



Suomen historialliset paikat ja kartat avoimena linkitettyinä datana

Eero Hyvönen, Esko Ikkala, Jouni Tuominen, Eetu Mäkelä
Aalto-yliopisto, Semanttisen laskennan tutkimusryhmä (SeCo)
Hankkeen kotisivu: <http://www.seco.tkk.fi/projects/histoplaces>

Hipla.fi on prototyyppi muistiorganisaatioissa ja kulttuurialan tutkimuksessa tarvittavasta historiallisten paikkojen kansallisesta nimistö- ja karttapalvelusta. Uuden palvelukonseptin innovaationa on yhteisöllinen historiallisen paikka- ja karttatiedon keruu ja ylläpito: tavoitteena on kansallisen avoimen datan pysyvä kartuttaminen keskitetyssä verkkopalvelussa, jossa datan keruuta voidaan integroida osaksi muistiorganisaatioiden normaalia luettelointi- ja sisälönkuvailutoimintaa. Palvelun ydinaineistoina olevat historialliset paikka- ja kartta-aineistot tulevat useista kotimaisista avoimen datan lähteistä ja yhteistyöverkoston organisaatioilta. Palvelu sisältää myös 800 000 nykyistä paikkaa Suomessa. Verkkopalvelun toteutuksessa on hyödynnetty uusinta semanttisen webin teknologiaa sekä kansallisten FinnONTO- ja Linked Data Finland -hankkeiden avointa koodia ja ontologioita. Järjestelmä on ensimmäinen laatuaan maailmassa.

1 MIKSI HIPLA.FI TARVITAAN?

1.1 Historiallisen paikannimistön ja karttojen tarve

Kun muistiorganisaatiot, museot, kirjastot ja arkistot, tallentavat kokoelmiinsa objekteja (tangible cultural heritage), kuten valokuvia, artikkeleita, esineitä, karttoja, maalauksia, kirjoja, tai perinnetietoa (intangible cultural heritage), sisällöt pitää kuvailla metatietona, jotta ne voidaan myöhemmin löytää. Metatieto on tärkeää myös kokoelmien hallinnassa (provenience) ja kokoelmatietojen yhdistämisessä kansallisissa ja kansainvälisissä verkkopalveluissa, kuten Finna¹, Kulttuurisampo², Kirjasampo³ tai yleiseurooppalainen Europeana⁴ ja USA:n Digital Public Library of America (DPLA)⁵. (Hyvönen, 2012) Näin esimerkiksi voidaan löytää talvisodan aikaan otettu valokuvan Olavi J. Paavolaisesta Kivennavan Vienolassa, sijoittaa kuvan kohde kartalle paikkaan, jossa Vienola-nimistä paikkaa ei enää ole olemassa, ja yhdistää kuva kirjailijan tuotantoon Kirjasammossa, Kivennavan pitäjän historiaan ja paikan muihin nimiin eri kielillä jne. Tärkeä käyttäjäryhmä historiallisten paikkojen nimistöille (historical gazetteers) ovat kulttuurin harrastajat, tutkijat ja historioitsijat (Southall et al., 2011): "historians often need specialised gazetteers listing places that no longer exist and names that are no longer used or whose spelling has significantly altered".

¹ <http://www.finna.fi>

² <http://www.kulttuurisampo.fi>

³ <http://www.kirjasampo.fi>

⁴ <http://europeana.eu>

⁵ <http://dp.la/>

Historiallisiin paikannimiin liittyvä yksi hankaluus on niiden näyttäminen nykyisillä kartoilla. Tähän tarvitaan usein historiallisia kartoja, koska vanhoja paikkoja ei uusilta kartoilta välttämättä löydy tai ne ovat merkittävästi muuttuneet. Paikan yksilöimiseksi tarvitaan nimen ja kartan ohella usein muutakin kontekstietoa (Ikkala et al., 2015). Vanhojen paikannimien ontologisena haasteena taas on ajassa elävän paikan käsitteen ja nimien merkitysten muuttuminen (Hyvönen et al, 2011). Mitä tarkoittaa esimerkiksi "Saksa" eri aikoina?

