

# Suomalainen historiallisten paikkojen ja karttojen ontologiapalvelu

Esko Ikkala

## Sähkötekniikan korkeakoulu

Diplomityö, joka on jätetty opinnäytteenä tarkastettavaksi diplomi-insinöörin tutkintoa varten Espoossa 3.8.2016.

### Työn valvoja:

Prof. Eero Hyvönen

### Työn ohjaaja:

FM Jouni Tuominen

Tekijä: Esko Ikkala		
Työn nimi: Suomalainen historiallisten paikkojen ja karttojen ontologiapalvelu		
Päivämäärä: 3.8.2016	Kieli: Suomi	Sivumäärä: 7+63
Tietotekniikan laitos		
Professori: Mediatekniikka		
Työn valvoja: Prof. Eero Hyvönen		
Työn ohjaaja: FM Jouni Tuominen		
<p>Historiallinen paikkatieto on keskeisessä asemassa muistiorganisaatioiden kokoelmien hallinnassa ja hyödyntämisessä sekä digitaalisten ihmistieteiden tutkimuksessa. Paikkatiedon käsitteleminen muissa kuin erikoistuneissa paikkatietojärjestelmissä sekä paikkatiedon ajallinen ulottuvuus tuovat mukanaan lukuisia haasteita, joihin linkitetyn datan teknologiat ovat tarjonneet lupaavia ratkaisuja. Tässä työssä esitellään kulttuurialan organisaatioiden tarpeeseen kehitetty uusi linkitetyn datan teknologioihin perustuva historiallisten paikkojen ja karttojen palvelumalli, HIPLA. HIPLA-palvelumallin tavoitteena on tarjota yhteinen näkymä eri organisaatioiden hallinnoimaan paikkatietoon ja mahdollistaa hajautettujen paikkatietoaineistojen yhteisöllinen täydentäminen, haku ja selailu sekä nykyisillä että historiallisilla kartoilla. Lisäksi työssä toteutettiin HIPLA-palvelumallin etuja havainnollistava prototyypisovellus Hipla.fi, jota pilotoitiin osana talvi- ja jatkosodan aineistoja linkitettyinä avoimena datana julkaisevaa Sotasampo-projektia. Pilotoinnin tuloksena syntyi talvi- ja jatkosodan paikkaontologia, joka tarjoaa työkalun sotiin liittyvien aineistojen automaattiselle linkitykselle ja aineistojen maantieteelliselle visualisoinnille.</p>		
Avainsanat: historiallinen paikkatieto, paikkaontologia, yksikäsitteistäminen, georeferointi, linkitetty data, RDF, SPARQL		

Author: Esko Ikkala

Title: Finnish Ontology Service of Historical Places and Maps

Date: 3.8.2016

Language: Finnish

Number of pages: 7+63

Department of Computer Science

Professorship: Media technology

Supervisor: Prof. Eero Hyvönen

Advisor: M.Sc. Jouni Tuominen

Historical geographic data plays an important part in the management of cultural heritage data and also in digital humanities research. This brings forth problems such as managing the spatial dimension of geographic data, and the fact that traditionally geospatial data has been isolated in geographic information systems. Linked data technologies are a promising approach to these problems. This thesis presents a new Linked Open Data brokering service model HIPLA for using and maintaining distributed historical place gazetteers and maps. HIPLA provides a common access to historical geographic data and facilitates collaborative maintenance of geo-ontologies and maps in real time as a side effect of annotating contents in legacy cataloging systems. Hipla.fi, the first prototype implementation of the HIPLA model, was created to test and demonstrate the HIPLA model. The Hipla.fi user interface and data service were also applied in creating a semantic portal of historical Second World War data.

Keywords: historical geographic data, place ontology, disambiguation, georeferencing, linked data, RDF, SPARQL

## Esipuhe

Ensiksi haluan kiittää professori Eero Hyvöstä ja ohjaajaani Jouni Tuomista kannustavasta ja asiantuntevasta ohjauksesta. Kiitos myös kaikille muille Semanttisen laskennan tutkimusryhmän jäsenille, joiden kanssa tämän diplomityön ratkaisuja kehitettiin ja pilotoitiin yhteistyössä osana Sotasampo-projektia.

Jyrki Tiittaselle suuri kiitos Karjalan karttanimistön keräyksestä, aineiston luovutuksesta vapaaseen käyttöön sekä neuvoista aineiston suhteen. Historiallisten karttalehtien georektifoinnista tunnustus kuuluu Wikimedia Suomi ry:lle, Susanna Änäsille ja Hanna Hyvöselle.

Lopuksi erityinen kiitos avovaimolleni Sannalle kannustuksesta ja käytännön tuesta diplomityön valmiiksi saattamisessa sekä äidilleni Hilikalle korvaamattomasta lapsenhoitoavusta. Ilman lähipiirin tukea tätä työtä tuskin olisi syntynyt.

Espoossa, 3.8.2016

Esko Ikkala

# Sisällysluettelo

Tiivistelmä	ii
Tiivistelmä (englanniksi)	iii
Esipuhe	iv
Sisällysluettelo	v
Lyhenteet	vii
<b>1 Johdanto</b>	<b>1</b>
<b>2 Digitaalinen paikkatieto</b>	<b>4</b>
2.1 Paikkaviittaukset . . . . .	4
2.2 Sijainti ja sijaintiin liittyvä tieto . . . . .	5
2.3 Georeferoinnin haasteet . . . . .	5
2.4 Nimistöluettelot . . . . .	7
2.5 Maantieteellinen tiedonhaku . . . . .	7
2.6 Ajallisen ulottuvuuden tuki paikkatiedon standardeissa . . . . .	8
2.7 Paikkatiedon visualisointi . . . . .	9
2.8 Linkitetty paikkatieto . . . . .	9
<b>3 Historiallisen paikkatiedon palvelut ja lähteet</b>	<b>13</b>
3.1 Perinteiseen teknologiaan perustuvat palvelut ja lähteet . . . . .	13
3.2 Linkitettyä dataa hyödyntävät palvelut ja lähteet . . . . .	15
3.3 Yhteenveto historiallisen paikkatiedon palveluista . . . . .	17
<b>4 Historiallisten paikkojen ja karttojen palvelumalli</b>	<b>18</b>
4.1 Visio . . . . .	18
4.2 Vaatimukset . . . . .	18
4.3 Käyttäjärühmät ja käyttötapaukset . . . . .	20
4.4 Arkkitehtuuri . . . . .	22
4.5 Historiallisen paikkatiedon harmonisointi . . . . .	23
4.6 Kehittyvän ontologian hallinnointiprosessi . . . . .	24
<b>5 Palvelumallin prototyyppitoteutus: Hipla.fi</b>	<b>26</b>
5.1 Käyttöliittymän yleiskuvaus . . . . .	26
5.2 Automaattisesti täydentyvä ja federoitu tekstihaku . . . . .	28
5.3 Karttahaku . . . . .	29
5.4 Semanttinen haku paikkakäsitteillä . . . . .	29
5.5 Georektifioitujen historiallisten karttalehtien hyödyntäminen . . . . .	30
5.6 Kontekstitiedon tarjoaminen muista avoimen linkitetyn datan lähteistä	31
5.7 Puuttuvien paikkojen yhteisöllinen ehdottaminen . . . . .	32
5.8 Tekninen toteutus . . . . .	33
5.9 Datalähteet . . . . .	37

<b>6</b>	<b>Hipla.fi-palvelun pilotointi Sotasampo-projektissa</b>	<b>40</b>
6.1	Talvi- ja jatkosodan paikkaontologia . . . . .	40
6.2	Sotasammon aineistojen linkittäminen talvi- ja jatkosodan paikkaontologiaan . . . . .	48
6.3	Sotasammon paikkanäkymä . . . . .	48
<b>7</b>	<b>Tulosten arviointi</b>	<b>51</b>
7.1	Hakutyökaluja ja kontekstietoa georeferoinnin avuksi . . . . .	51
7.2	Historiallisten karttojen tarve . . . . .	51
7.3	Paikkatietoaineiston hajautettu tuottaminen ja täydentäminen . . . . .	52
7.4	HIPLA-palvelumallin toiminnallisuudet muissa palveluissa . . . . .	52
<b>8</b>	<b>Yhteenveto ja jatkokehitys</b>	<b>55</b>
	<b>Viitteet</b>	<b>57</b>

## Lyhenteet

COST	Cooperation in the Fields of Science and Technology
GeoSPARQL	Geographic Query Language for RDF Data
GIR	Geographic Information Retrieval
GIS	Geographic Information System
GPS	Global Positioning System
HTTP	Hypertext Transfer Protocol
INSPIRE	Infrastructure for Spatial Information in Europe
IR	Information Retrieval
ISO	International Organization for Standardization
JulkICT Lab	Julkisen hallinnon palvelu- ja innovaatiotoiminnan kehittämisympäristö
Kotus	Kotimaisten kielten keskus
OGC	Open Geospatial Consortium
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
SAPO	Suomen ajallinen paikkaontologia
SeCo	Semanttisen laskennan tutkimusryhmä
SPARQL	SPARQL Protocol and RDF Query Language
TGN	Getty Thesaurus of Geographic Names
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
WGS84	World Geodetic System 1984
WWW	World Wide Web

# 1 Johdanto

Historiallinen paikkatieto on keskeisessä asemassa kun muistiorganisaatiot, kuten esimerkiksi museot, kirjastot ja arkistot, kartuttavat ja hallinnoivat kokoelmiaan. Sekä kokoelmiin talletettaville fyysisille objekteille että perinnetiedolle täytyy antaa mahdollisimman kattavat metatiedot, jotta kokoelmiin voidaan myöhemmin suorittaa hakuja. Esimerkiksi valokuvan metatiedoiksi voidaan tallentaa kuvauspaikka, kuvaajan nimi, kuvan ottamisen ajankohta, kuvan aihe ja siinä näkyvät henkilöt. Hakutoimintojen ja kokoelmien hallinnoimisen lisäksi metatietoa tarvitaan kansallisissa ja kansainvälisissä verkkopalveluissa, jotka yhdistelevät kokoelmatietoja useista eri lähteistä [33]. Koska paikannimet ja paikan käsite [85] liittyvät kulttuurin kentässä lähes kaikkeen tietoon, muistiorganisaatioiden kokoelmien metatiedoista löytyy suurella todennäköisyydellä viittaus joko nykyiseen tai historialliseen paikkaan.

Toinen tärkeä historiallisen paikkatiedon hyödyntäjä on tällä vuosikymmenellä vakiintunut digitaalisten ihmistieteiden (engl. digital humanities) [81] tutkimusala, jossa humanistiset näkökulmat ja digitaalinen tietojenkäsittely yhdistyvät. Paikkatieto mahdollistaa muun muassa maantieteellisten hakujen suorittamisen ja kokoelmatietojen visualisoinnin kartoilla. Lisäksi paikkatietoa voidaan käyttää uusien yhteyksien löytämiseen tietoaineistoista, jotka eivät suoraviivaisesti liity toisiinsa – yhteinen tapahtumapaikka voi liittää hyvinkin eri tyyppisiä asioita yhteen [31].

Suomessa Maanmittauslaitoksella on lakisääteinen tehtävä ylläpitää kansallista paikannimirekisteriä tällä hetkellä käytössä olevasta nimistöstä, joka on julkaistu avoimena tietona kaikkien käytettäväksi [61]. Historiallisen paikannimien osalta tällaista velvoitetta ei ole kenelläkään, joten historiallista paikannimistöä ei voi kootusti etsiä mistään lähteestä. Historiallinen paikannimistö tuo mukanaan paikkatiedon ajallisen ulottuvuuden, sekä lukuisia muita haasteita, jotka selittävät keskitetyn historiallisen paikannimirekisterin puutetta. Seuraavassa on listattu erityisesti historialliseen paikkatietoon liittyviä haasteita, joista osa tulee esille myös nykyhetkeä koskevaa paikkatietoa käsiteltäessä.

1. **Kontekstitiedon tarjoaminen.** Historiallisen sijainnin tunnistamiseen ja ymmärtämiseen ei yleensä riitä pelkkä tekstimuotoinen paikannimi [76, 32]. Samaan paikkaan voidaan viitata eri nimillä ja murreilmaisilla eri yhteyksissä (synonymia), ja lisäksi samaa nimeä käytetään samaan aikaan monista eri puolilla maailmaa olevista paikoista (homonymia).
2. **Nimistön muutokset.** Paikkojen nimet muuttuvat ajassa. Esimerkiksi luovutetussa Karjalassa paikannimet uusittiin systemaattisesti venäjäksi jatkosodan jälkeen. Lisäksi sama tekstimuotoinen paikkaviittaus, esimerkiksi Saksa, voi tarkoittaa hyvin eri asiaa riippuen siitä, mitä historian hetkeä tarkoitetaan.
3. **Monikielisyys.** Paikoilla on erikielisiä nimiä, joiden primaarisuus saattaa vaihtua ajan kuluessa esimerkiksi ruotsinkielisen kunnan vähitellen muuttuessa suomenkieliseksi.
4. **Aluemuutokset.** Vaikka nimi pysyisi samana, sen tarkoite (esimerkiksi alue)



voi muuttua ajan kuluessa, kuten vaikkapa Helsingin laajeneminen 1500-luvulta osoittaa.

5. **Historiallisten karttojen tarve.** Usein historialliset paikannimet löytyvät jo käytöstä poistetuilta kartoilta. Historiallisia karttalehtiä on yhä enemmän saatavilla skannattuna digitaalisessa muodossa, mutta niiden löydettävyyks on huonolla tasolla, jos niitä ei ole asemoitu moderneille kartoille tai muulla tavoin linkitetty nykyisiin paikannimiin.
6. **Päivittyvän historiallisen nimistöluettelon ylläpito.** Historiallisista paikoista ei voida rakentaa yhtenäistä ja kattavaa listausta. Uusia paikannimiä kohdataan luettelointityössä jatkuvasti, joten uusia paikkatietoresursseja täytyy voida lisätä ja jakaa käyttäjäkunnan kanssa dynaamisesti ja nopeasti. Vain nykyhetken keskittyvissä paikannimirekistereissä päivitystarpeet ovat maltillisempia.
7. **Hajautettujen historiallisten paikannimirekisterien hallinta.** Historiallinen paikkatieto on tyypillisesti hajautunut monien eri organisaatioiden ylläpitämiin rekistereihin. Tämä aiheuttaa päällekkäistä työtä ja vaikeuttaa nykyisten valtioiden rajojen yli levittäytyvien alueiden, kuten esimerkiksi Rooman valtakunnan tutkimista. Eri organisaatioiden ylläpitämissä paikkatietorekistereissä on käytössä monia erilaisia tunnisteita ja tietomalleja historialliselle paikkatiedolle, mikä asettaa rajoituksia ja haasteita historiallisen paikkatiedon keskitetylle hyödyntämiselle.

Edellä mainittujen haasteiden lisäksi historiallisilla paikannimillä ei ole niiden keskeisestä kulttuurisesta merkityksestä huolimatta samaa taloudellista merkitystä kuin nykyisillä paikannimillä. Sanastotyö on monitieteistä, sillä isojen nimistöjen käsittelyssä ja organisoimisessa tarvitaan kielitiedettä, terminologian tutkimusta, historian tutkimusta, geoinformatiikkaa, tietojenkäsittelytiedettä sekä aineistojen julkaisussa ja ylläpidossa osaamista web-teknologioista. Monille rahoittajille ja kehitysohjelmille monitieteisen tutkimuksen rahoittaminen voi olla haasteellista, koska niiden painopiste on usein vain jollain tietyllä tieteenalalla.

Tässä työssä esitetään uusi linkitettyyn dataan perustuva historiallisten paikkojen ja karttojen palvelumalli, HIPLA. Hypoteesina on, että paikkatietoon sovelletut linkitetyn datan teknologiat [56] yhdessä HIPLA-palvelumallin kanssa tarjoavat kulttuurialan organisaatiolle sekä kaikille muille historiallisen paikkatiedon kanssa työskenteleville ratkaisuja edellä mainittuihin haasteisiin. Tavoitteena on yhteisöllinen, keskitetty välityspalvelu, jonka avulla eri organisaatiot voivat yhteistyössä koota, julkaista ja säilyttää historiallisiin paikkoihin liittyvää nimistöä ja karttoja, ilman että samaa työtä tarvitsee tehdä moneen kertaan eri organisaatioiden toimesta.

Paikkatieto ja historiallinen paikkatieto ovat hyvin laajoja tutkimusalueita, joten tämä diplomityö rajataan seuraavien, erityisesti haasteisiin **1, 5, 6, ja 7** liittyvien tutkimuskysymyksien avulla:

1. Minkälaisia toiminnallisuuksia historiallisen aineiston tutkija tarvitsee, jotta hän voi löytää kohtaamalleen tekstimuotoiselle paikkaviittaukselle koordinaatit ja yksilöivän tunnisteen?

