentoida ja dokumentaatio tulee saattaa helposti saataville. Tietoturvallisuus ei ole este perustietovarannon sisällön ja tietopalvelurajapintojen julkistamiselle nyt tavoitellussa laajuudessa. Tietojärjestelmien tietoturvamennettelyt eivät saa perustua tietoturvamennettelyjen ominaisuuksien salaamiseen, vaan niitä tulee voida arvioida julkisesti, mikä lisää luottamusta järjestelmän turvallisuuteen.

Selkeästi on tarve julkishallinnon sanasto- ja skeemahakemistolle. Näillä on lisäksi läheinen yhteys tekniseen palveluhakemistoon. SOA-palvelut, Web Services -palvelut, sanastot ja XML-skeemat muodostavat kokonaisuuden, joka pitäisi määritellä yhdessä. Työn tavoitteena tulee olla määritellä ja rakentaa sanastojen ja skeemojen hallinnan infrastruktuuri julkishallinnon tasolla.

4.3 Yhteiset koodistot ja sanastot

Julkishallinnon sanastotyötä tulee jatkaa. Perustietovarantojen ”perustietoyksiköt” ovat tässä keskeisessä asemassa. Tärkeimmät käsitteet tulisi kuvata ja niiden omistajuuden suhde virastokohtaisiin käsitteisiin määritellä.

Sanastotyötä pitää tehdä konkreettisella tasolla ja pyrkiä nopeasti selkeyttämään avainkäsitteitä. Tietojärjestelmien semanttisen yhteentoimivuuden ongelmia ja ratkaisuja on käsitelty kattavasti valtiovarainministeriön työryhmäraportissa 3/2005 ”Sanastot siltana saumattomalle tiedonvaihdolle”. Siinä todetaan muun muassa että ”*Sanastotyössä on usein lähdekkäisyys liikkeelle siitä huolimatta, että kansainvälisiä tai kansallisia valmiita sanastoja ei olisikaan olemassa. Käytännön tarve ohjaa tekemään sanastotyötä pienin askelin siellä missä tarve on ajankohtainen. Tarvelähtöisyyden pitääkin olla ehdottomasti tärkein sanastotyötä ohjaava voima. Sanastotyön*

organisointi pelkästään yleisten sanastojen luomiseksi on tavallisesti resurssien tuhlaamista.”

Useimmissa järjestelmissä on käytössä samat peruskoodistot. Osa koodistoista perustuu JHS-suosituksiin. Koodistoja ylläpitävät eri tahot. Tärkeimmät koodistot on kuvattu Tilastokeskuksen luokitusoppaassa (<http://tilastokeskus.fi/tk/tt/luokitukset/luokitusopas2004.html>). On tärkeää laatia koodistojen käsittelyyn yhtenäinen ohjeistus ja saattaa koodistot ja niiden päivitykset helposti saataville koneellisesti hyödynnettävässä muodossa. Ohjeistuksessa tulee ottaa kantaa myös koodiston muutosten yhtenäiseen käsittelyyn ja historiatiedon säilyttämiseen eri tietojärjestelmissä.

Koodistoihin toteutetaan kyselyrajapinnat samoin periaattein kuin mitä perustietovarantoihin ehdotetaan.

4.4 Sovellus-sovelluskyselyille yhtenäiset Web Services-rajapinnat

Perustietovarantojen Web Services- ja XML-rajapintojen kuvaamisessa ja toteuttamisessa pitäisi pyrkiä yhtenäisiin integraatio- ja SOA-linjauksiin. Ne luovat pohjaa muun muassa monille sähköisen asioinnin ratkaisuille. Siksi ne tulisi määrittellä ja toteuttaa ”malliratkaisuuina”.

Usein kyselyt ovat yksinkertaisia: etsitään esimerkiksi henkilötunnukselle nimeä, rakennukselle osoitetta tai rekisterinumeron perusteella auton haltijaa. Jokaista perustietovarantoa kohti kartoitetaan tyypilliset kyselyt. Tavoitteena on määrittellä kyselyjoukko, joka kattaa kyselyvolyymissä mitattuna valtaosan jokaisen tietovarannon kyselyistä. Näin saadulle kyselyjoukolle määrittellään vakio- tai muotoiset XML-sanomat. Sanomina kuvataan kyselysanoman muoto (pakolliset parametrit, oletusparametrit ja optiot) sekä siihen saatava vastaussanoma.

Jokaiseen perustietovarantoon tulee toteuttaa yhtenäisen määrittelyn mukainen Web Services-rajapinta määrittelyjen sanomien välittämiseen. Rajapinnan ominaisuuksia ovat ainakin siirtotien SSL-salaus (HTTPS), käyttäjän ja palvelimen luotettava varmennepohjainen tunnistaminen sekä kyselyn tekijän käyttöoikeuksien tarkastus. Niissä rajapinnoissa, joissa käsitellään henkilötietoa, otetaan huomioon lain henkilötietojen käsittelylle asettamat erityisvaatimukset ja rajoitukset.