1.2 Koordinoitua yhteistyötä tarvitaan

Suomessa Maanmittauslaitoksella (MML) on lakisääteinen tehtävä ylläpitää kansallista paikannimirekisteriä tällä hetkellä käytössä olevasta nimistöstä, joka on jokin aika sitten julkaistu avoimena datana kaikkien käytettäväksi⁶. Rekisteri sisältää n. 800 000 paikannimeä Suomessa käytetyillä kielillä, joita ovat suomi, ruotsi, sekä saamen kielen kolme eri murretta. Historiallisten paikannimien osalta tällaista velvoitetta ei ole kenelläkään, eikä mistään lähteestä historiallista paikannimistöä voi kootusti etsiä. Monilla organisaatioilla, kuten esimerkiksi Kotimaisten kielten keskus (Kotus), Suomalaisen Kirjallisuuden Seura, Svenska litteratursällskapet, Museovirasto, Kansallisarkisto, Kansalliskirjasto, museot, sukuseurat jne. on kuitenkin runsaasti paikkasidonnaista historiallista tietoa ja tietoa paikoista, joskin tämä tieto on monasti osin puutteellista, epätarkkaa tai digitoimatta. Kotuksessa on käynnissä laaja historiallisten paikannimien digitointihanke - keskuksen arkistossa on peräti 2,7 miljoonaa nimilappua selityksineen. Paikannimet ja paikan käsite (Tuan, 1975) liittyvät kulttuurin kentässä lähes kaikkeen tietoon.

Tässä artikkelissa esiteltävän Hipla.fi-palvelun kunnianhimoisena tavoitteena on olla samalla kertaa organisatorinen, tekninen ja ontologinen ratkaisumalli historiallisten paikkojen nimistön ja karttojen julkaisemiseksi avoimena datana ja palveluna verkossa. Palvelun asiakasrajapinnan käyttäjiä ovat ihmiskäyttäjät ja data-rajapintojen hyödyntäjiä toiset verkkopalvelut ja sovellukset sekä näiden ohjelmoijat.

Hipla.fi-palvelusta on syksyllä 2015 valmistunut Aalto-yliopiston Semanttisen laskennan tutkimusryhmässä (SeCo) ensimmäinen demonstraattori (Hyvönen et al., 2015a; Ikkala et al., 2015) Suomen Kulttuurirahaston rahoituksen turvin. Seuraavassa esitetään lyhyt kuvaus palvelun asiakasrajapinnasta sekä tämän perustana olevasta avoimen linkitetyn datan (Linked Data) (Heath, Bizer, 2011) palvelusta, jolla Hipla.fi:n aineistoja voidaan hyödyntää erilaisissa sovelluksissa. Työn jatkoa SeCo-ryhmän ONKI-ontologiapalvelukonseptille (Viljanen et al., 2009), jonka Kansalliskirjasto on tuotteistanut nykyiseksi Finto.fi-palveluksi.

2 KÄYTTÖ ASIAKASRAJAPINNAN KAUTTA

Asiakasrajapinnan tavoitteena on: 1) Mahdollistaa historiallisten paikkojen hakeminen eri data-varannoista. 2) Mahdollistaa tulosten näyttäminen ja tutkiminen sekä nykyisillä että historiallisilla kartoilla. 3) Tarjota lukijoille paikkoihin liittyvää kontekstietoa muista avoimen datan lähteistä. 4) Mahdollistaa uusien paikkojen yhteisöllinen ehdottaminen ja käyttö palvelun kautta reaaliajassa.

Kuva 1 esittää järjestelmän asiakasrajapintaa osoitteessa <http://hipla.fi>. Ylhäällä vasemmalla on lista avoimista linkitetyn datan paikkatietolähteistä (SPARQL-endpoint), jotka on tuotettu palvelun ensimmäiseen pilottisovellukseen:

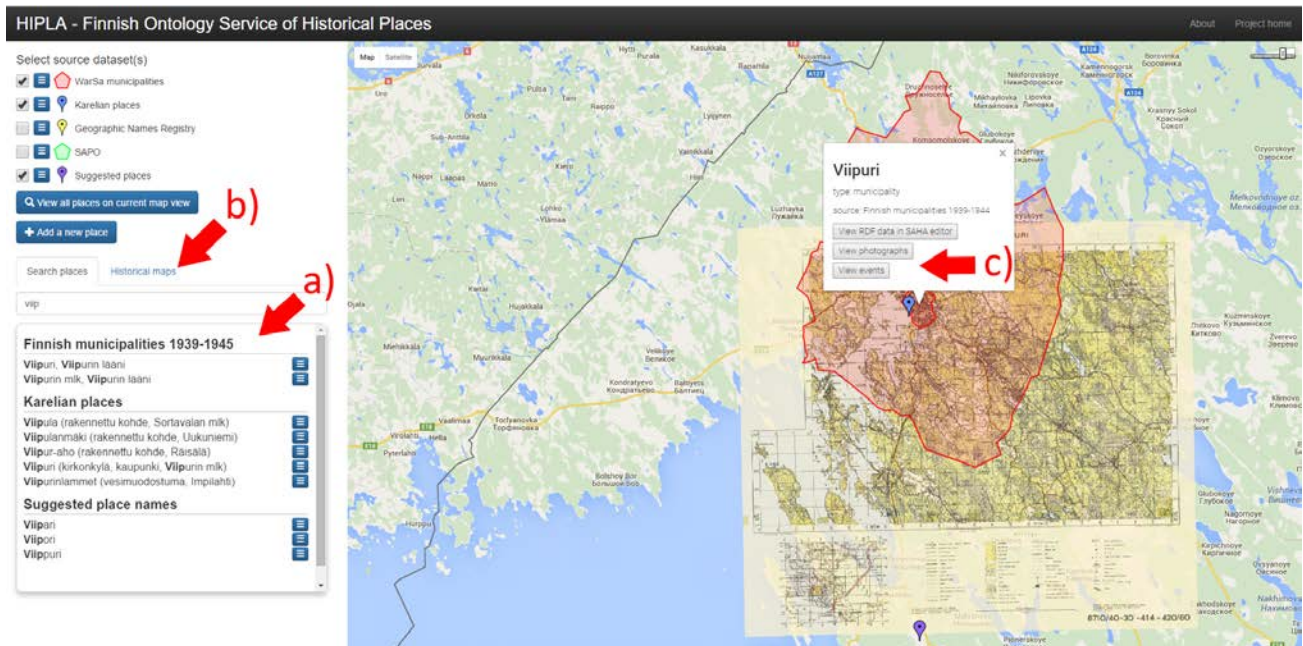
1. Sotasampo.fi-palvelun⁷ käyttämät historialliset kunnat.
2. Luovutetun Karjalan 35 000 paikkaa⁸, jotka on jalostettu RDF-muotoon ja WGS84 koordinaattijärjestelmään
3. Maanmittauslaitoksen avoin paikannimirekisteri; n. 800 000 nykyistä paikkaa, jotka on muutettu linkitetyksi avoimeksi dataksi.
4. Suomen ajallinen kuntahistoria SAPO (Hyvönen et al., 2011) FinnONTO-projektista.

⁶ <http://www.maanmittauslaitos.fi/digituotteet/nimisto>

⁷ <http://ww.sotasampo.fi>

⁸ Karjalan historialliset paikat on alunperin geokoodannut Jyrki Tiittanen.

5. Käyttäjien ehdotukset uusiksi historiallisiksi paikoiksi. Demonstraattorissa kuka tahansa voi lisätä paikkoja palveluun palvelukonseptiin tutustumista varten.



Kuva 1. Hipla.fi-palvelun asiakaskäyttöliittymä.

Haku tehdään samanaikaisesti kaikkiin käyttäjän valitsemiin tietolähteisiin check box -laatikoiden alla olevan automaattitäydennyskentän avulla (b). Kuvassa käyttäjä on kirjoittanut kenttään ”viip” ja tuloksena on viip-alkuiset paikat kolmesta datalähteestä (a) lisätiedoilla varustettuna tunnistamista varten. Lisätietoja ovat esimerkiksi laajempi kunta/lääni ja paikkatyyppi, kuten ”rakennettu kohde”. Listalta valittu kohde on merkitty Googlen kohde-merkillä (marker) (c) kartalle, jossa on näkyvissä myös liittyvä paikkaa liittyvä tietoikkuna (info window). Siinä on paitsi kerrottu paikan nimi myös annettu paikkaan liittyvää kontekstietoa, tässä tapauksessa toisen maailmansodan aikaisia tapahtumia (12 437 kpl), valokuvia puolustusvoimien avoimesta SA-kuva-arkistosta (160 000 kuvaa) sekä Kansa taisteli -lehtien artikkeleita (3 360 kpl). Nämä linkittyvät Hipla.fi:n URI-paikkatunnisteisiin Sotasampo.fi-palvelun avoimen datapalvelun kautta.