2. Miten nykyisillä koordinaattijärjestelmillä esitettyyn paikkatietoon voidaan yhdistää historiallisia karttalehtiä?
3. Miten paikkatietoaineiston jatkuva tuottaminen ja täydentäminen voidaan hajauttaa aineiston käyttäjille reaaliaikaisesti aineiston käyttötilanteessa?

Työssä toteutetaan soveltuvuus selvitys (engl. proof-of-concept), jossa uuden HIPLA-palvelumallin pohjalta kehitetään ja julkaistaan julkisesti verkossa saatavilla oleva Hipla.fi-prototyyppi, joka havainnollistaa HIPLA-palvelumallin ratkaisuja käytännössä. Tutkimuskysymyksiin haetaan vastauksia HIPLA-palvelumallin ja Hipla.fi-prototyypin avulla sekä vertaamalla työssä kehitettyjä ratkaisuja muihin verkossa julkaistuihin historiallista paikkatietoa ja karttoja käsitteleviin palveluihin. Lisäksi esitetään Hipla.fi-prototyypin pilotoinnin tuloksia Sotasampo-projektissa [34]. Työ tehdään Aalto-yliopiston Semanttisen laskennan tutkimusryhmässä, ja työn tuloksia on esitelty ryhmässä valmistuneissa kansainvälisissä tutkimusartikkeleissa [40, 35, 37, 34].

Diplomityön rakenne on seuraavanlainen. Luvussa 2 esitellään työhön liittyvää digitaaliseen paikkatietoon liittyvää teoreettista taustaa, johon kuuluvat muun muassa georeferoinnin haasteet, maantieteellinen tiedonhaku ja paikkatietoon liittyvät linkitetyn datan teknologiat. Tämän jälkeen luvussa 3 luodaan katsaus olemassa oleviin historiallista paikkatietoa tarjoaviin palveluihin. Luku 4 keskittyy HIPLA-palvelumalliin, sen vaatimuksiin, visioon, käyttäjäryhmiin ja käyttötapauksiin sekä arkkitehtuuriin. HIPLA-palvelumallin esittelyn jälkeen luvussa 5 kuvataan miten mallin perusteella rakennettiin ensimmäinen prototyyppipalvelu, Hipla.fi, ja sitä miten prototyyppipalvelu toteuttaa HIPLA-palvelumallin käyttötapauksia. Luku 6 esittelee Hipla.fi-palvelun pilotointia osana talvi- ja jatkosodan sotahistoriallisia aineistoja julkaisevaa Sotasampo-projektia. Luvussa 7 arvioidaan Hipla.fi-palvelua suhteessa muihin historiallista paikannimistöä tarjoaviin palveluihin, minkä jälkeen luvussa 8 esitetään yhteenveto ja jatkokehitystarpeet.

## 2 Digitaalinen paikkatieto

Tämän luvun tarkoituksena on esitellä digitaalisesti esitetyn paikkatiedon teoreettista taustaa. Paikkatiedon ja paikannimistön hallintaan tietojärjestelmissä on kehitetty erilaisia menetelmiä, joita ovat muun muassa nimistöluettelot (engl. gazetteer), tesaukukset, paikkatietoa mallintavat linkitetyn datan sanastot ja paikkatietojärjestelmät (engl. Geographic Information System, GIS) sekä edellä mainittujen menetelmien yhdistelmät. Tässä työssä keskitytään paikkatiedon käsittelyyn varsinaisten paikkatietojärjestelmien ulkopuolisten tietojärjestelmien kannalta, koska paikkatietojärjestelmien integrointi esimerkiksi muistiorganisaatioiden tietojärjestelmiin on ollut puutteellista [31].

### 2.1 Paikkaviittaukset

Fotheringhamin ja Wegenerin mukaan “mitä tahansa tapahtuu, se tapahtuu jossakin ajassa ja paikassa” [20], joten paikka ja aika ovat tärkeässä roolissa, kun informaatiota käsitellään, haetaan ja visualisoidaan. Useimmiten paikkoihin viitataan paikannimillä, ja on arvioitu että vähintään 70 prosenttia kaikista maailman tekstidokumenteista sisältää paikannimillä tehtyjä viittauksia [31].

Historiallisen paikkatiedon kannalta oleelliset kulttuuritietoaineistot sisältävät runsaasti viittauksia paikkoihin [42, 80]. Kokoelmien metatiedot tai bibliografiset tiedot sisältävät yleensä joko kontrolloiduista sanastoista valittuja paikannimiä tai vapaasti kirjoitettuja paikannimiä [31]. Esimerkiksi pääkaupunkiseudun yleisten kirjastojen kirjastoverkko HelMetin haussa “Viipuri”-termillä löytyy yhteensä 485 hakutulosta, joista ensimmäinen on esitetty kuvassa 1. Kuvasta 1 nähdään, että paikkaviittaukselle ei HelMet-järjestelmässä ole omaa kenttää, vaan se löytyy muiden paikkatietoon liittymättömien asiasanojen seasta. Paikkoihin viittaamiseen sisältyy lukuisia muitakin haasteita, joita on käsitelty tarkemmin luvussa 2.3.

Teoksen tiedot	
Tekijä	Gardberg, C. J.
Teoksen nimi	<b>Viipuri</b> : kivestä rakennettu kaupunki / C. J. Gardberg, P. O. Welin ; [käsikirjoituksesta .. suomentanut Irma Savolainen]
Julkaisutiedot	Helsingissä : Otava, 1996
Kontrollinumero	9511139304
ISBN	951-1-13930-4 sid.
Kieli	suomi
Aineistolaji	KIRJA / BOK
Ulkoasu	143 s. : kuv., kartt.
Alkuteos	98
Muu nimeke	Viborg: en stad i sten, suomi
Asiasana	arkkitehtuuri historia rakennuskulttuuri kulttuurihistorialliset rakennukset kaupunkisuunnittelu paikallishistoriat luovutetut alueet <b>Viipuri</b> Karjala

Kuva 1: Paikkaviittaus asiasanana Helmet-aineistohaussa

## 2.2 Sijainti ja sijaintiin liittyvä tieto

Kuvassa 2 on esitetty käsitekartta, joka sisältää tämän työn kannalta olennaisia paikkatietoon liittyviä termejä. Kuva ja termien määrittelyt perustuvat Geoinformaatiikan sanastoon [75]. Keskeisin käsite on sijainti, joka tarkoittaa paikkaa, jossa jokin tarkasteltava asia sijaitsee. Sijainti voidaan esittää koordinaattien avulla, jolloin puhutaan suorasta sijainnista. Suora sijainti määrittää paikan yksikäsitteisesti. Sen sijaan, jos sijainti ilmoitetaan esimerkiksi paikannimen tai osoitteen avulla, kyseessä on epäsuora sijainti, jolloin paikka ei enää välttämättä ole yksikäsitteisesti määritelty.

Koordinaattien avulla ilmaistut sijainnit määritellään käytännössä yhden tai useamman koordinaattipisteen avulla, jolloin saadaan kuvattua pistemäisiä kohteita tai erimuotoisia alueita. Historialliseen paikkatietoon liittyen oleellisia koordinaatteja ovat pituus- ja leveyskoordinaatit, koska aineistoissa on harvoin saatavilla tietoa korkeudesta.

Kun missä tahansa tietojärjestelmässä liitetään sijaintia koskevaa tietoa muuhun tietoon, eli esimerkiksi merkitään valokuvan metatietoihin kuvauspaikka, puhutaan georeferoinnista. Jos paikkaan viitataan epäsuoran sijainnin avulla, kyseessä on vapaamuotoinen georeferointi, kun taas suoralla sijainnilla paikkaan viittaaminen tarkoittaa formaalia georeferointia [31, 50].

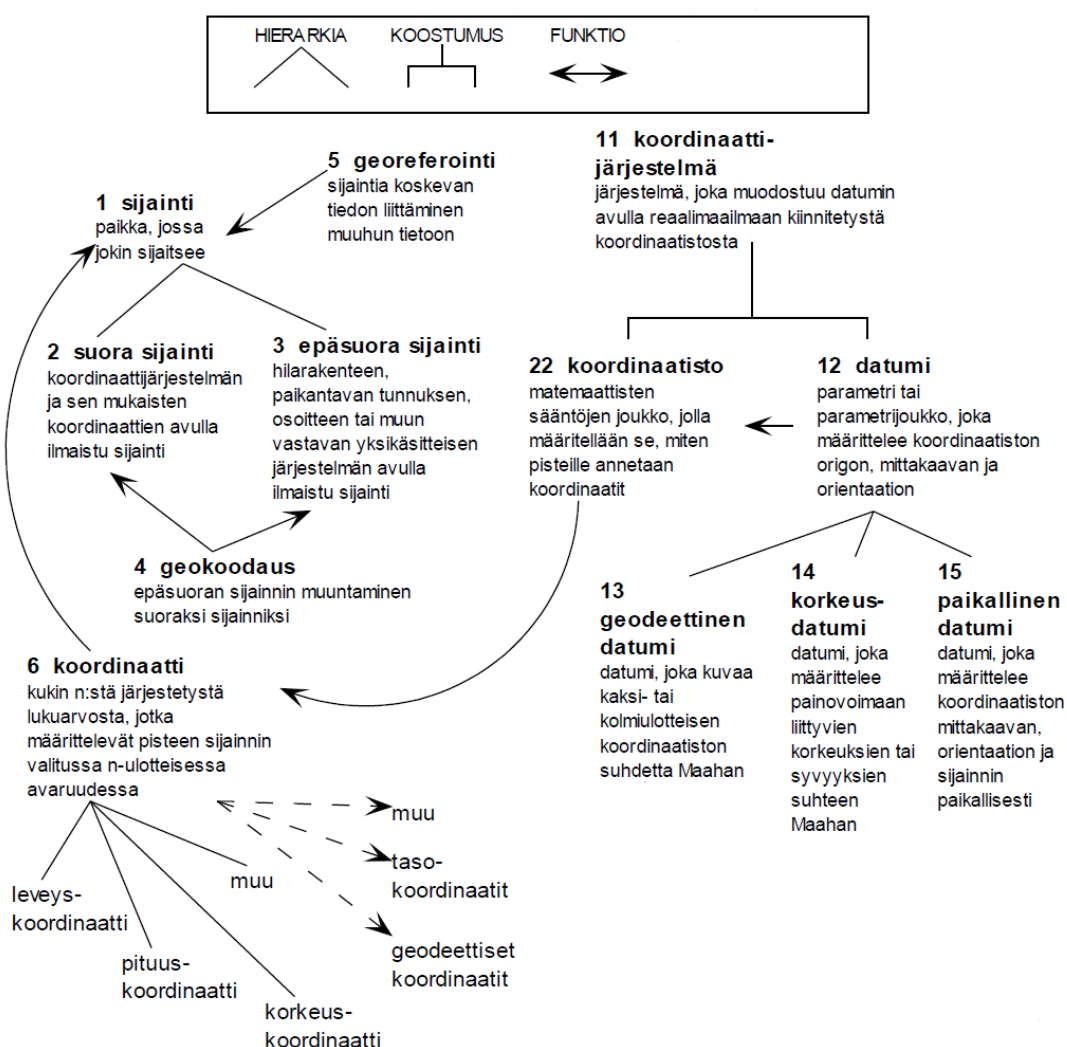
## 2.3 Georeferoinnin haasteet

Formaali georeferointi on GPS-järjestelmän yleistymisen ja verkossa saatavilla olevien karttasovelluksien kehityksen myötä helpompaa kuin koskaan aiemmin, mutta silti suuri osa esimerkiksi muistiorganisaatioiden aineistojen paikkaviittauksista muodostuu erilaisilla tarkkuustasoilla ilmaistuista tekstimuotoisista paikannimistä, eli epäsuorista sijainneista. Tähän on osaltaan syynä se, että paikkatiedon käsittely on perinteisesti keskittynyt erillisiin paikkatietojärjestelmiin, joiden integroiminen muihin tietojärjestelmiin on ollut puutteellista tai olematonta [31]. Verrattuna formaaliin georeferointiin vapaamuotoinen georeferointi sisältää useita kielellisiä ja semanttisia haasteita, joita on esitelty seuraavassa listauksessa.

- **Paikannimet eivät ole yksilöiviä.** Synonymian vuoksi samaan paikkaan voidaan viitata eri nimillä ja murreilmaisuuilla eri yhteyksissä. Lisäksi samaa nimeä käytetään samaan aikaan monista eri puolilla maailmaa olevista paikoista, jolloin puhutaan homonymiasta. Klassinen esimerkki homonymiasta on suomenkielinen paikannimi Pyhäjärvi: eri puolilla Suomea on kaikkiaan 39 eri järveä, joiden nimi on Pyhäjärvi. Tämän lisäksi Pyhäjärvi on esiintynyt myös kolmen kunnan nimenä, muiden asuttujen paikkojen nimenä ja sukunimenä.
- **Monikielisyys ja murteet.** Paikoilla on monia nimiä eri kielillä ja murteilla.
- **Paikannimet muuttuvat ajan kuluessa.** Esimerkiksi luovutetun Karjalan alueen paikkojen suomenkieliset nimet muutettiin venäjänkieliseksi.
- **Sijainnit muuttuvat ajan kuluessa.** Esimerkiksi Helsinki 1500-luvulla kat-  
taa hyvin erilaisen alueen kuin 2010-luvulla. Monet aineistot ja tilastot perus-

tuvat hallinnolliseen aluejakoon, joka on ollut voimassa aineistoa laadittaessa. Tällöin hallinnollisissa alueissa tapahtuneet muutokset aiheuttavat ongelmia jälkeenpäin; vuosien 1865 ja 2015 välillä pelkästään Suomen kunnissa on tapahtunut yhteensä 1101 muutostapahtumaa, jotka koskevat 1736 kuntaa [88].

Yllä mainituista haasteista huolimatta on selvää, että paikkatiedon kanssa työskenneltäessä tarvitaan keinoja, joilla itse aineistosta, tai aineiston metatiedoista löytyvä epäsuora sijainti voidaan ymmärtää ja sijoittaa oikeaan kohtaan kartalla. Tässä tehtävässä tärkeänä tiedonlähteenä toimivat nimistöluettelot (engl. gazetteer), joita esitellään tarkemmin luvussa 2.4.



Kuva 2: Käsittekaavio sijaintiin liittyvistä käsitteistä [75]

## 2.4 Nimistöluettelot

Nimistöluettelo [25] on luettelo maantieteellisistä nimistä, joihin on liitetty sijaintitietoa ja muuta kuvailevaa tietoa. Nimistöluetteloon tallennettavien tietojen vähimmäisvaatimuksena voidaan pitää paikannimeä eli toponyymiä [4] sekä paikkatyyppejä. Näiden lisäksi koordinaateilla määritelty sijaintitieto on hyvin olennainen, mutta sen pakollisuudesta nimistöluettelossa ei ole yhtenäistä näkemystä. Esimerkiksi Hill listaa sijaintiedon välttämättömäksi nimistöluettelon elementiksi [31], kun taas Southall ym. päätyy johtopäätöksissään siihen, että nimistöluetteloista menetetään olennaista tietoa, jos niissä ei listata myös paikkoja, joiden sijaintia ei tiedetä tarkkaan [79].

Tietojärjestelmien kontekstissa nimistöluettelo on tietämuskanta, jota voidaan käyttää muun muassa kaksisuuntaiseen geokoodaukseen: 1) paikannimestä voidaan siirtyä koordinaattipisteeseen tai alueelle kartalla, 2) koordinaattipisteestä tai alueesta voidaan siirtyä paikannimiin [32]. Teoriassa tämä johtaa siihen, että sekä formaalilla että vapaamuotoisella georeferoinnilla saavutetaan sama lopputulos. Käytännössä tämä ei kuitenkaan toteudu, koska nimistöluetteloiden tiedot eivät ole kattavia, eikä niitä ole saatavilla laajoilta alueilta yhtenäisessä formaatissa. Nimistöluetteloiden alueellinen kattavuus sekä luvussa 2.3 esitettyjen haasteiden huomioiminen nimistöluetteloiden toteutuksessa ja käytössä ovat merkittävimpiä georeferoinnin onnistumiseen vaikuttavia tekijöitä [28]. Tulevaisuuden visiona on kuitenkin esitetty, että nimistöluettelopalvelujen ja automaattisen tekstintunnistuksen kehityksen myötä paikkatietojärjestelmien ulkopuolella toimineisiin ja vapaamuotoista georeferointia hyödyntäviin tietojärjestelmiin voidaan suorittaa yhä enemmän maantieteellistä tiedonhakuja ja paikkoihin viittaavaan tiedon visualisointia [31].