Yksittäiskyselyitä varten tietovarantoihin toteutetaan web-pohjainen kyselyliittymä. Web-liittymät toteutetaan sovellus-sovellusyhteytenä Web Services-rajapintaa hyödyntäen.

4.5 Tietoluovutusten painopiste eräsiirrosta sovellus-sovellusyhteyksiin

Osa tiedonsiirtotarpeista on sellaisia, että eräsiirron käyttäminen on perusteltua. Tällaisia tapauksia ovat esimerkiksi jokin kaikkia kansalaisia koskeva rekisteri tai vaalien äänioikeutettujen rekisteri. On järkevää, että näissä tapauksissa eräsiirrot jäävät rinnalle mutta on tavoiteltava, että niiden määrä muuten tulee vähenemään huomattavasti.

Erityisesti silloin kun tiedonsiirto johtaa päällekkäisen erikseen ylläpidettävän rekisterikopion syntyyn (yleensä johtaa) tai epätarkoituksenmukaisen tietojoukon poimintaan (esimerkiksi kaikkien kuntalaisten tiedot poimitaan kunnan terveydenhuollon rekisteriin ”varmuuden vuoksi”), on pyrittävä tarpeen mukaan tapahtuviin tosiaikaisiin sovellus-sovelluskyselyihin.

Esitettyjen muiden kehitysehdotusten on määrä tukea tätä tavoitetta.

4.6 Siirrettyjen tietojen aikaleimojen määrittely ja yhtenäinen toteuttaminen

Siirretyille tiedoille, jotka asiakas tallentaa omaan järjestelmäänsä, määrittellään yksikäsitteisesti ymmärrettävät aikaleimat. Aikaleimoja on ainakin kolme. Ne kuvaavat tiedon syntymisen ajankohtaa, suositeltavaa käyttöaikaa, jonka jälkeen tietoa pitää yrittää virkistää, sekä ajankohtaa,

jolloin tieto ei ole enää käyttökelpoista ja sitä ei saa käyttää. Menettelyssä tavoitellaan Internetin DNS-nimipalvelun kaltaista mekanismia: kerran kysytyt tiedot voidaan tallettaa hyödyntävään järjestelmään ikään kuin välimuistiin ja niitä voidaan käyttää uudelleen kunnes suositeltava käyttöaika umpeutuu. Tämä jälkeen niitä pitää yrittää virkistää lähdejärjestelmästä, mutta jos se ei onnistu, välimuistissa olevaa tietoa voidaan käyttää niin kauan kuin se on aikaleiman mukaan poistettava.

(”tekopäivä”, ”parasta ennen”, ”viimeinen käyttöpäivä”)

4.7 Tietoturvaluksuusmenettelyjen kehittäminen

Jatkossa eräsiirtojen siirtotiedostot allekirjoitetaan ja tarvittaessa salataan digitaalisesti. Ottamalla käyttöön siirtotiedostojen digitaalinen allekirjoitus, voidaan yksinkertaisesti toteuttaa sekä siirtotiedoston alkuperän varmistaminen että ennen kaikkea varmistaa tiedon muuttumattomuus siirron aikana. Tarvittaessa siirrettävä tieto salataan. Samat periaatteet ovat sovellettavissa myös XML-sanomien siirrossa (XML-encryption ja XML-signature).

Tietojen käytön seuranta on kehitettävä siten, että tyyppillisesti käytetyn passiivisen kyselytapah- tumien koneellisen kirjaamisen rinnalle otetaan aktiivisia tapoja seurata lokikirjauksia koneellisesti niin, että esimerkiksi poikkeava käyttöprofiili voisi laukaista hälytyksen mahdollisesta väärinkäytöstä.

Käytön seurannan vaatimus on laajennettava kattavasti ja johdonmukaisesti koskemaan myös hyödyntäviä järjestelmiä niin, että luovutetun tiedon asiallinen käyttö on varmistettu myös niissä.

4.8 Tietopalvelun tekninen erottaminen perusjärjestelmästä

Perus- ja tietopalvelujärjestelmän erottaminen toisistaan on usein monin tavoin hyödyllistä. Erottamisella voidaan saavuttaa operatiivisen järjestelmän ja sen tietojen parempi suojaustaso. Erottamisella saadaan myös skaalautuvuutta ja redundanssia tietopalvelujärjestelmään.