Haku-valintalehden (tab) ohella käyttäjä voi hakea ja valita historiallisia karttoja toisella valintalehdellä (b). Historialliset kartat on asemoitu Googlen karttojen päälle osin läpinäkyvästi, jolloin voi samalla kertaa tarkastella sekä historiallisia että nykyisiä paikkoja. Kuvassa käyttäjä on valinnut kartan, jolla valittu paikka Viipuri näkyy. Kartalta löytyy mm. Monrepos’n puisto ja muuta vanhaa suomalaista nimestöä, joita voi verrata nykyiseen venäjän kieliseen nimestöön.

Karttojen aseointi ja näyttäminen Googlen kartoilla eri tarkkuustasoilla (zoom-tasot) tapahtuu erillisen MapWarper -palvelun avulla. Se on käytettävissä ONKI-palvelun aliosoitteessa <http://mapwarper.onki.fi>. Järjestelmä perustuu New York Public Libraryn alun perin kehittämään Open Source -ohjelmistoon. Sen avulla projektiryhmä on yhteistyössä Suomen Wikimedia ry:n kanssa aseoinut kaikki Karjalan topologiset kartat 1:100 000 sekä vanhemman Kansallisarkiston Senaatin karttojen karttakokoelman Hipla.fi-palveluun, yhteensä n. 460 karttaa. Senaatin kartasto kattaa Karjalaa laajemman osan Suomea. Uusien karttojen aseointi ei MapWarperia käytettäessä vaadi ohjelmointiosaamista.

3 KÄYTTÖ LINKITETYN DATAN PALVELUNA

Hipla.fi:n datapalvelun tavoitteena on, että sen tarjoamia historiallisia paikka- ja kartta-aineistoja voidaan hyödyntää linkitettyinä datana muissa sovelluksissa. Lisäksi järjestelmässä kehitetty linkitetyn datan URI-järjestelmä mahdollistaa eri datasiiloissa olevien aineistojen sisällöllisen rikastaminen toistensa avulla. Näin voidaan esimerkiksi Puolustusvoimien SA-kuva-arkiston valokuvat ja Kansallisarkiston kokoelmista saatava tieto joukko-osastojen toiminnasta ja taisteluista sijoittaa kartalla ja liittää vielä kuvat ja toiminta toisiinsa automaattisesti aika- ja paikkatietoa yhdistämällä (Hyvönen et al., 2015b).

Datan hyödyntämisen lisäksi myös Hipla.fi:n omia käyttöliittymäkomponentteja voidaan käyttää uudelleen URI-järjestelmän kautta. Esimerkki tällaisesta sovelluksesta on Sotasampoon sisältyvä paikanäkymäsovellus⁹, jossa toiseen maailmansotaan liittyvät paikat voidaan vaivattomasti näyttää loppukäyttäjälle Hipla.fi-palvelun HTML-komponenttien avulla kartoilla, eräänlaisina historiallisten paikkojen ”kotisivuina”. Esimerkiksi Suomen kunnat 1939–44 datajoukon Viipurin kuntaan liittyvä ”kotisivu” on osoitteessa

http://www.sotasampo.fi/places?uri=http://ldf.fi/warsa/places/municipalities/m_place_614

Tässä osoitemuodossa HTTP-osoitteen URI-parametrina annetaan Linked Data Finland (ldf.fi) -nimiavaruuden Sotasampo-palvelun (warsa) paikka, jonka Sotasampo.fi-sovelluksen paikkasovellus (places) lukee ja näyttää loppukäyttäjälle Sotasammon omana sivuna, johon Hipla.fi:n tuottama HTML-aineisto on upotettu.