## 2.5 Maantieteellinen tiedonhaku

Tietojenkäsittelytieteessä tiedonhaku (engl. information retrieval, IR) tarkoittaa tietyn tiedontarpeen täyttävän materiaalin (yleensä dokumenttien) hakemista laajoista, tietojärjestelmiin tallennetuista kokoelmista [63]. Tiedonhaun menetelmät ovat olleet pääasiassa tekstipohjaisia, keskittyen indeksointiin ja hakuun kontrolloituja sanoja tai vapaata tekstiä käyttäen. Ray Larsonin vuonna 1996 ensimmäistä kertaa kuvailema maantieteellinen tiedonhaku [57] (engl. geographic information retrieval, GIR) lisää tiedonhaakuun spatiaalisen ulottuvuuden.

Paikkatieto mahdollistaa hyvin erilaisten tieteenalojen aineistojen yhdistelemisen. Pelkkä perinteinen tekstihaku ei kuitenkaan riitä. Jos käyttäjällä ei ole tarkkaa paikannimeä tiedossa, vaan ainoastaan ajatus siitä, missä päin haettu paikka sijaitsee, tarvitaan ehdottomasti karttapohjainen hakukäyttöliittymä. Seuraavassa on listattu eräitä tyypillisiä maantieteellisen tiedonhaun käyttötapauksia.

- **Aluekysely** - haetaan määritellyn alueen (useimmiten suorakaide) sisäpuolella sijaitsevat entiteetit.
- **Geokoodaus** (engl. geocoding) - muunnetaan tekstimuotoinen kuvaus sijainnista maantieteelliseksi koordinaateiksi.

- **Käänteinen geokoodaus** - haetaan entiteettejä maantieteellisten koordinaattien avulla.
- **Maantieteellinen yhdistely** - haetaan toisiaan osittain tai kokonaan leikkaavia maantieteellisiä entiteettejä.

Edellä mainittujen maantieteellisen tiedonhaun käyttötapauksien käytännön toteutukseen etenkin suuremmille aineistoille tarvitaan useimmiten maantieteellisiä indeksejä. Esimerkiksi laajan paikkatietoaineiston koko sisällön selaamista karttanäkymässä voidaan nopeuttaa rajoittamalla hakutuloksia vain käyttäjän näytöllä näkyvissä olevalle alueelle aluekyselyn avulla [95].

## 2.6 Ajallisen ulottuvuuden tuki paikkatiedon standardeissa

Luvussa 2.3 mainittuun paikkatiedon ajallisen ulottuvuuden problematiikkaan on monia syitä. Vaatii suurta asiantuntemusta määrittellä esimerkiksi tietylle paikannimelle oma ajanjaksonsa, jolloin se on ollut voimassa. Vain tietyt poikkeukselliset paikkatietoaineistot, kuten Suomen ajallinen paikkaontologia (SAPO) [38] mallintavat paikkatiedon ajallista ulottuvuutta kattavasti. Ajallisen ulottuvuuden huomiotta jättäminen johtuu myös osaltaan tietojärjestelmien ja standardien sisältämistä puutteista. Kenties merkittävin standardeja julkaiseva organisaatio International Organization for Standardization (ISO) on julkaissut standardin *ISO 19112:2003 Geographic Information — Spatial Referencing by Geographic Identifiers* [41], joka on tarkoitettu epäsuorilla sijainneilla tapahtuvien paikkaviittausten määrittelyyn. Tämä standardi sivuuttaa täysin paikkatiedon ajallisen ulottuvuuden, koska ainut paikkoihin liittyvä ajanmääre, nimeltään *Temporal extent*, on tarkoitettu kuvaamaan ainoastaan ajanhetkeä, jolloin paikkainstanssi on luotu järjestelmään, eikä esimerkiksi sitä, minä aikana kyseinen paikka on ollut olemassa.

Myöskään Eurooppalaisen paikkatietoinfrastruktuurin määrittelevä Inspire-direktiivi ei ota vahvaa kantaa paikkatiedon ajalliseen ulottuvuuteen. Direktiivin liittyvässä teknisessä ohjeessa metatietojen laatimiseksi [18] mainitaan, että paikkainstanssille täytyy antaa vähintään yksi ajallinen määre seuraavista kategorioista: 1) ajallinen ulottuvuus, 2) julkaisuaika, 3) viimeisimmän päivityksen ajanhetki tai 4) luomisaika. Näin ollen ajallista ulottuvuutta ei määritellä Inspire-direktiivin piirissä paikkatiedon pakolliseksi ominaisuudeksi.

Käytännön sovelluksissa paikkatietoa mallinnetaan lukemattomilla eri tavoilla ja harvoin täysin paikkatiedon standardien mukaisesti. Tämän vuoksi luvussa 3, jossa esitellään historiallista paikkatietoa tarjoavia palveluita, on kerrottu tarkemmin, millä tavalla kussakin esitellyssä palvelussa on otettu kantaa paikkatiedon ajalliseen ulottuvuuteen.

## 2.7 Paikkatiedon visualisointi

Paikkatiedon visualisointi on yksi yleisimmistä paikkatiedon sovellusalueista. Paikkatiedon visualisointiin liittyen on tunnistettu kolme erilaista käyttötapausta [55].

- Ennalta tuntemattoman paikkatiedon tutkiminen, jolloin paikkatietoa sisältävän datajoukon sisällön visualisoiminen kartalle auttaa hahmottamaan datajoukon sisältöä, rakennetta, kattavuutta ja mahdollisia käyttötarkoituksia
- Analyysin apukeinona käyttäminen, esimerkiksi yhdistämällä kahden paikkatietolähteen tiedot samalle kartalle, voidaan tehdä lähteiden sisällön välistä maantieteellistä vertailua.
- Analyysin tuloksien synteesi, jossa visualisoinnin avulla paikkatietoon liittyvän analyysin tulokset voidaan esittää suurelle yleisölle helposti ymmärrettävässä muodossa.

Paikkatieto visualisoidaan luonnollisesti useimmiten karttojen avulla, mutta koska paikkatietoon liittyy usein tilasto- ja ominaisuustietoa, visualisointikäyttöliittymissä täytyy yhdistellä tekstimuotoista tietoa karttoihin. Visualisointiin voidaan käyttää paikkatietojärjestelmiä tai lukuisia, erityisesti paikkatiedon visualisoimiseen tarkoitettuja ohjelmointirajapintoja, joita ovat esimerkiksi Google Maps JavaScript API [27], OpenStreetMap API [73] ja Leaflet [5]. Visualisointityökalun valinta riippuu muun muassa siitä, millaista pohjakarttaa halutaan käyttää.

## 2.8 Linkitetty paikkatieto

### Linkitetty data

Linkitetty data (engl. linked data) tarkoittaa joukkoa standardeja ja periaatteita, joiden tavoitteena on julkaista ja liittää yhteen rakenteellista tietoa ja metatietoa käyttäen hyväksi nykyistä WWW:n (World Wide Web) arkkitehtuuria. WWW:n keksijänä parhaiten tunnettu Tim Berners-Lee on määritellyt neljä linkitetyn datan pääperiaatetta [9]:

1. Käytä asioiden nimeämiseen yksilöiviä URI-tunnisteita (Uniform Resource Identifier).
2. Käytä URI-tunnisteina HTTP-osoitteita (Hypertext Transfer Protocol), jotta ihmiset voivat löytää niistä lisätietoa.
3. Kun joku etsii lisätietoa URI:n avulla, tarjoa sitä käyttäen linkitetyn datan standardeja, joita ovat muun muassa RDF (Resource Description Framework) [16], RDFS (RDF Schema) [12], OWL (Web Ontology Language) [64] ja SPARQL (SPARQL Protocol and RDF Query Language) [90].
4. Tarjoa URI-tunnisteeseen liittyviä linkkejä muihin URI-tunnisteisiin, jotta niitä seuraamalla löytää uusia asioita.



Linkitetyn datan tietomallina toimii RDF, jossa kaikki tieto rakentuu subjekti-predikaatti-objekti-kolmikoista (engl. triple). RDF ei ota kantaa tiedon esitysmuotoon, joten RDF-muotoinen tieto voidaan esittää useiden eri formaattien avulla.

Tietomallin lisäksi tarvitaan kyselykieli, jolla voidaan hyödyntää tietovarastoja. Kyselykielenä toimii SPARQL, joka jakaantuu W3C:n (World Wide Web Consortium) standardin mukaan kolmeen osaan: 1) kyselykieli ja sen syntaksi ja semantiikka 2) kyselyn tulosten esitystapa, ja 3) SPARQL:n sidonta HTTP-protokollaan, mikä mahdollistaa kyselyiden ja niiden tuloksien siirtämisen internetiin kytkettyjen koneiden välillä.

Linkitetyn datan visiona on, että edellä mainittujen periaatteiden mukaisesti julkaistut tietoaineistot muodostovat maailmanlaajuisen linkitetyn datan pilven (linked data cloud), jossa tähän pilveen liittyneet tietoaineistot täydentävät ja rikastavat toisiaan. Kun data esitetään koneluettavana ja koneymmärrettävänä, aineistoja voidaan yhdistellä helpommin ja uutta tietoa voidaan päätellä annetun tiedon pohjalta [30].

### **Paikkatietojärjestelmien ja linkitetyn datan teknologioiden yhdistäminen**

Perinteisesti paikkatietoa on käsitelty pääasiallisesti muista tietojärjestelmistä erillisissä paikkatietojärjestelmissä. Tämä on aiheuttanut ongelmia, koska paikkatietojärjestelmissä ei usein ole tarjolla työkaluja paikkatiedon semanttisen yhteentoimivuuden mahdollistamiseksi. Osaltaan tämän vuoksi paikkatietojärjestelmiin säilötyn paikkatiedon hyödyntäminen paikkatietojärjestelmien ulkopuolisissa tietojärjestelmissä on vaikeaa.

Paikkatieto on merkittävä osa globaalia linkitetyn datan pilveä [56]. Lisäksi monia informaation integroimis- ja kerryttämistehtäviä ei voida suorittaa ilman paikkatietoa. Paikkatiedon kanssa työskenteleville toimijoille linkitetyn datan lähestymistapa pyrki tarjoamaan työkaluja, joilla dataa voidaan tuottaa itsenäisesti, ja sen jälkeen mahdollisimman vaivattomasti integroida toimijan oma data muihin datajoukkoihin. Paikkatiedon avulla voidaan kytkeä yhteen monia hyvin erilaisten aihealueiden datajoukkoja ja löytää uusia lisäarvoa tuovia linkityksiä.

Paikkatieto sisältää myös usein monimutkaisia paikkatyypin muodostamia hierarkioita. Esimerkiksi joki voi olla samaan aikaan vesireitti, kuljetusväylä tai hallinnollinen alue. Tällä osa-alueella perinteisten tietokantateknologioiden ja RDF-tietokantojen päättelykykyjen yhdistäminen on osoittautunut lupaavaksi sekä linkitetyn datan alueella että paikkatietojärjestelmissä [56].

Luvussa 2.4 esitellyissä nimistöluetteloissa paikkannimien yksikäsitteistäminen tapahtuu tyypillisesti käyttämällä hyväksi paikkannimen lisäksi muita paikkaan liitettyjä tietoja. Yksi nimistöluetteloiden suurimmista puutteista on, että tätä yksikäsitteistämistä on hyvin vaikeaa saada mukaan sekä tiedon indeksoinnissa että tietoa hakiessa. Linkitetyn datan lähestymistavassa kaikille paikoille annetaan yksilöivät URI-tunnisteet, jolloin kertaalleen suoritettu yksikäsitteistäminen takaa oikean lopputuloksen sekä indeksoinnissa että haussa.

Linkitetyn datan periaatteiden tuominen paikkatiedon käsittelyyn on kuitenkin vasta alkutaipaleella. Linkitetyn datan ja paikkatietoalan merkittävimmät toimijat World Wide Web Consortium (W3C) ja Open Geospatial Consortium (OGC) jul-

kistivat tammikuussa 2015 yhteistyöprojektin [70], jonka tarkoituksena on edistää paikkatiedon yhteentoimivuutta ja integroimista web-ympäristössä.

### Linkitetyn datan sanastoja paikkatiedon mallinnukseen

Paikkatiedon mallinnukseen on vuosien varrella esitetty useita erilaisia RDF-sanastoja, mutta mistään sanastosta ei ole vielä muodostunut vallitsevaa standardia. Kenties tunnetuin sanasto on W3C:n Basic Geo -sanasto [93], joka dokumentaation mukaan “tarjoaa muutaman perustermin, joita voi käyttää RDF-datassa silloin, kun syntyy tarve kuvailla koordinaattitietoa.” Basic Geo -sanasto tukee lähtökohtaisesti vain pistemäisiä kohteita. GeoRSS [1] lisäsi tuen monipuolisemmille geometrisille objekteille, kuten viivoille, suorakulmioille ja polygoneille. GeoOWL [94] ja NeoGeo Geometry Ontology [2] tarjoavat vielä edellä mainittuja sanastoja kehittyneempää mallinnusta paikkatiedon käsitteille. GeoJSON [13] on paikkatiedon tallennus- ja välitysformaatti, joka perustuu JavaScript Object Notation (JSON) -kieleen ja sisältää tuen seuraaville geometrisille primitiiveille: *Point*, *LineString*, *Polygon*, *Multipoint*, *MultiLineString*, *MultiPolygon* ja *GeometryCollection*. GeoJSON-LD [3] on linkitetyn datan laajennus GeoJSON-formaatille.

Vahva ehdokas tulevaksi vallitsevaksi linkitetyn datan paikkatietoon liittyväksi sanastoksi on Open Geospatial Consortiumin (OGC) kehittämä GeoSPARQL-standardi [7], jonka pyrkimyksenä on tarjota yhtenäinen ja alustasta riippumaton pääsy RDF-muotoiseen paikkatietoon. Toisin kuin edellä mainitut sanastot, GeoSPARQL on tarkoitettu sekä paikkatiedon esittämiseen RDF-muodossa että paikkatiedon kyselyihin SPARQL:n laajennosten avulla. Lisäksi GeoSPARQL ei rajoitu pelkästään WSG84-koordinaatistoon, kuten esimerkiksi W3C:n Basic Geo -sanasto, mikä helpottaa yhteensopivuutta erilaisiin koordinaattijärjestelmiin perustuvien paikkatietoaineistojen kanssa. GeoSPARQL on kuitenkin tätä kirjoitettaessa vasta alkutaipaleella oleva standardiehdokas, ja vain harvat nykyisin käytössä olevat RDF-tietokannat sisältävät edes osittaisen GeoSPARQL-toteutuksen.

### RDF-tietokantojen maantieteelliset indeksit

Perinteiset tietokantojen indeksit, kuten tekstihakuun tarkoitetut indeksit, eivät sisällä tukea maantieteellisen tiedonhaun kyselyille. Maantieteelliset indeksit (engl. spatial index) tarjoavat ratkaisuja maantieteellisten kyselyiden optimointiin. Yksinkertaisia maantieteellisen tiedonhaun kyselyjä, kuten suorakulmiolla rajoitettuja aluekyselyjä, voidaan toki suorittaa myös kokonaan ilman maantieteellisiä indeksejä. Maantieteellisiä indeksejä tarvitaan kuitenkin etenkin silloin, kun maantieteellisten kyselyiden kohteena olevat aineistot ovat laajoja [22]. Useita monimutkaisempia maantieteellisen tiedonhaun kyselyitä, esimerkiksi kahden polygonin leikkausta palauttavaa kyselyä on käytännössä mahdotonta suorittaa ilman maantieteellistä indeksiä.

Tämän työn kannalta oleellinen Linked Data Finland -julkaisualusta [36] sisältää Fuseki-tietokannan [83] laajennuksena Jena Spatial -indeksin [84], joka on tarkoitettu yksinkertaiseen maantieteelliseen tiedonhaakuun ilman, että sen käyttäjältä vaaditaan syvällistä tuntemusta paikkatietojärjestelmistä. Käytännössä Jena Spatial tarjoaa

maantieteellisen indeksin ja joukon valmiita SPARQL-funktioita muun muassa erilaisten aluekyselyiden luomiseen. Myös esimerkiksi RDF-tietokanta Virtuoso on lisätty tuki maantieteellisille kyselyille osin GeoSPARQL-standardiin perustuvilla laajennuksilla [72].