Tietopalvelurajapinnat

LIITE 2

	Pääte	C-S	Web	So-So	Eräsiirto	Muuta
Väestötietojärjestelmä (Henkilö- ja rakennustietojärjestelmät)	3270 (poistuu)	Ei	On, korvaa pääteyhetydet	On (ei toistaiseksi massakäytössä)	Muutostiedot, Poiminnat (määrämuotoiset tiedostot, XML tulossa).	Paperituloste myös
Kiinteistötietojärjestelmä						
Kiinteistörekisteri	Ei	Sisäisessä käytössä	KTJ-selaintietopalvelu (ePortti), Ammatillaisen kartapaikka	Web-service sanomarakajapinta	(S)FTP aineistonvälityshakemistojen kautta. Muutokset kunnille XML-muotoisena. CD, irtolevyasemat.	XML laajasti käytetty, Paperitulosteina "rekisteriotteet"
Lainhuuto- ja kiinnitysrekisteri	Sisäisessä käytössä 3270	Ei	Sisäisessä käytössä		FTP aineistonvälityspalvelu	Tietopalvelukäyttö Kiinteistörekisterin kautta
Yritys- ja yhteisötiedot						
Kaupparekisteri (PRH)	Katka yksittäiskäyttö 3270 (poistuu)	Ei	Ei	Asiakaskohtaisesti toteutettavissa, uudessa vakiorajapinta, mm. Poliisi ja Tulli EBR:n kautta	FTP. Peruspoiminta muutoksista joko yö. CD-toimitukset loppuvat	Paperitulosteina "rekisteriotteet"
Yhdistysrekisteri	Katka yksittäiskäyttö 3270 (poistuu)	Ei	On	Haluttaessa asiakaskohtaisesti	Osoitepoiminta alustetaan www-liittymän kautta, FTP-siirto määrämuotoinen tiedosto. CD-toimitukset loppuvat	Paperitulosteina "rekisteriotteet"
Arvonlisävero ja työnantajarekisterit (VH)					Eräsiirto 1-2 krt vuodessa: levyke, CD, magneettinauha	Käyttö YTJ:n kautta, tietojen saanti myös suoraan mahdollista
Yritys- ja yhteisötietojärjestelmä YTJ	Ei	Ei	On	mm. poliisi EBR:n kautta		
Yritys- ja toimipaikkarekisteri			Yhteenvetotietojen tarkastelu		FTP. Kerran kuukaudessa määrämuotoinen tiedosto. CD:llä voidaan myös siirtää	
Ajoneuvoliikennerekisteri	Ei	Ei	On	Yksittäisen ajoneuvon tietojen kysely	FTP/XML: muutostietopalvelut päivittäin, kuukausittaiset poiminnat	
Kansallinen tie- ja katutietojärjestelmä Digiroad	Ei	Ei	Ylläpitoon kunnille		CD/DVD koko aineisto, 2007 pelkät päivitykset mahdollisia	
Paikkatiedot	Ei	Ei	On	Ei	FTP	
Verotustiedot					FTP	
Kansaneläkelaitoksen etuustiedot						

Lisäarvopalvelujen rajapinnat

LIITE 3

	Pääte	Web	So-So	Eräsiirto	Muu
Väestötietojärjestelmä		ePortti		ePortti, WM-data, TietoEnator	Kunnat, rakennustietopalvelut myös maistraatit
Kiinteistötietojärjestelmä					Kunnat, maistraatit
Kiinteistörekisteri		ePortti			
Lainhuuto- ja kiinnitysrekisteri		ePortti			
Yritys- ja yhteisötiedot					Maistraatit ja TE-keskukset
Kaupparekisteri (PRH)		ePortti, Asiakastieto	(EBR = European Business Register)	ePortti	
Yhdistysrekisteri		ePortti			Teleoperaattori WAP
Arvonlisävero ja työnantajarekisterit (VH)					
Yritys- ja yhteisötietojärjestelmä YTJ		ePortti, Asiakastieto	(EBR)	ePortti	Teleoperaattori WAP
Yritys- ja toimipaikkarekisteri		ePortti, Asiakastieto			YTJ:n kautta
Ajoneuvoliikennerekisteri		ePortti			Teleoperaattorit SMS/SoSo
Kansallinen tie- ja katutietojärjestelmä Digiroad					Useita – ei suoria liittymiä
Paikkatiedot		"Ei ole vastuutettu eri paikkatietovarantoja integroivan palvelun toteutusta"			
Verotustiedot		ePortti, Asiakastieto			
Kansaneläkelaitoksen etuustiedot					

Keskeisten lisäarvopalvelujen tarjoajien rajapinnat

Asiakastieto		Kyllä			PDF-tiedosto
Kauppalahden palvelu = ePortti		Kyllä	Tarvittaessa asiakaskohtaisesti	Kyllä	PDF-tiedosto