

JHS XXX

Paikkatiedon yksilöivät tunnisteet

JUHTA - Julkisen hallinnon tietohallinnon neuvottelukunta

Esa Tiainen

Maanmittauslaitos (MML)

Paikkatiedon yksilöivät tunnisteet -keskustelutilaisuus 14.4.2014

Paja-auditorio

Vähän historiaa aluksi...

- Yksilöivät tunnisteet vakiintuivat hakuavaimiksi tietojärjestelmissä relaatiotietokantojen myötä 1970-luvulta alkaen
- Jo tätä ennenkin tunnisteita käytettiin ns. indeksimuodossa tehostamaan tiedonhakua ja löytämistä verkkotietokannoissa
- Opittiin käyttämään myös surrogaattiavaimia, joilla oli pseudotunnisteita, vaikkapa tiedon anonymisointiin (henkilö-ym. tietosuoja)
- Internet-ajan vastineeksi edellä mainituille 2000-luvulla milleniumpalkinnon saaja Sir Tim Berners Lee lanseerasi Linked Data –käsitteen
- [Liittyköhän tämä ollenkaan päivän aiheeseen, katsotaanpa...](#)

Hankkeen perustamisen tausta

- Samaa kohdetta eri tietoresursseissa kuvaavia tietoja voidaan yhdistää yksilöivien tunnisteiden avulla
- Ne mahdollistavat kertaalleen kerätyn tiedon yhteiskäytön tehostamisen sekä päällekkäisten rekisterien ja tiedonkeruun vähentämisen
 - JHS-painopistealueet: Yhteisten tietovarantojen hyödyntäminen, Tietojärjestelmien yhteentoimivuus
- Inspire-direktiivin (2007/2/EY) toimeenpanon osana on kansallisten järjestelmien mukaisten tunnistekäytäntöjen yhdenmukaistaminen
- Aihe on JHS-työsuunnitelmassa

Inspire-direktiivi

- Yksilöivä tunniste
 - [Liitteen I](#) aineistoille joulukuuhun 2017
 - [Liitteen II](#) aineistoille n. vuoden 2020 loppuun mennessä
 - [Liitteen III](#) aineistoteemoille (21 kpl) ja paikkatietokohteille, joita useat käyttäjätahot tarvitsevat
 - Uusilla aineistoilla tiukemmat aikarajat alkaen v. 2015
- Tunnisteet julkaistaan HTTP URI -muodossa koostuen
 - paikkatietokohteen tietolähteen yksilöivästä nimiavaruudesta ja
 - paikkatietokohteen paikallisesta tunnisteesta, joka on tiedon tarjoajan antama tunniste

Semantiikkaa ja ontologioita

- Ontologiat ovat käsiteverkkoja, joiden avulla voidaan helposti tunnistaa, yksilöidä ja paikantaa internetissä samaan ilmiöön liittyvää tietoutta mm. internetin yleisten hakukoneiden kautta eri tietolähteissä.
- Esim. Googlen hakukoneet eivät kuitenkaan alunperin käyttäneet aitoja ontologioita, joskin Googlessakin on tehty ja tehdään edelleen kehitystä tähän suuntaan.
- Aineistoliitteiden tietotuotemäärittelyt muodostavat harmonisoidun käsitteistön, joka on yhteentoimivuuden perusta

Paikkatiedot ja semantiikka

- Paikkatiedon semantiikan hallintaan on kolme peruslähestymistä:
 - Harmonisointi
 - Yksilöivät tunnisteet (edellyttäen, että yhteinen käsitteistö eli controlled vocabulary on kiinnitetty = Inspire Data specifications)
 - Ontologiat (joko tukeutuen yhteiseen käsitteistöön = Inspire Data specifications tai kiinnittäen käsitteistö annotoinnin kautta jne., taikka näitä molempia, mikä on lähestymistapamme TSK-yhteistyössä mahdollistaen käsitteistön laajentamisen edelleen yhteisen käsitteistön perustalta).
- Käytännössä nämä kolme lähestymistapaa tukevat ja täydentävät toisiaan tietojen hyödyntämisessä tehokkaimalla mahdollisella tavalla, mikä nykyään on myös ehdoton vaatimus tiedon käyttöön saattamiselle laajalle koko yhteiskunnassa, mihin Linked Data tähtää