### 3 Historiallisen paikkatiedon palvelut ja lähteet

Tässä luvussa luodaan katsaus verkossa saatavilla oleviin historiallista paikkatietoa tarjoaviin palveluihin, jotka tarjoavat käyttöliittymän. Lisäksi mukaan on otettu tämän työn kannalta olennaisia historiallisen paikkatiedon lähteitä, joille ei ole olemassa varsinaista käyttöliittymää. Koska linkitetyn datan tuki on tämän työn kannalta olennainen seikka, palvelut ja lähteet on jaoteltu kahteen ryhmään: 1) perinteiseen teknologiaan perustuvat, ja 2) linkitettyä dataa hyödyntävät.

#### 3.1 Perinteiseen teknologiaan perustuvat palvelut ja lähteet

##### **The Historical Gazetteer of England's Place-names**

The Historical Gazetteer of England's Place-names [17] on yksi tunnetuimmista kansallisista historiallisten paikkojen palveluista. Palvelussa on noin neljä miljoonaa historiallista paikkaa ja se perustuu perinteiseen teknologiaan. Palvelu sisältää monipuolisen tekstihaun: paikkoja voi etsiä sekä modernilla että historiallisella kirjoitusasulla. Paikkatiedon ajallinen ulottuvuus on huomioitu siten, että historiallisten paikannimien kirjoitusasun yhteyteen on merkitty päivämäärä, jolloin kyseistä kirjoitusasua on käytetty alkuperäisessä lähteessä. Näiden päivämäärien lisäksi hakutuloksia voi suodattaa lähteen, läänin ja paikkatyypin mukaisesti.

Karttapohjaista hakua ei ole toteutettu lainkaan. Aineistojen selailu tapahtuu valitsemalla ensin läänin ja etenemällä tämän jälkeen kyseisen läänin alueella sijaitseviin paikkoihin. Vasta valitsemalla yksittäisen paikannimen käyttäjä saa esiin paikan sijainnin visualisoituna historiallisella kartalla. Paikkojen selailu yhtenäisessä karttanäkymässä ei ole mahdollista, eikä paikkoihin liittyvää kontekstietoa ole tarjolla.

##### **Gazetteer for Scotland**

Gazetteer for Scotland on tarjonnut vuodesta 1998 lähtien verkkokäyttöliittymän, jonka käyttämä tietokanta sisältää tietoa Skotlannin alueen kaupungeista, kylistä, vuorista, laaksoista, turistikohteista sekä historiallisista rakennuksista ja paikoista. Maantieteellisen tiedon lisäksi palveluun on lisätty paikkoihin liittyviä tapahtumia, kuuluisia henkilöitä, klaanien nimiä, tilastotietoa ja yritystietoa. [24]

Palvelu sisältää noin 23 000 paikkaa, joita on mahdollista hakea sekä tekstihaulla että karttahaulla. Tekstihaun tuloksia voidaan suodattaa paikkatyypin mukaan ja käyttäjä voi valita ottaako tekstihaku huomioon paikannimen, vaihtoehtoisen nimen, esittelytekstin tai edellä mainittujen yhdistelmän. Osaan paikoista on liitetty kontekstietoa, esimerkiksi paikkoihin liittyvien perheiden ja klaanien historiatietoa ja historiallisia tapahtumia.

##### **Paikannimirekisteri**

Maanmittauslaitoksen ylläpitämä paikannimirekisteri on Suomen paikannimitietojen ajantasainen tietovarasto, joka koostuu paikannimirekisteristä ja karttanimirekis-

teristä. Paikannimirekisteriin on tallennettu nimettyjen paikkojen ja paikannimien tiedot, kun taas karttanimirekisteri sisältää paikannimien kartografisen esittämisen tiedot karttatuotteittain. Paikannimien kirjoitusasu on Kotimaisten kielten keskuksen (Kotus) asiantuntijoiden tarkistama. Paikannimirekisteri sisältää noin 800 000 paikkaa ja se avattiin vapaaseen käyttöön vuonna 2011 [61].

Paikannimirekisteriä ei ole vielä virallisesti julkaistu linkitetyn datan periaatteiden mukaisesti, mutta keväällä 2016 on käynnistynyt Julkisen hallinnon digitaalisten palvelujen kehittämissympäristön (JulkICT Lab) hanke [45], jonka tavoitteena on tuottaa virallinen linkitetyn datan julkaisu paikannimirekisteristä. Tässä työssä on hyödynnetty Semanttisen laskennan tutkimusryhmän toteuttamaa paikannimirekisterin linkitetyn datan julkaisua [59], jota on käytetty aiemmin muun muassa ONKI-Paikka-palvelussa [58].

## Old Maps Online

Sveitsiläisen Klokian Technologies -nimisen yrityksen ylläpitämä Old Maps Online [52] on globaali historiallisten karttojen indeksointi- ja hakupalvelu, jonka aineistossa on kirjoitushetkellä noin 380 000 linkkiä historialliseen karttaan. Karttalehtiä voi etsiä siirtämällä modernia karttanäkymää tai tekstihaulla syöttämällä paikannimen, jolloin moderni karttanäkymä siirtyy kyseisen paikan kohdalle. Sivupalkissa listataan esikatselukuvat ja metatiedot kaikista modernia karttanäkymää sivuavista historiallisista kartoista. Esikatselukuvan kautta käyttäjä pääsee kartan julkaisijan sivulle, josta voi ladata varsinaisen karttakuvan. Pääosaa historiallisista kartoista ei ole asemoitu tukipisteiden avulla modernien karttojen päälle (georektifointi).

Old Maps Online on luonteeltaan historiallisten karttojen indeksointipalvelu, ja se ei tarjoa historiallisten karttojen metatietoja ohjelmointirajapinnan kautta. Lisäksi osa palvelun tarjoamista linkeistä on vanhentuneita. Erityisesti moni Suomen aluetta koskeva karttalinkki palauttaa virheilmoituksen.

## Old-Maps

Old-Maps [68] on Ison-Britannian historiallisiin karttoihin keskittynyt palvelu, joka sisältää myös nimistöluettelon. Karttoja voi etsiä viidellä eri tapaa: 1) siirtämällä modernin karttanäkymän haluttuun kohtaan, 2) kaupunkien nimillä, 3) postinumerolla, 4) koordinaateilla tai 5) nimistöluettelon avulla, jolloin valitaan ensin lääni ja sen jälkeen kaupunki.

Modernin karttanäkymän kattavat historialliset kartat näytetään valikoimana sivupalkista, josta voi valita yhden historiallisen karttalehden kerrallaan näkyviin modernin karttanäkymän päälle. Historiallisen karttalehden läpinäkyvyyttä ei voi säätää. Palvelu on osin maksullinen. Jotta historiallisia karttalehtiä voi zoomata lähelle, täytyy kyseinen karttalehti ensin ostaa palvelusta. Ohjelmointirajapintaa tai karttojen metatietoa ei ole saatavilla ulkoisissa palveluissa hyödynnettäväksi.

## 3.2 Linkitettyä dataa hyödyntävät palvelut ja lähteet

### Suomen ajallinen paikkaontologia

Suomen ajallinen paikkaontologia (SAPO) [38] perustuu linkitettyyn dataan ja Kauppisen ja Hyvösen kehittämään menetelmään [49], jolla mallinnetaan paikkaontologiassa tapahtuvia muutoksia ajan kuluessa, sekä Jari Väätäisen mittavaan sisältötyöhön. SAPO sisältää muutostiedot Suomen hallinnollisista alueista eli Suomen alueella vaikuttaneista kunnista, lääneistä ja valtioista. Kuntien tiedoista löytyy niiden olemassaolon alkuaikajankohta, sekä loppuaikajankohta, mikäli kunta on lakannut olemasta. Osalle kunnista on tallennettu niiden rajat polygoneina. Tämän lisäksi SAPO mallintaa eri aikakausina voimassa olleiden kuntien välisiä peittävyksiä.

### Getty Thesaurus of Geographic Names

Kansainvälisten historiallisten paikkojen nimistöluetteloista tunnetuin on [79] amerikkalaisen Getty-säätiön ylläpitämää Thesaurus of Geographical Names (TGN) -luettelo [82], jota käytetään laajasti USA:ssa ja monissa muissa maissa, esimerkiksi Hollannissa. Palvelun pääsijallisia käyttäjäryhmiä ovat arkistonhoitajat ja luettelot ja kuvataide- ja museoaloilla.

TGN sisältää nykyisin lähes 1,5 miljoonaa historiallista paikkaa ja data julkaistiin vuonna 2015 avoimena linkitettyinä datana [15]. Suomi ei tässä laajassakaan sanastossa ole juurikaan edustettuna. Esimerkiksi pääkaupunkiseudun alueelta löytyy ainoastaan neljä paikkainstanssia. Vaikka suurin osa TGN:n paikkainstansseista sisältää koordinaattitiedot, ei TGN tarjoa omassa hakukäyttöliittymässään maantieteellisiä hakutoimintoja, kuten esimerkiksi aineiston selausta karttapohjalla. Dokumentaation mukaan paikkojen koordinaatit on tarkoitettu vain paikkojen summittaiseen paikantamiseen, mutta TGN:n linkittäminen olemassa oleviin paikkatietojärjestelmiin tuo sisällön paremmin saataville myös kuvataide- ja museoalojen ulkopuolisille toimijoille. Tähän liittyen linkitetyn datan rajapinnat tarjoavat lupaavia keinoja hyödyntää TGN:n sisältöä osana linkitetyn datan pilveä.

Kuvassa 3 on esitetty Viipurin tietosivu TGN-hakukäyttöliittymässä. Kuvasta nähdään, että TGN sisältää useita kirjoitusasuja paikoille. Kirjoitusasuille käytetään muun muassa seuraavia tunnisteita: *“nykyinen”* (C), *“historiallinen”* (H), *“sekä nykyinen että historiallinen”* (B) ja *“paikallinen”* (V tai O). TGN mahdollistaa myös paikkainstanssien voimassaoloajan määrittelyn. Tästä huolimatta suuri osa paikkatiedon ajalliseen ulottuvuuteen liittyvistä tiedoista on sisällytetty tekstimuotoisesti kuvauskenttiin, mikä estää tehokkaan ohjelmallisen hyödyntämisen. Kuvasta 3 ilmenee myös, että TGN sisältää eriteltyt lähteet paikkojen erilaisille kirjoitusasuille.

### Pelagios-hanke ja Pleides-nimistöluettelo

Pelagios [78] 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

## Vyborg (inhabited place)

### Coordinates:






Lat: 60 45 00 N *degrees minutes* Lat: 60.7500 *decimal degrees*  
 Long: 028 41 00 E *degrees minutes* Long: 28.6833 *decimal degrees*

**Note:** Located on Gulf of Finland; member of Hanseatic League by 14th cen.; captured by Peter the Great in 1710; included in Finland under Russian sovereignty in 1812 & under complete Finnish control 1918-1940; ceded back to Russia in 1947.

### Names:

**Vyborg** (**preferred**,C,V)  
**Viborg** (C,O)  
**Viipuri** (C,O)  
**Wyborg** (C,O)

### Hierarchical Position:

 World (facet)  
 .... Asia (continent) (P)  
 ..... Russia (nation) (P)  
 ..... Leningradskaya Oblast' (oblast) (P)  
 ..... Vyborg (inhabited place) (P)

### Place Types:

inhabited place (**preferred**, C) ..... founded as a fortress by Swedes in 1293  
 city (C)  
 port (C)

### Sources and Contributors:

Viborg..... [VP]  
 ..... Rand McNally Atlas (Reprinted 1994) I-185  
 Viipuri..... [VP]  
 ..... Rand McNally Atlas (Reprinted 1994) I-185  
 Vyborg..... [BHA, VP Preferred]  
 ..... Canby, *Historic Places* (1984) II, 1005  
 ..... Encyclopaedia Britannica (1988) XII, 440  
 ..... Times Atlas of the World (1994) 211  
 ..... USBGN: Foreign Gazetteers  
 Wyborg..... [BHA]  
 ..... BHA, Authority file (1973-)  
**Subject:** ..... [BHA, VP]  
 ..... BHA, Authority file (1973-)  
 ..... Canby, *Historic Places* (1984) II, 1005  
 ..... Encyclopaedia Britannica (1988) XII, 440  
 ..... Rand McNally Atlas (Reprinted 1994) I-185  
 ..... Times Atlas of the World (1994) 211  
 ..... USBGN: Foreign Gazetteers

Kuva 3: Viipurin tietosivu Getty Thesaurus of Geographic Names -palvelussa

aasialaisiin paikkoihin, rahoittajana Carnegie-Mellonin säätiö.

Pelagios eroaa edellä esiteltyistä palveluista siten, että sen tavoitteena on kerätä hankkeeseen liittyviltä toimijoilta erilaisissa historiallisissa aineistoissa esiintyviä paikkaviittauksia, joille annetaan hankkeeseen liittyvässä Pleiades-nimistöluettelossa [6] yksilöivät URI-tunnisteet. Tavoitteena ei siis ole yksi yhtenäinen nimistöluettelo, vaan listaus viittauksista paikkojen ja erilaisten aineistojen välillä. Pleiades-nimistöluetteloon viittaville aineistoille ja niiden tietomalleille ei aseteta mitään rajoituksia, mutta hankkeeseen liittyviltä toimijoilta edellytetään, että ne julkaisevat Pleiades-viittausten tuloksena syntyvän paikkojen metatiedon julkisesti verkossa Pelagios-hankkeen määrittelemässä yhteisessä formaatissa.

Pelagios hyödyntää linkitetyn avoimen datan teknologioita, ja paikkaviittauk-

set mallinnetaan Open Annotation -sanaston [29] avulla. Tarjolla on myös HTTP-ohjelmointirajapinta, jonka avulla Pelagioksen sisältöä voidaan hyödyntää ulkopuolisissa sovelluksissa.

### 3.3 Yhteenveto historiallisen paikkatiedon palveluista

Tässä luvussa on esitelty joukko historiallisen paikkatiedon palveluita ja lähteitä, joissa on käytetty hyvin erilaisia ratkaisuja paikkatiedon mallinnuksessa ja käyttöliittymien suunnittelussa. Suuri osa palveluista keskittyy tietyn tarkasti rajatun historiallisen paikkatietoaineiston julkaisuun, jolloin käyttöliittymä on räätälöity juuri kyseiselle aineistolle sopivaksi. Tämä tekee ulkopuolisten aineistojen käyttöönoton hankalaksi. Palveluissa käytettyjen aineistojen kattavuuksien hahmottaminen ja vertailu on vaikeaa, koska Gazetteer for Scotland -palvelua lukuunottamatta kaikista muista palveluista puuttuu mahdollisuus selata aineistojen sisältöä kartalla. Pelkkään tekstihakuun tukeutuviissa palveluissa uudelle käyttäjälle palveluun ja aineistoihin tutustuminen on hyvin hankalaa, jos heti ensimmäiseksi täytyy tietää millä tietyillä hakusanoilla saa tuloksia esiin.

Linkitettyä dataa hyödyntävät palvelut ja lähteet mahdollistavat aineistojen hyödyntämisen ulkopuolisissa sovelluksissa. Sen sijaan esitellyissä perinteiseen teknologiaan perustuvissa palveluissa on hyvin rajoitetut mahdollisuudet saada selville aineistojen alkuperä tai hyödyntää niitä ulkopuolisissa sovelluksissa. Tämä on merkittävä rajoite, koska historiallisen paikkatiedon luonteeseen kuuluu, että paikkatieto on vääjäämättä hajautunut useisiin eri lähteisiin. Lisäksi paikkoihin liittyvää kontekstietoa on saatavilla vain Gazetteer for Scotland -palvelussa, mutta tätä kontekstietoa ei tuoteta dynaamisesti, vaan se löytyy paikkatietoaineistosta valmiina.



## 4 Historiallisten paikkojen ja karttojen palvelumalli

Tässä luvussa esitellään uusi linkitetyn datan teknologioihin perustuva historiallisten paikkojen ja karttojen palvelumalli, HIPLA. Aluksi esitetään visio siitä, minkälaista hyötyä mallin käytöllä voidaan saavuttaa. Seuraavaksi käydään läpi vaatimuksia, joita historiallisen paikkatiedon erityispiirteet asettavat HIPLA-palvelumallille. Tämän jälkeen pureudutaan mallin eri käyttäjäryhmiin ja näitä käyttäjäryhmiä vastaaviin konkreettisiin käyttötapauksiin. Luvun lopussa esitellään HIPLA-palvelumallin arkkitehtuuria yleisellä tasolla. HIPLA-palvelumallin mukaisesti toteutetun palvelun kuvaus ja tekniset ratkaisut on erotettu omaksi kokonaisuudekseen luvussa 5.