Avoim datajoukko	SPARQL-endpoint	Palvelun RDF-graafi	URI-esimerkki
Suomen kunnat 1939–44	http://ldf.fi/warsa/sparql	http://ldf.fi/warsa/places/municipalities	http://ldf.fi/warsa/places/municipalities/m_place_509
Karjalan karttanimistö 1922–44	http://ldf.fi/warsa/sparql	http://ldf.fi/warsa/places/karelian_places	http://ldf.fi/warsa/places/karelian_places/k_place_48
Paikannimirekisteri	http://ldf.fi/pnr/sparql	http://ldf.fi/pnr/	http://ldf.fi/pnr/P_10342733
Suomen ajallinen paikkaontologia (SAPO)	http://ldf.fi/hipla/sparql	http://www.yso.fi/onto/sapo	http://www.yso.fi/onto/sapo/Helsinki(1946-1965)
Käyttäjien paikkaehdotukset	http://ldf.fi/demo/sparql	http://ldf.fi/demo/hipla-suggestions	http://ldf.fi/demo/hipla-suggestions/ff998be1-8d24-4a38-9383-9cdc88c797a9
Karjalan kartat 1:100 000 topografinen	http://ldf.fi/hipla/sparql	http://ldf.fi/hipla/maps/karelian_maps/topographic_100000	http://ldf.fi/hipla/maps/karelian_maps/topographic_100000/map_4022
Senaatin kartasto 1:20 000 topografinen	http://ldf.fi/hipla/sparql	http://ldf.fi/hipla/maps/senate_maps/topographic_20000	http://ldf.fi/hipla/maps/senate_maps/topographic_20000/map_1156163

Taulukko 1. Hipla.fi-palvelun datajoukkoja.

Hipla.fi käyttää datan julkaisualustana Linked Data Finland -portaalia¹⁰. Ensimmäisen pilottisovelluksen käyttämät datajoukot on esitetty taulukossa 1. Palvelun avoin rajapinta perustuu W3C:n standardoimaan SPARQL-kyselykieleen. Linkitetyn datan periaatteiden mukaisesti jokaiselle paikalle ja kartalle annetaan yksilöivä tunniste (URI), joka mahdollistaa aineistoihin viittaamisen ja niihin linkityksen. LDF-portaalin ohjesivuilta löytyy tarkempaa tietoa URI-käytännöistä¹¹ sekä rajapinnan käytöstä ja aineistojen lataamisesta¹².

4 HIPLA.FI:N UUTUUSARVO

Kansainvälisten historiallisten paikkojen sanastoista tunnetuin lienee amerikkalaisen Getty säätiön ylläpitämä Thesaurus of Geographical Names (TGN)¹³, jota käytetään laajasti USA:ssa ja monissa muissa maissa, esimerkiksi Hollannissa. Siinä on nykyisin lähes 1,5 miljoonaa historiallista paikkaa ja data julkaistiin v. 2015 avoimena linkitetynä datana¹⁴. Suomi ei tässä laajassakaan sanastossa ole juurikaan edustettuna. Pelagios¹⁵ on yhteisöllinen kansainvälinen hanke kulttuurin tutkimuksessa tarvittavien paikannimien ja tiedon tuottamiseksi ja ylläpitämiseksi. Vaiheissa 1 ja 2 hanke keskittyi erityisesti antiikin Kreikan ja Rooman paikkoihin ja työtä rahoitettiin Englannissa. Vaiheessa 3 fokus on siirtynyt keskiajan kristillisiin, arabialaisiin ja aasialaisiin paikkoihin rahoittajana Carnegie-

⁹ <http://www.sotasampo.fi/places/>

¹⁰ <http://www.ldf.fi/>

¹¹ <http://www.ldf.fi/uri-data-services.html>

¹² <http://www.ldf.fi/sparql-services.html>

¹³ <http://www.getty.edu/research/tools/vocabularies/tgn/>

¹⁴ <http://vocab.getty.edu/>

¹⁵ <http://pelagios-project.blogspot.fi/p/about-pelagios.html>

Mellonin säätiö. Pleiades¹⁶ on antiikin muinaiseen paikannimistöön keskittyvä hanke ja paikannimirekisteri. Sekä Pelagios- että Pleiades-hankkeissa hyödynnetään yhdistetyn avoimen tiedon menetelmiä, kuten Hipla.fi:ssä. Kansallista historiallisten paikkojen palveluista tunnetuin on The Historical Gazetteer of England's Place-names¹⁷. Palvelussa on n. 4 miljoonaa historiallista paikkaa, ja se perustuu perinteiseen teknologiaan.