Viitekehys

- Yksilöivät tunnisteet sisältyvät myös julkisen hallinnon kokonaisarkkitehtuuriin ja paikkatiedon viitearkkitehtuuriin
 - sovellettava yhtenäistä käytäntöä ja rakennemallia
 - paikkatiedon kytkeminen paremmin muihin sovellusalueisiin
- Avoimen tiedon ohjelma 2013-2015, VM 17.5.2013:
 - ” suurimmat hyödyt ...saadaan, kun eri tietolajeja voidaan yhdistää toisiinsa”
- Linked data
 - avainteknologioita ovat URIt, http ja RDF
- EU ([ISA](#) – the Interoperability Solutions for European Public Administrations programme)
 - suosituksia tunnistekäytäntöjen (URI) yhdenmukaistamiseksi eri sovellusalojen kesken (2013)

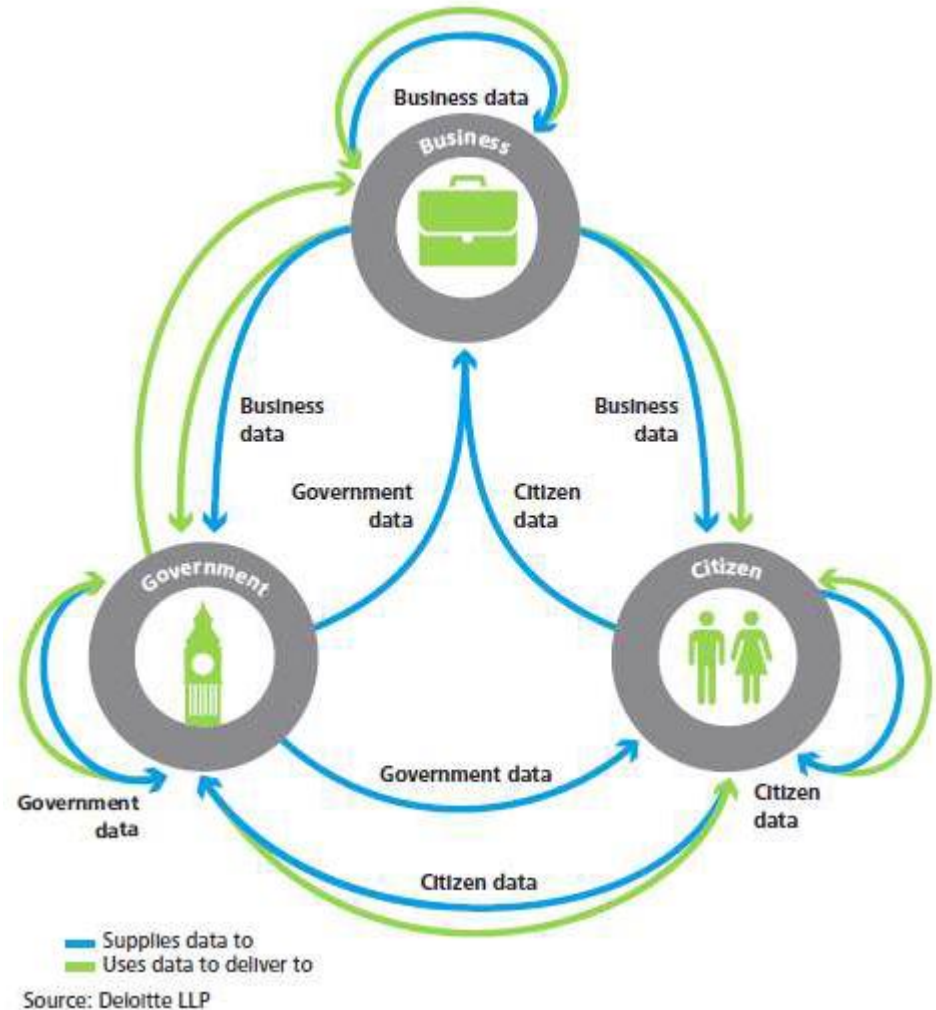
Hankkeen tavoitteet

- Määritellä ja vakiinnuttaa
 - paikkatiedon yksilöivien tunnisteiden yhtenäinen rakenne
 - niiden julkaiseminen koneluettavassa muodossa sekä
 - johdonmukaiset menettelyt uudelleenohjauksiin tunnisteita käytettäessä
- Inspire-vaatimusten sovittaminen olemassa oleviin tunnistekäytäntöihin
- URI-rakenteen sopiminen ja järjestäminen kansallisella tasolla, jolloin yhteiskunnan tietoresursseista saadaan suurimmat hyödyt
- URI-tunnisteiden julkaiseminen ilmentymätason lisäksi myös käsitetasolla, jotta käsitteiden avulla voidaan hakea niihin liittyvää paikkatietoa ilmentymätasolla; yhteentoimivuus- ja Linked Data-tavoitteiden toteuttaminen

Hyödyntäminen

- Tiedon keruu
 - uudelleenkäyttö eri tietolähteistä
 - tiedon ylläpito
 - muutostietojen ylläpito
- Tietojen yhdistely
 - analyysikäyttö
 - monitorointi
 - abstraktien kohteiden luonti
- Tiedon jakelu, tietopalvelut