### 4.1 Visio

HIPLA-palvelumalli on tarkoitettu hajautettujen historiallisten sekä nykyisten paikkaontologioiden ja historiallisten karttojen hyödyntämiseen ja ylläpitoon. Tavoitteena on tarjota yhteinen näkymä eri organisaatioiden hallinnoimaan paikkatietoon. Linkitetyn datan ja web-teknologioiden avulla HIPLA-palvelumalli integroituu perinteisiin tiedonkuvailujärjestelmiin niin, että historiallista paikkatietoa hyödyntävät organisaatiot voivat kytkeä oman järjestelmänsä HIPLA-järjestelmään samaan tapaan kuin aiemmin Semanttisen laskennan tutkimusryhmässä toteutetussa ONKI-järjestelmässä [86].

Ideana on, että HIPLA-mallin avulla historiallista paikkatietoa voidaan kerätä erilaisten organisaatioiden ja käyttäjien toimesta niin, että kertaalleen jonkun tahon keräämä aineisto saadaan julkaistua linkitettynä avoimena datana kaikkien tahojen yhteisesti hyödynnettäväksi. Näin samaa työtä ei tarvitse tehdä turhaan moneen kertaan eri organisaatioissa, jolloin säästetään esimerkiksi kulttuurialan niukkoja resursseja ja mahdollistetaan eri tahoilla tuotettavan historiallisen paikkatiedon semanttinen yhteentoimivuus yhteisissä verkkopalveluissa ja historian tutkimuksessa.

### 4.2 Vaatimukset

Historiallisen paikkatiedon palveluun liittyy monia haasteita liittyen erityisesti maantieteelliseen tiedonhakuun sekä paikkatiedon ajalliseen ulottuvuuteen. Jones ja Purves [43] ovat tiivistäneet maantieteelliseen tiedonhakuun liittyviä haasteita seuraavasti:

1. maantieteellisten viittausten tunnistaminen teksteistä ja käyttäjien hakulausekkeista
2. käyttäjän tarkoittaman paikannimen yksikäsitteistäminen
3. epämääräisesti ilmaistujen maantieteellisten suhteiden tulkitseminen, esimerkiksi: “Viipurin lähellä”, “Kivennavan pohjoispuolella”, “Turun saaristossa” tai “Karjalan kannaksella”.
4. dokumenttien indeksointi sekä maantieteellisesti että niiden sisällön mukaan

5. dokumenttien relevanssin määrittäminen sekä maantieteellisesti että niiden sisällön mukaan
6. käyttäjäystävällisten, tehokkaiden hakukäyttöliittymien kehittäminen
7. maantieteellisten tiedonhaun evaluointimenetelmien kehitys

HIPLA-mallin vaatimukset liittyvät pääasiassa kohtiin 2, 4 ja 6. Tämän lisäksi HIPLA-malli pyrkii vastaamaan vaatimuksiin, joita hajautetun historiallisen paikkatiedon hyödyntäminen, päivittäminen ja ylläpito aiheuttavat. Seuraavissa aliluvuissa on kerrottu tarkemmin HIPLA-mallin vaatimuksista.

### **Paikannimellä määritetyn kohteen yksikäsitteistäminen**

Tekstimuotoisesta aineistosta voidaan viitata maanpäälliseen sijaintiin käyttämällä paikannimeä. Tällainen viittaus ei kuitenkaan ole yksikäsitteinen, koska paikannimet eivät ole yksilöiviä. Samaan paikkaan voidaan viitata eri nimillä ja murreilmaisuuilla eri yhteyksissä. Lisäksi sama tekstimuotoinen paikkaviittaus, esimerkiksi Saksa, voi tarkoittaa hyvin eri asiaa riippuen siitä, mitä historian hetkeä tarkoitetaan. Paikannimien yksikäsitteistämiseen tarvitaan siis muitakin tietoja kuin paikannimi – tärkeimpinä paikan tyyppi ja sijainti. Tämän vuoksi yksikäsitteistämistä tukevan palvelun täytyy pystyä tarjoamaan käyttäjälle mahdollisimman paljon paikannimiin liittyvää kontekstietoa, jotta palvelun käyttäjä voi valita oikean paikkatietoresurssin mahdollisesti kymmenien tai jopa satojen samannimisten paikkatietoresurssien joukosta.

### **Historiallisten karttojen tarve**

Historiallista paikkatietoa käsiteltäessä syntyy ilmeinen tarve historiallisille karttalehdille. Digitointihankkeet eri puolilla maailmaa ovat tuoneet suuren määrän skannattuja historiallisia karttoja digitaalisesti saataville. Pelkkä saatavuus ei kuitenkaan riitä, vaan historiallisia karttoja päästään toden teolla hyödyntämään vasta sitten, kun kartat ovat jollakin tapaa linkitetty olemassa oleviin paikkatietoaineistoihin.

Eräs tapa parantaa historiallisten karttalehtien löydettävyyttä on asemoida karttalehdet moderniin koordinaatistoon eli suorittaa karttalehtien georektifointi. Näin tietyn alueen karttoja eri aikakausilta voidaan hakea nykyisin saatavilla olevien maantieteellisen haun toimintojen avulla, mikä on suuri edistysaskel verrattuna esimerkiksi pelkkään tekstihakuun, joka kohdistuu kartan nimeen ja mahdollisesti puutteellisiin metatietoihin. Toki myös karttojen metatietoja tarvitaan, ja metatietojen tulisi olla saatavilla avoimen rajapinnan kautta.

### **Hajautetun historiallisen paikkatiedon hallinta**

Historiallisista paikoista on mahdotonta tehdä kattavaa globaalia nimistöluettelo, sillä historiallisen paikkatiedon lähteet, kuten nimistöluettelot, tesaurokset ja paikkaontologiat keskittyvät tyypillisesti jonkin tietyn alueen nimistöön, joten ne ovat

usein siiloutuneita ja hajallaan. Hyvin erityyppiset kulttuurialan organisaatiot käyttävät historiallisia paikannimiä; Suomessa esimerkiksi Kotimaisten kielten keskus, Suomalaisen Kirjallisuuden Seura, Svenska litteratursällskapet, Museovirasto, Kansallisarkisto, Kansalliskirjasto, museot ja sukuseurat pitävät hallussaan ja hyödyntävät paikkasidonnaista historiallista tietoa.

Siiloutuneisuus vaikeuttaa yhteistyön koordinoimista yli hallinnollisten organisaatorajojen. Käytössä on monia erilaisia historiallisia nimistöluetteloita, jotka käyttävät omia tunnisteitaan ja tietomallejaan ja kehittyvät toisistaan riippumatta. Mahdollisimman globaali kattavuus on kuitenkin historiallisen paikkatiedon hyödynnettävyyden kannalta selkeä vaatimus. Esimerkiksi suurten alueiden, kuten toisen maailmansodan tapahtumapaikkojen tai Rooman valtakunnan kattaminen ei onnistu ilman useiden eri organisaatioiden hallinnoimaa paikkatietoa. Lisäksi siiloutunut toimintapa johtaa usein päällekkäisen työn tekemiseen, kun eri organisaatiot ylläpitävät paikkatietoa samoilta alueilta tietämättä toisistaan. On myös huomattava, että historiallisen paikkatiedon ajallinen ulottuvuus tuo entistä suuremman vaatimuksen organisaatioiden väliseen yhteistyöhön verrattuna esimerkiksi ainoastaan nykyhetkeä koskevaan paikkatietoon.

Tarvitaan siis menetelmiä, joiden avulla voidaan hyödyntää eri lähteistä tulevaa historiallista paikkatietoa systemaattisesti ja keskitetysti, ilman että eri organisaatioiden hallussa olevan paikkatiedon erityispiirteitä menetetään.

### **Päivittyvien historiallisten paikkaontologioiden ylläpito**

Luettelointityössä kohdataan jatkuvasti uusia paikannimiä, joita ei löydy olemassa olevista järjestelmistä. Tämä synnyttää tarpeen dynaamisille ja nopeille ontologian päivitysmekanismille. Päällekkäisen työn välttämiseksi paikkaontologian päivityksien täytyy olla heti kaikkien käyttäjien saatavilla. Verrattuna nykyhetkeä kuvaavin paikkaontologioihin historialliset paikkaontologiat ovat usein puutteellisempia ja siten tarve uusien paikkainstanssien lisäämiselle on suurempi.

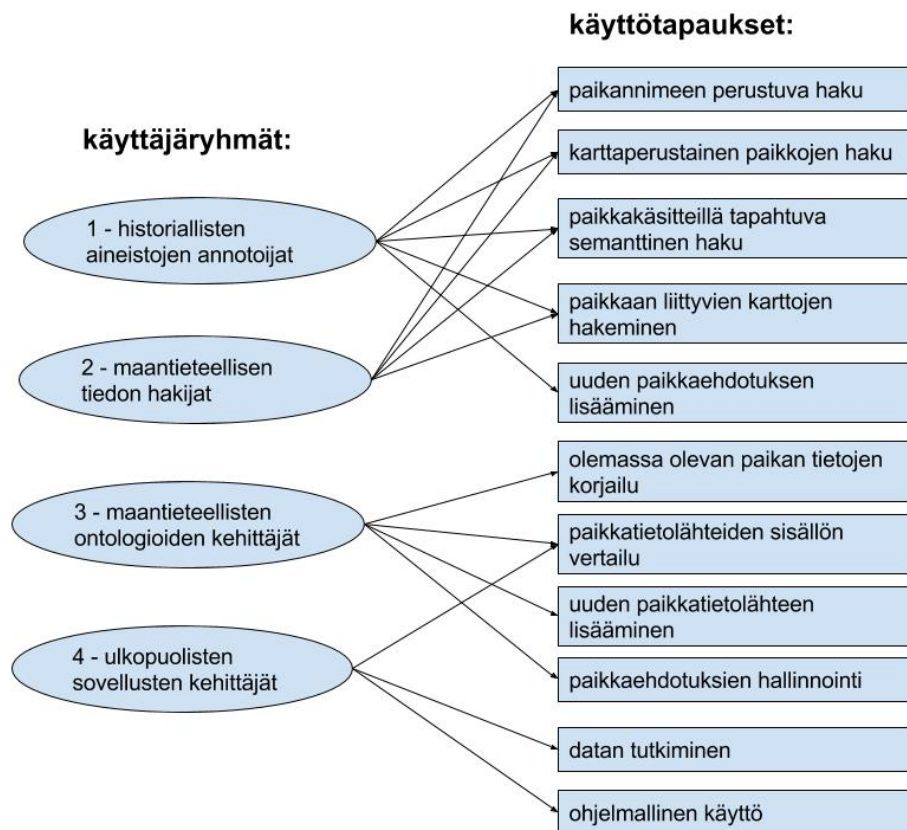
### **4.3 Käyttäjärühmät ja käyttötapaukset**

Käyttäjärühmiä voidaan käyttää apuna käyttötapauksien määrittelyssä. Työn edetessä HIPLA-mallin mukaan toteutettavalla palvelulle tunnistettiin lopulta 4 erilaista käyttäjärühmää:

1. historiallisten aineistojen annotoijat
2. maantieteellisen tiedon hakijat
3. maantieteellisten ontologioiden kehittäjät
4. ulkopuolisten sovellusten kehittäjät

Alkuperäisessä työsuunnitelmassa oli 4 käyttötapausta liittyen historiallisten aineistojen annotointiin, mutta ei erillisiä palvelun käyttäjärühmiä. Eri käyttäjärühmien tarpeiden perusteella alkuperäisiä käyttötapauksia tarkennettiin, ja uusia

käyttötapauksia lisättiin. Kuvassa 4 on esitetty eri käyttäjäryhmille muodostetut käyttötapaaukset. Käyttötapaaukset on kuvattu seuraavassa listauksessa tarkemmin.



Kuva 4: HIPLA-järjestelmän käyttötapaaukset

1. **Paikannimeen perustuva haku.** Automaattisesti täydentyvä tekstihaku, joka kohdistuu dynaamisesti valittavissa oleviin paikkatietolähteisiin, auttaa käyttäjää löytämään ja yksikäsitteistämään hakemansa paikan.
2. **Karttaperustainen paikkojen haku.** Silloin kun käyttäjä ei tiedä tarkkaa nimeä hakemalleen paikalle, tarvitaan karttihakua. Lisäksi karttihakua mahdollistaa aineistoihin tutustumisen ja paikkojen selailun ilman, että käyttäjän täytyy keksiä, mitä hakusanoja tulisi käyttää.
3. **Paikkakäsitteillä tapahtuva semanttinen haku.** Paikkaontologian ontologisten suhteita hyödynnetään teksti- ja karttahaussa. Esimerkiksi paikkaontologiassa määritellyt kuuluvuussuhteet tai paikkatyyppien hierarkia mahdollistavat kehittyneitä hakutoimintoja.
4. **Paikkaan liittyvien karttojen hakeminen.** Vanhoja skannattuja karttalehtiä on nykyisin saatavilla runsaasti, mutta niitä on vaikeaa löytää tai hyödyntää, ellei niitä ole georeferoitu. Tarjotaan helppokäyttöinen työkalu historiallisten

karttalehtien georeferointiin, jonka avulla karttalehdet saadaan piirrettyä modernin karttapohjan päälle.

5. **Uuden paikkaehdotuksen lisääminen.** Palveluun rekisteröityneelle käyttäjälle annetaan mahdollisuus ehdottaa paikkatietolähteistä puuttuvia paikkoja, ja nämä paikkaehdotukset ovat välittömästi kaikkien käyttäjien käytettävissä.
6. **Olemassa olevan paikan tietojen korjailu** Palveluun rekisteröityneelle käyttäjälle annetaan mahdollisuus ehdottaa korjauksia paikkatietolähteiden tietoihin.
7. **Paikkatietolähteiden sisällön vertailu.** Karttahakua ja dynaamisesti valittavia paikkatietolähteitä voidaan käyttää hajautettujen paikkatietoaineistojen sisällön ja kattavuuksien vertailuun.
8. **Uuden paikkatietolähteen lisääminen.** Uusia paikkatietolähteitä voidaan lisätä konfiguroimalla SPARQL-palvelupiste ja paikkatietolähteen tietomallin mukaiset SPARQL-kyselyt paikkojen hakutoimintoihin ja hakutulosten visualisointiin.
9. **Paikkaehdotuksien hallinnointi.** Paikkaehdotuksia hallinnoidaan luvussa 4.6 myöhemmin kuvattavan prosessin mukaisesti.
10. **Datan tutkiminen.** Dynaaminen paikkatietolähteiden valinta sekä teksti- ja karttahakutoiminnot tarjoavat monipuolisen työkaluvalikoiman paikkatietolähteiden sisällön tutkimiseen.
11. **Ohjelmallinen käyttö.** Ulkopuoliset tahot voivat hyödyntää HIPLA-järjestelmää joko ottamalla käyttöön HIPLA:n tarjoamia valmiita käyttöliittymäkomponentteja tai käyttämällä avointa SPARQL-ohjelmointirajapintaa.

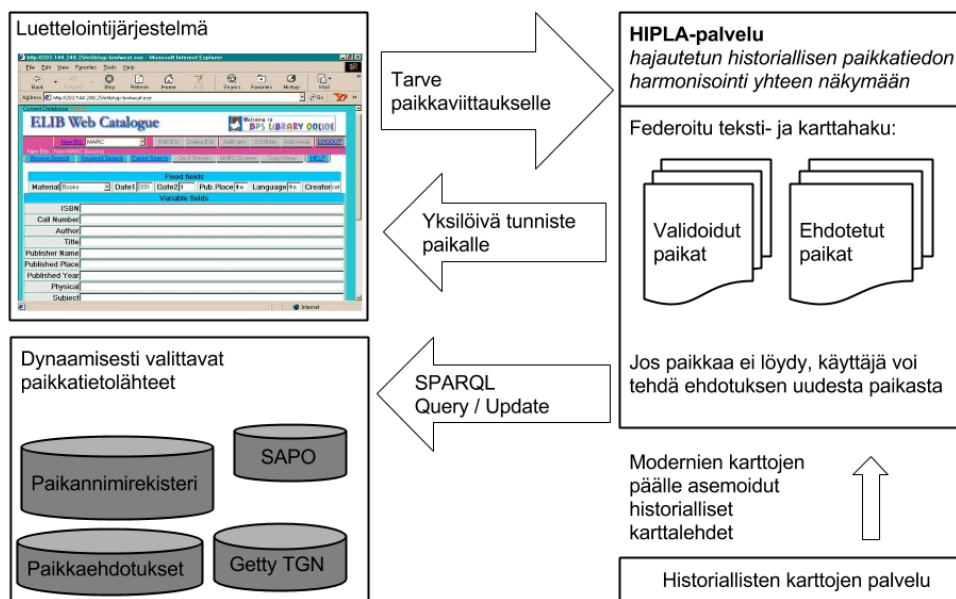
#### 4.4 Arkkitehtuuri

Kuvassa 5 on esitetty HIPLA-palvelumallin pääkomponentit. Vasemmasta alakulmasta löytyvät paikkatiedon lähteinä toimivat julkiset tai yksityiset SPARQL-palvelupisteet. Paikkatiedon lähteiden yläpuolella on perinteinen luettelointi- tai annotointijärjestelmä, jossa syntyy tarve tehdä viittaus paikkaan. Luettelointijärjestelmä kytkeytyy HIPLA:an esimerkiksi HIPLA:n tarjoamalla integroitavalla käyttöliittymäkomponentilla, joka mahdollistaa automaattisesti täydentyvän federoidun tekstihaun paikkatietolähteisiin. HIPLA toimii välityspalveluna, joka noutaa hakutulokset kaikista kulloinkin käyttöön määritellyistä SPARQL-palvelupisteistä ja kerää tulokset keskitetysti yhteen, jotta järjestelmän käyttäjä saa näkyviin kaiken käytettävissä olevan, mahdollisesti ristiriitaisenkin tiedon hakemastaan paikannimestä.