Hipla.fi on maailman ensimmäinen kansallinen historiallisten paikannimien palvelu, jossa pyritään yhteisöllisesti kokoamaan rekisteriä eri aikakausien nimistöstä ja näiden välisistä semanttisista suhteista integroimalla palvelu osaksi organisaatioiden normaalia luettelointi- ja tutkimustoimintaa. Uutta on myös se, että palveluun sisältyy paikkatiedon ohella myös historiallisten karttojen datapalvelu, ja että asiakasrajapinnan ohella Hipla.fi on linkitetty avoimen datan infrastruktuuripalvelu erilaisten sovellusten hyödynnettäväksi.

Kiitokset

Kiitos Jyrki Tiittaselle geokoodatusta Karjalan paikkadatatista. Hipla.fi:n kartat asemoi MapWarperilla Hanna Hyvönen. Tutkimustyötä rahoitti Suomen Kulttuurirahasto ja karttojen asemointia Suomen Wikimedia ry:n Wikimaps-projekti Susanna Änäsin johdolla.

KIRJALLISUUTTA

1. Tom Heath, Christian Bizer: *Linked Data: Evolving the Web into a Global Data Space*. Morgan & Claypool, Palo Alto, CA, 2011.
2. Eero Hyvönen: *Publishing and Using Cultural Heritage Linked Data on the Web*. Morgan & Claypool, Palo Alto, CA, 2012.
3. Eero Hyvönen, Jouni Tuominen, Esko Ikkala and Eetu Mäkelä: *Ontology Services Based on Crowdsourcing: Case National Gazetteer of Historical Places*. Proceedings of 14th International Semantic Web Conference 2015 (ISWC 2015), Posters and Demos, CEUR-WS Proceedings, Bethlehem, PA, USA, October, 2015.
<http://www.seco.tkk.fi/publications/2015/hyvonen-et-al-hipla.pdf>
4. Eero Hyvönen, Jouni Tuominen, Eetu Mäkelä, Jérémie Dutruit, Kasper Apajalahti, Erkki Heino, Petri Leskinen and Esko Ikkala: *Second World War on the Semantic Web: The WarSampo Project and Semantic Portal*. Proceedings of 14th International Semantic Web Conference 2015 (ISWC 2015), Posters and Demos, CEUR-WS Proceedings, Bethlehem, PA, USA, October, 2015.
<http://www.seco.tkk.fi/publications/2015/hyvonen-et-al-warsampo.pdf>
5. Eero Hyvönen, Jouni Tuominen and Tomi Kauppinen, Jari Väättäinen: *Representing and Utilizing Changing Historical Places as an Ontology Time Series*. Luku teoksessa *Geospatial Semantics and Semantic Web: Foundations, Algorithms, and Applications* (Naveen Ashish and Amit Sheth (eds.)), Springer-Verlag, 2011.
6. Esko Ikkala, Jouni Tuominen and Eero Hyvönen: *Contextualizing Historical Places in a Gazetteer by Using Historical Maps and Linked Data*. October, 2015. Submitted.
<http://www.seco.tkk.fi/publications/>
7. Humphrey Southall, Ruth Mostern, and Merrick Lex Berman: *On historical gazetteers*. *International Journal of Humanities and Arts Computing* 5 (2), pp. 127–145, 2011.
8. Yi Fu Tuan, *Place: An Experiential Perspective*, *Geographical Review*, 65 (2), pp. 151-165, 1975.
9. Kim Viljanen, Jouni Tuominen, and Eero Hyvönen. *Ontology libraries for production use: The Finnish ontology library service ONKI*. Proceedings of the 6th European Semantic Web Conference (ESWC 2009). Springer-Verlag, 2009. <http://www.seco.tkk.fi/publications/2009/viljanen-et-al-www.yso.fi-2009.pdf>

¹⁶ <http://pleiades.stoa.org/>

¹⁷ <http://www.placenames.org.uk>