 - kohteista kiinnostuneet tahot
- Toiminnan ja prosessien kehittäminen
 - primaaritiedon uudelleenkäyttö
- Tieto- ja kokonaisarkkitehtuurit
 - tuottavuus, tehokkuus
- Tietomäärittely, sovelluskehitys
 - nopea kehityssykli
 - ylikansallinen käyttö
- Yhteiskäyttö, sähköinen asiointi
 - julkinen, yksityinen, kansalaiset
- Crowdsourcing
 - tietojen ylläpito
 - kohdetasoinen palaute (avoin tieto)

Figure 1. The open data ecosystem



Soveltamisala

- Inspire-aineistot muodostavat laajan kansallisen paikkatietoinfrastruktuurin perustavan, ympäristöön mahdollisesti vaikuttavia politiikkoja tai toimintoja tukevan ja harmonisoidun tietoresurssin, ja
- kattavat useiden perusrekistereiden sisältöjä
- Julkishallinnon kokonaisarkkitehtuuri
- Avoin tieto
 - Yritykset, yhteisöt, kansalaiset (palautetieto, crowdsourcing)
- Kansallinen ratkaisumalli

JHS-työryhmä

- Geodeettinen laitos
- Geologian Tutkimuskeskus
- Kansalliskirjasto
- Maanmittauslaitos
- Suomen Kuntaliitto
- Suomen ympäristökeskus
- Tampereen kaupunki
- Valtiovarainministeriö
- Mahd. täydennystä: Liikennevirasto...
- Editorit: Poscon Oy, Maanmittauslaitos

Aikataulu

- Aikataulutettu kahdelle vuodelle 2013-2015
 - Jälkimmäinen pääosin laajaan palautekierrokseen
 - Yhteistyö: Kansalliskirjaston ONKI-projekti, Linked Open Data Finland (Tekes)
- Varaus teknisen asiantuntemuksen vahvistamiseen
 - URI-resoluutiopalvelu
 - Käyttöönottosuunnitelman tekniset kysymykset

Kiitos mielenkiinnosta!

- Hankevalmistelun vetäjä

Esa Tiainen

esa.tiainen(at)nls.fi

Johtava asiantuntija

Maanmittauslaitos/Inspire-sihteeristö