HIPLA on siis pohjimmiltaan kyselyjen välityspalvelu, eikä sen tarkoituksena ole säilöä paikkatietoaineistoja. Palvelu sisältää kuitenkin sisäisen paikkaehdotuksien tietokannan (käytännössä myös SPARQL-palvelupiste), jotta käyttäjät voisivat ehdottaa paikkatietoaineistoista puuttuvia paikkoja tai korjauksia nykyisin tietoihin.

Lisäksi HIPLA käyttää ulkoista karttojen georektifointipalvelua (engl. georectification service), jonka avulla HIPLA:n käyttöliittymään voidaan tuoda historiallisia karttalehtiä asemoituna nykyisiä koordinaattijärjestelmiä käyttävien karttojen päälle.

Historiallisten karttalehtien metatiedot tallennetaan RDF-tietokantaan omaan graafiinsa. Tämä mahdollistaa kehittyneiden karttalehtien hakupalveluiden rakentamisen SPARQL-palvelupisteen avulla. Historiallisten karttojen palvelu puolestaan tarjoaa karttalehtien yksilöllisten tunnisteiden avulla varsinaiset asemoidut karttakuvat oman rajapintansa kautta HIPLA-palvelun käyttöön.



Kuva 5: HIPLA-mallin arkkitehtuuri

## 4.5 Historiallisen paikkatiedon harmonisointi

Koska eri SPARQL-palvelupisteiden kautta käytettävillä paikkatiedon lähteillä, eli käytännössä nimistöluetteloilla, on kullakin oma yksilöllinen tietomallinsa, HIPLA-palvelumallissa ei voida tehdä yleisiä oletuksia sille, millä tavoin historiallista paikkatietoa on mallinnettu. Luvussa 2.4 esitettyjen seikkojen perusteella voidaan kuitenkin olettaa, että saatavilla on aina vähintään paikannimi ja paikkatyyppe, useimmiten lisäksi sijaintitietoa. Jo näiden kolmen ominaisuuden avulla on mahdollista yhdistellä eri lähteistä tulevaa, erilaisia tietomalleja käyttävää paikkatietoa yhtenäiseen teksti- ja karttahaun tarjoavaan käyttöliittymään, menettämättä eri tietomallien erityispiirteitä. Käytännössä tämä paikkatiedon harmonisointi tapahtuu SPARQL-kyselyillä: jokaiselle paikkatietolähteelle konfiguroidaan kyseistä tietomallia vastaava SPARQL-kysely. Näin voidaan esimerkiksi huomioida mahdollisesti saatavilla oleva paikkatiedon ajallisen ulottuvuuden mallinnus.

Paikkatietolähteille erikseen konfiguroidut SPARQL-kyselyt tuottavat erimuotoisia vastauksia, joten myös hakutulosten käsittely ja visualisointi täytyy käsitellä

käyttöliittymässä lähdekohtaisesti, mutta samalla hyödyntäen kaikille lähteille yhteisiä tietoja, kuten paikkatyyppejä ja koordinaatteja.

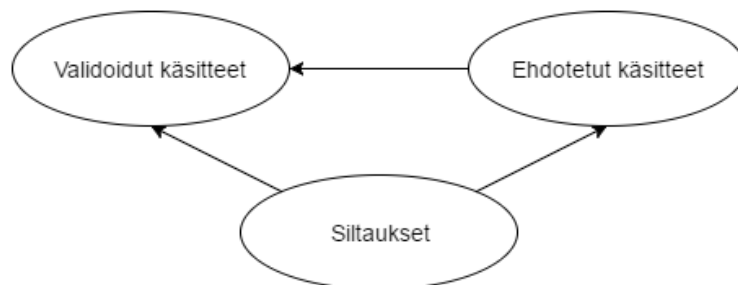
## 4.6 Kehittyvän ontologian hallinnointiprosessi

HIPLA-mallissa pyritään vastaamaan johdantoluvussa esitettyyn tutkimuskysymykseen “Miten paikkatietoaineiston jatkuva tuottaminen ja täydentäminen voidaan hajauttaa aineiston käyttäjille reaaliaikaisesti aineiston käyttötilanteessa?” ottamalla käyttöön kehittyvän ontologian hallinnointiprosessi, joka perustuu Semanttisen laskennan tukimusryhmän julkaisussa [37] esiteltyyn hallinnointiprosessiin.

Hallinnointiprosessiin kuuluu kolme käyttäjäryhmää:

1. **Ylläpitokäyttäjät**, joilla on oikeus lisätä, poistaa ja muokata ontologian käsitteitä.
2. **Rekisteröidyt käyttäjät**, jotka voivat lisätä ehdotuksia uusiksi käsitteiksi.
3. **Peruskäyttäjät**, jotka voivat hyödyntää ontologiaa sellaisenaan.

Ontologian infrastruktuuri jakaantuu kolmeen osaan, jotka on esitetty kuvassa 6. Ensimmäinen osa, *validoidut käsitteet*, sisältää ylläpitokäyttäjien hyväksymät ja validoimat käsitteet. Toinen osa, *ehdotetut käsitteet*, koostuu käsitteistä, joita rekisteröidyt käyttäjät ovat ehdottaneet, ja joita muut käyttäjät ovat mahdollisesti jo käyttäneet annotoinnissa, mutta joita ylläpitokäyttäjät eivät ole vielä hyväksyneet. Kolmas osa, *siltaukset*, sisältää siltauksia ehdotetuista käsitteistä validoituihin käsitteisiin.



Kuva 6: Kehittyvän ontologian hallinnointiprosessi

Kehittyvän ontologian hallinnointiprosessissa ontologia kehittyy joukkoistamalla ontologian rekisteröityjen käyttäjien toiminta. Seuraavassa listauksessa on esitetty ontologian kehittymisen vaiheet.

1. Rekisteröitynyt käyttäjä etsii tarvitsemaansa käsitettä esimerkiksi käyttäen automaattisesti täydentyvää tekstihakua. Haun tuloksissa näytetään sekä validoituja että ehdotettuja käsitteitä omissa kategorioissaan. Näin käyttäjä näkee heti, löytyykö hänen tarvitsemansa käsite jo ontologiasta, tai onko joku muu ehdottanut kyseistä käsitettä lisättäväksi ontologiaan. Sekä validoituja että ehdotettuja käsitteitä voi käyttää annotointiin.

2. Jos haluttu käsite löytyy validoiduista tai ehdotetuista käsitteistä, annotointi on valmis.
3. Muussa tapauksessa rekisteröitynyt käyttäjä voi tehdä ehdotuksen uudesta käsitteestä. Tällöin täytyy määritellä, mihin järjestelmään kytkettyyn ontologiaan käsitettä ehdotetaan. Valittu ontologia määrittelee edelleen sen, millaiset metatiedot ehdotetulle käsitteelle täytyy syöttää. Usein esiintyviä metatietoja ovat esimerkiksi nimi, koordinaatit ja paikkatyypit. Kun paikkaehdotus lisätään, järjestelmä luo automaattisesti ehdotetulle käsitteelle pysyvän yksilöivän tunnisteiden.
4. Järjestelmä lisää ehdotuksen ehdotettujen käsitteiden graafiin, jolloin ehdotus on välittömästi kaikkien käyttäjien saatavilla ja käytettävissä annotointiin.
5. Ylläpitokäyttäjä käy läpi uudet ehdotetut käsitteet tasaisin väliajoin.
6. Jos ehdotettu käsite on validi, ylläpitokäyttäjä siirtää sen ehdotuksen kohteena olevan ontologian validoitujen käsitteiden graafiin. Tässä vaiheessa ylläpitokäyttäjä voi luoda ehdotetulle käsitteelle uuden tunnisteiden, joka on kohteena olevan ontologian sääntöjen mukainen. Lisäksi ylläpitokäyttäjä voi muokata ehdotetun käsitteen metatietoja. Koska ehdotetun käsitteen aiempaa tunnistetta on voitu jo aiemmin käyttää annotoinnissa muiden käyttäjien toimesta, sitä ei ole syytä poistaa. Sen sijaan järjestelmä luo uuden siltauksen siltausgraafiin. Siltaus on RDF-kolmikko (ehdotetun käsitteen URI, siltaus, validoidun käsitteen URI), jossa siltaus on predikaatti, joka kuvaa ehdotetun ja validoidun käsitteen suhdetta. Siltausten ideana on siis tukea käyttäjiä, jotka ovat käyttäneet annotoinnissa ehdotettuja käsitteitä, koska siltausten avulla nämä annotoinnit voidaan jälkikäteen sillata validoituihin käsitteisiin.
7. Jos ehdotettu käsite ei ole validi, ylläpitokäyttäjä joko merkitsee sen hylätyksi, jolloin sitä ei voi enää käyttää annotointiin, tai lisää uuden siltauksen, joka osoittaa mitä käsitettä tulisi käyttää hylätyn käsitteen sijaan.



## 5 Palvelumallin prototyypitoteutus: Hipla.fi

Tässä luvussa esitellään HIPLA-palvelumalliin perustuva prototyyppi Hipla.fi: suomalainen historiallisten paikkojen ja karttojen ontologiapalvelu ja palvelun käytössä olevat paikkatietolähteet. Luvun alkupuoli keskittyy HIPLA-palvelumallin mukaisen käyttöliittymän toiminnallisuuksien kuvailemiseen, ja lopuksi kuvaillaan näiden toiminnallisuuksien tekninen toteutus.

### 5.1 Käyttöliittymän yleiskuvaus

Hipla.fi-palvelun HTML5-sovelluksena toteutettu käyttöliittymä on käytettävissä osoitteessa <http://hipla.fi>. Käyttöliittymän pääsivu on esitetty kuvassa 7. Käyttöliittymä on jaettu kahteen osaan: vasemmassa reunassa sijaitsee työkalupalkki, ja oikealla on karttanäkymä. Koska karttanäkymä on palvelussa pääosassa, sille annettiin enemmän tilaa leveyssuunnassa kuin työkalupalkille. Kuvaan 7 on merkitty punaisella tekstillä tärkeimmät HIPLA-palvelumallin mukaiset käyttötapaukset, jotka toteutettiin tämän työn puitteissa.

Ensimmäinen toiminnallisuus kuvan 7 vasemmassa laidassa sijaitsevassa työkalupalkissa on dynaaminen paikkatietolähteiden valinta. Käyttäjä voi valita palvelun käyttämät data-lähteet rastittamalla data-lähteiden valintaruutuja. Paikkatietolähteiden valinta voi tapahtua myös kesken käyttäjän tekemiä hakuja, jolloin hakutulokset päivittyvät automaattisesti uusien data-lähteiden mukaisiksi. Tämä oli mahdollista toteuttaa helposti asiakaspään JavaScript-koodissa, koska paikkatietolähteinä toimivat SPARQL-palvelupisteet. Toinen toiminnallisuus on uusien paikkojen ehdottaminen, joka avaa lomakkeen paikan tietojen syöttämistä varten. Kolmas toiminnallisuus työkalupalkin yläosassa on karttahaku, jonka avulla käyttäjä voi selata valittujen paikkatietolähteiden sisältöä karttanäkymässä.

Työkalupalkin alalaidassa on kaksi erilaista hakutoiminnallisuutta: automaattisesti täydentyvä tekstihaku paikoille sekä käyttäjän määrittämää aluetta kattavien historiallisten karttalehtien haku. Tekstihaun hakutulosriveiltä löytyy painike linkitetyn datan selailulle ja semanttiselle haulle. Tässä työssä toteutettiin esimerkki semanttisesta hausta, joka hakee tietyn kunnan rajojen sisäpuolelta löytyvät paikat. Semanttinen haku on käytössä ”Suomen kunnat 1939–1944”-paikkatietolähteelle.

Kuvan 7 oikean laidan karttanäkymää käytetään karttahakuun, tekstihaun tulosten visualisointiin sekä paikkoihin liittyvän kontekstiedon tarjoamiseen. Karttanäkymän vasemmasta yläkulmasta löytyy Googlen karttaan liittyvät painikkeet, joilla voi valita käytetäänkö karttapohjana maastokarttaa vai satelliittikuvaa. Oikeasta yläkulmasta löytyy säädin, jolla voi muuttaa historiallisten karttalehtien läpinäkyvyyttä Googlen karttapohjan päällä. Vasemmassa alakulmassa ovat Googlen kartan zoomauspainikkeet sekä kontekstikartta, jonka avulla käyttäjä voi hahmottaa karttanäkymän sijainnin zoomaustasosta huolimatta.

**hipla.fi** Finnish Ontology Service of Historical Places and Maps

About Project home

Select source dataset(s)

- Finnish municipalities (1939-44)
- Karelian map names (1922-44)
- Finnish Geographic Names (contemp.)
- SAPO (1865-2010)
- Getty TGN
- Suggested places

+ Add a new place

View all places on current map view

Search places **Maps**

kiviniemi

**Karelian map names (1922-44)**

- Kiviniemi (kylä, Sakkola)
- Kiviniemi (maastokohde, Kurkijoki)
- Kiviniemi (maastokohde, Lumivaara)
- Kiviniemi (maastokohde, Salmi)
- Kiviniemi (rakennettu kohde, Saari)
- Kiviniemi (maastokohde, Salmi)
- Kiviniemi (maastokohde, Uukuniemi)
- Kiviniemi (maastokohde, Impilahti)
- Kiviniemi (maastokohde, Suistamo)
- Kiviniemi (maastokohde, Suistamo)
- Kiviniemi (rakennettu kohde, Suojärvi)
- Kiviniemi asema (rakennettu kohde, Sakkola)

**Finnish Geographic Names (contemp.)**

- Kiviniemi (Talo, Parainen)
- Kiviniemi (Talo, Salo)
- Kiviniemi (Talo, Parainen)

**dynaaminen paikkatietolähteiden valinta**

**uusien paikkojen ehdottaminen**

**karttahaku**

**historialliset karttalehdet**

**automaattisesti täydentyvä, federoitu tekstihaku**

**semanttinen haku paikkakäsitteillä**

**kontekstittieto ja linkitetyn datan selaus**

**Kiviniemi**  
 type: kylä  
 broader area: Sakkola  
 source: Karelian map names (1922-44)  
 View in linked data browser  
 Related linked data from the WarSampo project:  
 Events  
 Photographs

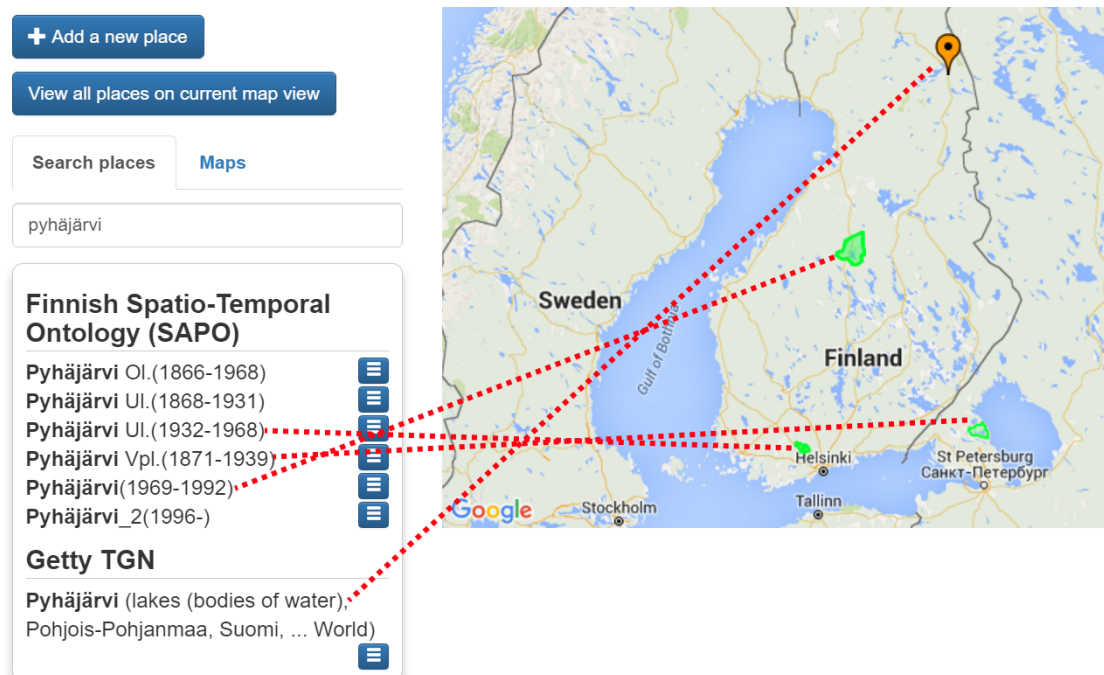
Kuva 7: Hipla.fi-palvelun käyttöliittymän yleiskuvaus

## 5.2 Automaattisesti täydentyvä ja federoitu tekstihaku

Kun käyttäjä on kirjoittanut tekstihakukenttään vähintään kaksi kirjainta, tehdään kysely kaikkiin kulloinkin valittuna oleviin paikkatietolähteisiin. Hakutulokset listataan tekstihakukentän alle paikkatietolähteittäin ryhmiteltyinä. Kirjoitushetkellä käytössä olevien paikkatietolähteiden hakutulosriveiltä löytyy paikkojen yksikäsitteistämistä varten aina vähintään paikannimi ja paikkatyyppi sekä painike, jolla pääsee katsomaan kyseisen paikan linkitettyä dataa SAHA-editorilla [62]. Lisäksi hakutulosrivillä näytetään kunta, lääni tai muu laajempi alue, johon kyseinen paikka kuuluu, jos tällainen tieto on saatavilla.

Ryhmitellyn listauksen lisäksi hakutulokset visualisoidaan aina kartalle. Myös tämä on tärkeä ominaisuus paikannimien yksikäsitteistämistä varten. Kuvassa 8 näkyy kahden eri paikkatietolähteen ”Pyhäjärvi”-hakusanan tulosrivit ja tulosrivejä vastaavat karttamerkit. Kun käyttäjä vie cursorin hakutulosrivin päälle, tätä hakutulosta vastaava karttamerkki liikkuu kartalla. Näin käyttäjä voi helposti ja systemaattisesti käydä läpi kaikki samalla hakusanalla eri lähteistä palautuvat tulokset ja nähdä, miten ne sijoittuvat kartalla.

Valitsemalla hakutulosrivin karttanäkymä siirtyy automaattisesti kyseisen paikan kohdalle, ja karttamerkin infoikkuna aukeaa. Infoikkunan painikkeista on kerrottu tarkemmin luvussa 5.6. Paikkatietolähteiden valintoja voi muuttaa kesken hakutulosten tutkimisen, jolloin hakutulokset päivittyvät automaattisesti.



Kuva 8: Tekstihaun tulosten yksikäsitteistäminen

### 5.3 Karttahuaku

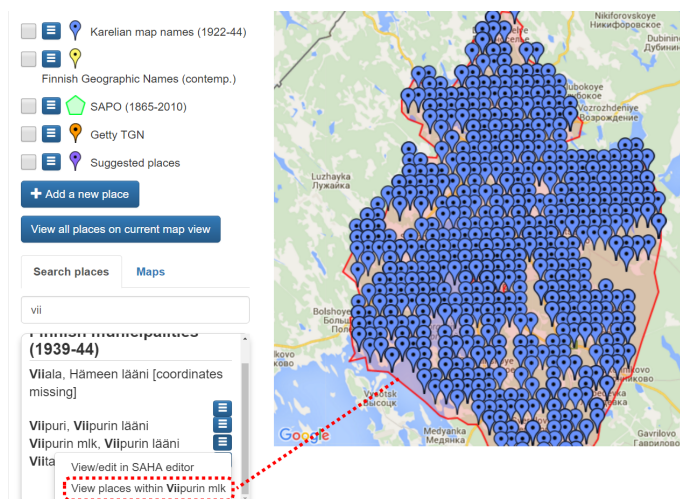
Karttahuakua tarvitaan silloin, kun käyttäjä ei tiedä hakemansa paikan tarkkaa nimeä eikä siten voi käyttää tekstihuakua. Tekstihuun tapaan karttahuussa on ensin mahdollista valita käytössä olevat paikkatietolähteet. Tämän jälkeen karttanäkymä on Google Maps API:n tarjoamien zoomaus- ja vieritystoimintojen avulla helppo siirtää halutulle alueelle, jonka jälkeen käyttäjä voi “View all places on current map view”-painiketta käyttämällä saada kaikki sen hetkiselältä karttanäkymältä löytyvät paikat kartalle tutkittavaksi. Karttamerkkien värikoodaus mahdollistaa karttahuun käyttämisen myös eri paikkatietoaineistojen vertailuun ja niiden kattavuuksien vertailuun.

Koska osa paikkatietolähteistä sisältää miljoonia paikkainstansseja, karttahuakua varten toteutettiin rajoitus, joka pyytää käyttäjää siirtymään vähintään zoomaustasolle 12, ennen kuin kaikki karttanäkymään sisältyvät paikat näytetään. Rajoitus estää liian suuria määriä vastauksenaan palauttavien kyselyiden syntymisen.

### 5.4 Semanttinen haku paikkakäsitteillä

Teksti- ja karttahuun lisäksi paikkaontologioille voidaan tietomallista riippuen toteuttaa semanttinen haku. Hipla.fi-prototyyppeihin toteutettiin esimerkki semanttisesta hausta ”Suomen kunnat 1939–1944”-paikkatietolähteelle. Hakemalla ensin tekstihaualla haluamansa kunnan käyttäjä voi valita kartalle näkyviin kaikki kyseisen kunnan rajojen sisäpuolella sijaitsevat paikat.

Kuvassa 9 on tehty ensin tekstihaku “vii”, ja klikattu hakutulostin “Viipurin mlk” valikkopainiketta ja valittu “View places within Viipurin mlk”, jolloin kartalle on saatu käyttäjän tutkittavaksi kaikki Viipurin maalaiskunnan rajojen sisällä olevat paikat. Tieto rajojen sisäpuolella sijaitsevista paikoista löytyy valmiiksi ”Suomen kunnat 1939–1944”-paikkatietolähteen RDF-datasta, mutta sama toiminnallisuus olisi ollut mahdollista toteuttaa myös dynaamisesti luvussa 6.1 esitellyillä keinoilla.



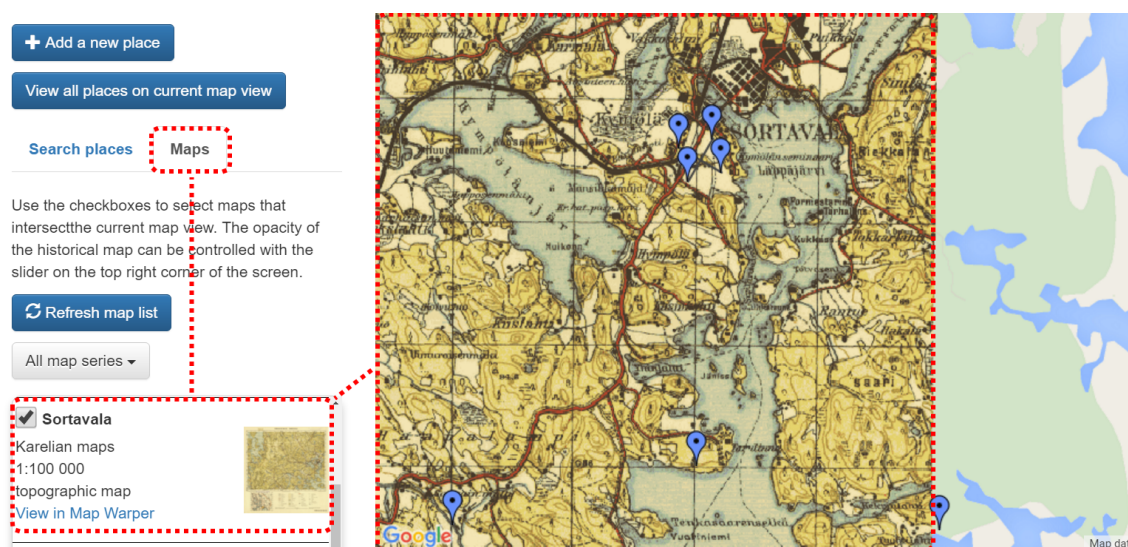
Kuva 9: Semanttinen haku “Viipurin mlk”-paikkakäsitteellä

## 5.5 Georektifioitujen historiallisten karttalehtien hyödyntäminen

Käyttöliittymään integroitu historiallisten karttalehtien georektifointi- ja hakupalvelu Map Warper [91] tarjoaa HIPLA-mallin vision mukaisesti käyttäjälle mahdollisuuden tarkastella hakutuloksia modernin karttapohjan päälle asemoiduilla historiallisilla karttalehdillä. Map Warper -palvelun integroinnin teknisistä yksityiskohdista on kerrottu tarkemmin luvussa 5.8.

Avaamalla “Maps”-välilehden käyttäjä saa listauksen historiallisista karttalehdistä, jotka leikkaavat sen hetkistä karttanäkymää. Listauksessa näytetään jokaisen karttalehden kohdalla nimi, karttasarja, mittakaava sekä lisäksi esillä on linkki, josta kyseistä karttalehteä pääsee tutkimaan Map Warper -palvelussa. Karttalehden nimen vieressä sijaitsevan valintaruudun avulla karttalehti piirtyy Googlen kartan päälle. Karttanäkymän oikeasta yläkulmasta löytyy liukusäädin, jolla voi säätää historiallisen kartan läpinäkyvyyttä. Näin käyttäjä voi hakutuloksia tutkiessaan joustavasti vertailla historiallista ja nykyajan karttaa. Jos käyttäjä siirtää karttanäkymää, hän voi päivittää historiallisten karttalehtien listauksen “Refresh map list”-painikkeella.

Karttanäkymässä rajaamisen lisäksi historiallisia karttoja voi suodattaa karttasarjan perusteella. Oletuksena näytetään kaikkien palvelusta löytyvien karttasarjojen karttoja, mutta “All map series”-painikkeella voi valita tietyn karttasarjan. Kaikki yhden karttasarjan karttalehdet saa esiin valitsemalla ensin karttasarjan ja sen jälkeen käyttämällä “Select all”-painiketta. Kuvassa 10 on hakutuloksien tutkimisen avuksi valittu “Karjalan kartat 1:100 000 topografinen kartta”-karttasarjasta Sortavalan karttalehti, jonka avulla saadaan Googlen karttapohjalta puuttuva historiallinen paikannimistö esiin.

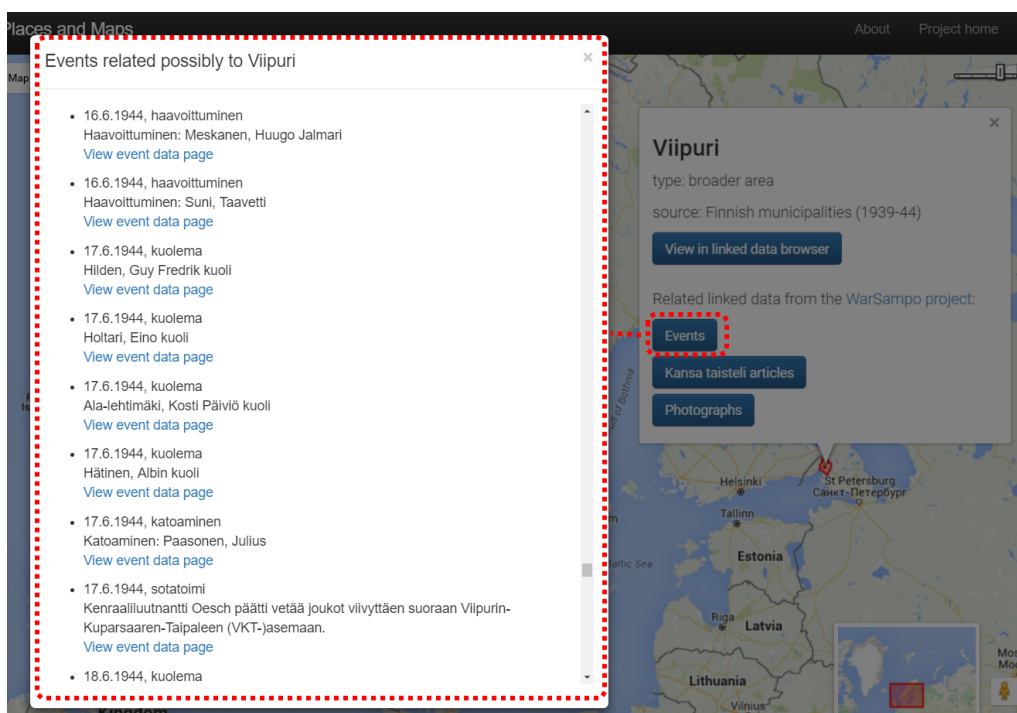


Kuva 10: Hakutuloksien tutkiminen Sortavan alueen historiallisen karttalehden avulla

## 5.6 Kontekstiedon tarjoaminen muista avoimen linkitetyn datan lähteistä

Kun käyttäjä valitsee Hipla.fi-käyttöliittymässä joko hakutulostin tai karttamarkin, kyseisen paikan infoikkuna aukeaa. Kuvassa 11 valittuna on talvi- ja jatkosodan paikkaontologiaan kuuluva Viipuri-instanssi, jonka kuvan oikeassa laidassa sijaitsevan infoikkunan yläosaan on kerätty yleistiedot paikasta: paikkatyyppi, mahdollinen laajempi alue sekä paikkatiedon lähde. Perustietojen alla on aina “View in linked data browser”-painike, jolla pääsee selaamaan paikan tietoja linkitetyn datan selaimella. Uuteen selainikkunaan aukeava selain riippuu siitä, mikä paikkatietolähde on kyseessä. Esimerkiksi Linked Data Finland -portaalissa julkaistujen paikkatietoaineistojen linkitetyn datan selailu tapahtuu SAHA-editorissa, kun taas muualla julkaistujen paikkatietoaineistojen linkitetyn datan selailun toteutus riippuu paikkatiedon julkaisijasta.

Linkitetyn datan selailun lisäksi ”Suomen kunnat 1939-1944”- ja ”Karjalan paikat”-aineistoille tarjotaan paikkoihin liittyvää kontekstietoa, kuten esimerkiksi paikkaan liittyviä historiallisia tapahtumia tai valokuvia Sotasampo-datapalvelusta. Kuvassa 11 infoikkunasta on valittu Viipuriin linkitetyt historialliset tapahtumat. Historialliset tapahtumat aukeavat omaan ikkunaansa, jossa niistä näytetään perustiedot, kuten tapahtuman päivämäärä, tyyppi ja kuvausteksti. Lisäksi jokaisen tapahtuman kohdalta löytyy linkki, jonka avulla käyttäjä voi siirtyä tutkimaan tapahtumaa Sotasampo-portaalissa. Sotasammosta ja sen datapalvelusta haettavasta kontekstiedosta on kerrottu lisää luvussa 6.



Kuva 11: Viipurin liittyvää kontekstietoa: Talvi- ja jatkosodan tapahtumia Sotasampo-portaalista

## 5.7 Puuttuvien paikkojen yhteisöllinen ehdottaminen

HIPLA-palvelumallin kehittyvän ontologian hallinnointiprosessin (esitelty luvussa 4.6) mukaisesti Hipla.fi-käyttöliittymään toteutettiin käyttäjille mahdollisuus lisätä uusia paikkaehdotuksia, jos käyttäjän tarvitsema paikka ei löydy palveluun kytketyistä paikkatietolähteistä tai aiemmista ehdotuksista.

Paikkaehdotuksia voi lisätä ”Add a new place”-painikkeella, jota käyttämällä avautuu kuvan 12 mukainen lomake. Lomakkeen avulla käyttäjä voi syöttää uuden paikkaehdotuksen perustiedot, joita ovat:

- ensisijainen paikannimi
- toissijaiset paikannimet
- kuvausteksti
- kunta
- laajempi alue
- paikkatyyppi
- kommentti
- ehdottajan nimi
- koordinaatit
- paikkatietolähde, johon ehdotus kohdistuu

Paikkaehdotuksen koordinaatit voi syöttää osoittamalla kuvan 12 oikean reunan karttanäkymää. Koordinaattien määrittelyssä voi käyttää apuna karttanäkymää sivuavia georektifioituja historiallisia karttalehtiä (”Search historical maps intersecting the map view”), mikä mahdollistaa moderneilta kartoilta puuttuvan paikannimistön lisäämisen paikkaontologioihin. Paikkaehdotus tallennetaan ”Save changes”-painikkeella, jonka jälkeen paikkaehdotus on välittömästi käytettävissä kaikille palvelun käyttäjille, esimerkiksi annotointityötä varten.

Koska Hipla.fi-palveluun kytketyissä eri paikkatietolähteissä on käytössä erilaisia tietomalleja paikkatiedon mallinnukseen, paikkaehdotuslomakkeelle kerättiin yleisluontoiset ja useimmiten paikkatietolähteistä löytyvät tiedot. Tulevaisuudessa lomake on mahdollista räätälöidä halutuille paikkatietolähteille erikseen, jolloin lomakkeen tiedoissa voidaan ottaa huomioon tietomallien erityispiirteet. Kehittyvän ontologian hallinnointiprosessiin kuuluvia toiminnallisuuksia, kuten paikkaehdotuksien validointia, siltausta ja hyväksyntää virallisiin paikkaontologioihin, ei toteutettu tämän työn puitteissa Hipla.fi-prototyypin.

**Place name**

**Alternative name**

**Description**

**Municipality**

**Broader area**

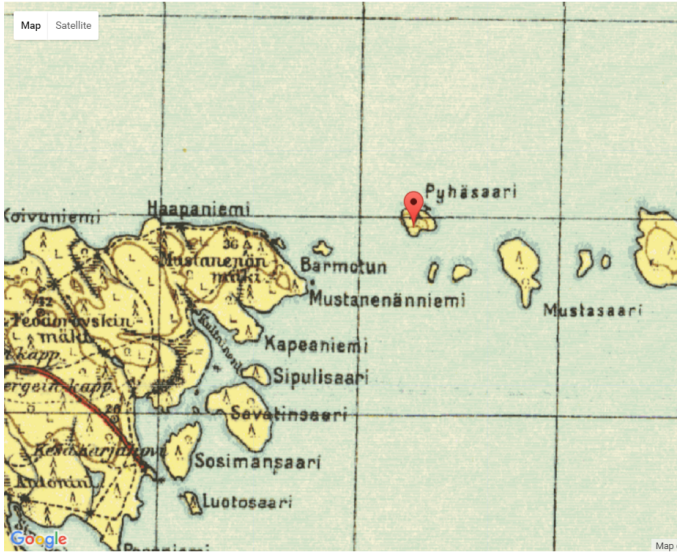
**Place type**

**Comment**

**Suggestor**

**Coordinates (click on the map)**

**Target dataset**



Map Satellite

Search historical maps intersecting the map view Hide historical maps

Historical maps from Map Warper, intersecting the current map view

Use the slider on the top right corner of the map to control the opacity of the historical map.

Sortavala Lahdenpohja Käkisalmi **Valamo** Impilahti Käsnäselkä Salmi

Kuva 12: Uuden paikkaehdotuksen lisääminen

## 5.8 Tekninen toteutus

### Käyttöliittymä

Hipla.fi-palvelun käyttöliittymä suunniteltiin ja toteutettiin Single-page Application -arkkitehtuurin (SPA) [65] mukaisesti, jolloin käyttöliittymä aukeaa aina käyttäjän selainikkunan kokoiseksi ilman vierityspalkkeja. Ensimmäisen sivulatauksen jälkeen uusia sivulatauksia ei tehdä, vaan kaikki toiminnallisuus tapahtuu samalla sivulla. Näin saavutetaan sujuva käyttökokemus, koska sivun sisällön muuttuessa ei tarvitse ladata koko sivua uudelleen. Hakukoneoptimoinnin kannalta Single-page Application -arkkitehtuuri ei ole paras mahdollinen ratkaisu, koska kaikille sivuille ei automaattisesti synny omaa URL-osoitetta. HTML5 History -ohjelmointirajapinnan [66] avulla on kuitenkin mahdollista luoda URL-osoitteita SPA-sovelluksen sisällä, mikä mahdollisti muun muassa omien URL-osoitteiden luomisen paikkainstanssien infosivuille.

Käyttöliittymän ensimmäisen version pääasiallisena JavaScript-kirjastona käytettiin ainoastaan jQueryä [44], joka on käytännössä vain kokoelma funktiota, joita kehittäjä voi kutsua haluamallaan tavalla. Tähän valintaan päädyttiin sen vuoksi, että käyttöliittymän toteutusta ei haluttu ensimmäisessä vaiheessa sitoa vielä mihinkään tiettyyn JavaScript-sovelluskehikseen, joka olisi määrännyt jonkin spesifisen tavan käyttöliittymän toiminnallisuuksien toteuttamiselle. Tulevaisuudessa lähdekoodia muokkaamalla on toki mahdollista siirtyä käyttämään esimerkiksi AngularJS-sovelluskehystä [26].



Käyttöliittymän ulkoasun rakentamiseen valittiin Bootstrap [10], joka on yksi tämän hetken suosituimmista asiakaspuolen sovelluskehyksistä. Bootstrap tarjoaa runsaasti valmiita käyttöliittymää varten tarvittavia elementtejä, kuten painikkeita ja valikoita, mikä nopeuttaa käyttöliittymän kehitystä, mutta toisaalta voi johtaa siihen että samaa sovelluskehystä käyttävistä sovelluksista tulee liikaa toisiaan muistuttavia. Yhtenäinen ja tuttu ulkoasu voi kuitenkin olla myös etu käyttökokemuksen muodostuessa. Valmiiden elementtien lisäksi Bootstrap tarjoaa responsiivisen [21] ruudukon, jonka avulla käyttöliittymän elementit voidaan asemoida sivulle. Minimivaatimukseksi asetettiin käyttöliittymän ensimmäisessä versiossa tabletti-kokoluokan laite, joten elementtien sijoittelu optimoitiin vähintään 1024x768-resoluutiota tukeville näytöille. Tähän vaatimukseen päädyttiin, koska monipuolisen hakukäyttöliittymän ja kartan yhdistäminen toimivaksi kokonaisuudeksi tätä pienempien näyttöjen kokoluokassa on hyvin haastavaa ja aikaa vievää, ja oli perusteltua olettaa että karttapainotteisen Hipla.fi-palvelun käyttäjät haluavat käyttää palvelua matkapuhelimen näyttöä suuremmilla ruuduilla.

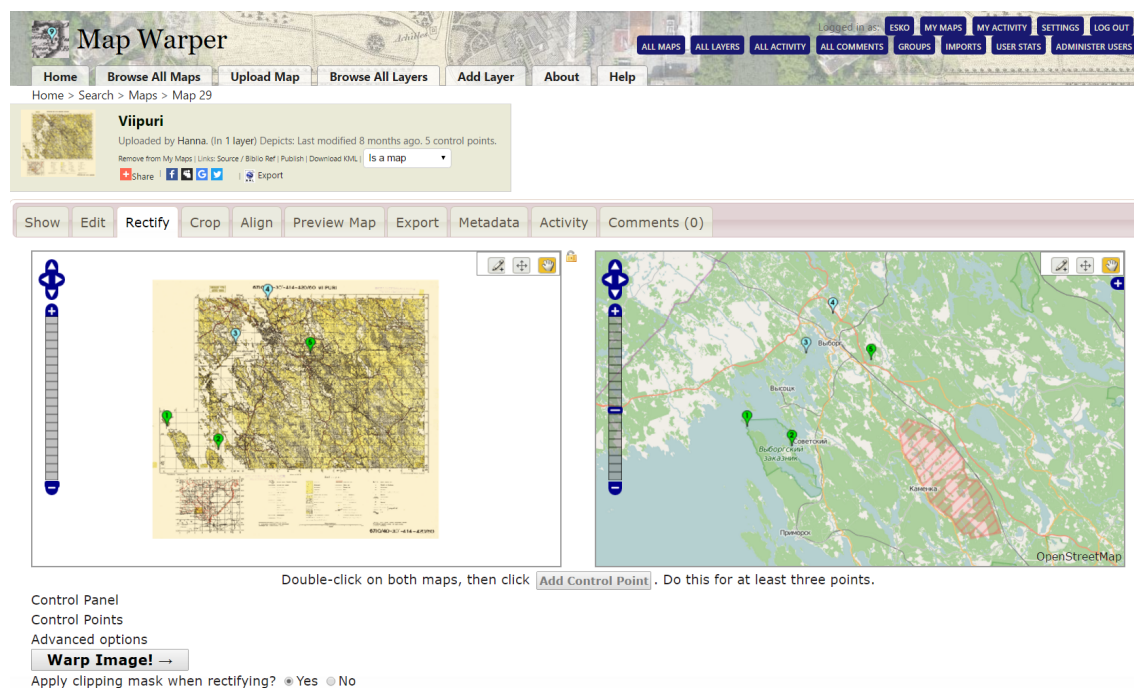
Twitterin kehittämää typeahead.js-kirjastoa [87] käytettiin tekstihaun automaattiseen täydennyksen toteutukseen. Typeahead.js-kirjastosta löytyy tuki usealle samanaikaiselle datalähteelle, mikä osoittautui olennaiseksi Hipla.fi-käyttöliittymän kannalta. Näin samaan hakutulostilaan voitiin yhdistellä tuloksia useasta eri datalähteestä, ja datalähteiden päälle- ja poiskytkemisen vaikutukset hakutuloksiin voitiin päivittää dynaamisesti.

Karttanäkymän nykyhetkeä kuvaavaan karttaan liittyvä toiminnallisuus tehtiin Google Maps JavaScript API:n [27] avulla. Googlen kartta valittiin karttanäkymän pohjakartaksi, sillä tämä mahdollisti aiemmassa kulttuurireitteihin liittyvässä tutkimuksessa [39] valmistuneiden Google Maps API:in perustuvien käyttöliittymäkomponenttien hyödyntämisen. Pohjakarttana olisi ollut mahdollista käyttää myös esimerkiksi OpenStreetMap-karttaa, koska Googlen pohjakartan tiilet on mahdollista vaihtaa OpenStreetMap-tiiliin [74]. Tätä toiminnallisuutta ei kuitenkaan toteutettu käyttöliittymän ensimmäiseen versioon.

## Historiallisten karttalehtien palvelun integrointi

Vanhojen skannattujen karttalehtien asemointia varten otettiin käyttöön New York Public Library:n käyttämä ja tukema avoimen lähdekoodin Map Warper -palvelu [53], joka julkaistiin osoitteessa <http://mapwarper.onki.fi>. Map Warper -palvelun avulla historiallisten karttalehtien georektifiointi on suoraviivainen toimenpide, eikä se vaadi erityistä asiantuntemusta. Palveluun ladataan ensin historiallisen karttalehden skannauksesta syntynyt kuvatiedosto. Seuraavaksi etsitään historialliselta karttalehdeltä ja nykyiseltä kartalta samoja kohteita, esimerkiksi paikallaan pysyneitä rakennuksia, teitä tai rantaviivan muotoja, minkä jälkeen samat kohteet merkitään tukipisteiksi. Kuvassa 13 on esitetty vasemmalla puolella skannattu ja Map Warper -palveluun ladattu Viipurin alueen historiallinen karttalehti ja oikealla puolella nykyinen kartta sekä kummaltakin kartalta poimitut tukipisteet. Tukipisteitä tarvitaan vähintään kolme kappaletta, mutta georektifioinnin tulos parantuu, jos niitä valitaan enemmän. Tukipisteiden määrittelyn jälkeen Map Warper suorittaa kartan

georektifoinnin eli skaalaa ja venyttää historiallista karttalehteä, jotta siitä tulee yhteensopiva nykyisen koordinaattijärjestelmän kanssa.



Kuva 13: Viipurin aluetta kuvaavan historiallisen karttalehden georektifointi tukipisteiden avulla Map Warper -palvelussa

Kun georektifointi on suoritettu, karttalehti on saavilla Map Warper -palvelusta eri formaateissa, joita ovat muun muassa PNG (Portable Network Graphics), Googlen alunperin kehittämä ja nykyisin OGC:n standardoima KML (Keyhole Markup Language) [69], WMS (Web Map Service) [71] sekä Googlen kartan kanssa yhteensopivat karttatiilet (engl. tiled web map). Karttatiilien ideana on, että kartta muodostetaan useista, esimerkiksi 256x256 pikselin kokoisista karttatiilistä, jotka asetetaan saumattomasti vierekkäin, kun taas muut edellä mainitut formaatit esittävät kartan yhtenä kuvana. Historiallisten karttalehtien esittäminen modernilla karttapohjalla, jota käyttäjä saa vapaasti zoomata ja liikuttaa, asettaa merkittäviä suorituskykyvaatimuksia historiallisten karttalehtien lataukselle. Tähän ratkaisuna ovat karttatiilet, koska muita yllä mainittuja formaatteja käytettäessä koko historiallinen karttalehti täytyisi ladata uudelleen joka kerta, kun käyttäjä zoomaa tai liikuttaa Hipla.fi-palvelun karttanäkymää. Karttatiiliä käytettäessä vain muuttuneet karttatiilet täytyy ladata uudelleen karttanäkymää päivitettäessä, ja lisäksi osa karttatiilistä voidaan ladata välimuistista. Huomattavasti parantuneen suorituskyvyn vuoksi suurimmat kaupalliset verkkopohjaisten karttapalvelujen tarjoajat, kuten esimerkiksi Google, Yahoo ja Microsoft ovat ottaneet karttatiilimenetelmän käyttöön. Karttatiilien spesifikaatiosta ei kuitenkaan ole olemassa yleistä standardia, vaan suuryritykset ovat päätyneet käyttämään osin omia määrittämiään, mikä vaikeuttaa karttojen hyödyntämistä avoimissa karttapalveluissa [23].

Hipla.fi-käyttöliittymään tietyn alueen historialliset karttalehdet haetaan Map

Warper -palvelun tarjoaman hakurajapinnan avulla. Rajapintakutsulle annetaan parametriksi suorakulmion muotoinen alue, ja vastauksesta saadaan kyseistä aluetta leikkaavien historiallisten karttalehtien tunnisteet. Tunnisteiden avulla historiallisten karttalehtien karttatiilet voidaan piirtää Google-kartan päälle Map Warper -palvelun karttatiilirajapinnan avulla. Karttatiiliä käytettäessä useita kymmeniä historiallisia karttalehtiä voidaan näyttää Googlen kartan päällä, ilman että kartan käyttö hidastuu merkittäväällä tavalla.

### Datapalvelu ja integrointi ulkoisiin luettelointijärjestelmiin.

Hipla.fi-datapalvelu koostuu käyttäjien paikkaehdotuksien ja historiallisten karttalehtien metadatan julkaisusta SPARQL-palvelupisteessä, paikkainstanssien URI:en resolvoinnista sekä tekstihakuleijukkeesta, jonka avulla Hipla.fi:n välittämä data voidaan ottaa tehokkaasti käyttöön ulkoisissa luettelointijärjestelmissä. Näin ollen Hipla.fi-palvelu ei ainakaan ensimmäisessä prototyypivaiheessa itse julkaise varsinaisia paikkaontologioita, vaan on ensisijaisesti paikkatiedon välityspalvelu.

Resolvoinnin yleinen idea on, että HTTP-standardissa määritellyn *content negotiation* -mekanismin avulla järjestelmä voi palauttaa saman URI:n perusteella erilaisen vastauksen riippuen kyselyn tehneestä toimijasta. Esimerkiksi selaimelle voidaan palauttaa HTML-dokumentti ja RDF-lukijalle RDF-tiedosto.

Välityspalveluna toimiva Hipla.fi ei toteuta paikkatietolähteiden linkitetyn datan paradigman mukaista URI:en resolvointia, vaan kunkin paikkatietolähteen julkaisija hoitaa tämän toiminnallisuuden itse. Sen sijaan Hipla.fi esittää usean eri paikkatietolähteen sisältöä samalla kartalla HTML-sivuiksi renderöitynä. Esimerkiksi merkkijonohaku "helsinki" tuottaa sivun <http://hipla.fi/?query=helsinki>, jolla esitetään kaikkien käyttäjän valitsemien paikkatietolähteiden tulokset kyseisellä merkkijonolla. Koska Helsinki löytyy useista eri paikkatietolähteistä, hakutulokset valitsemalla päästään taulukossa 1 listatuille eri paikkatietolähteiden Helsinki-instanssien HTML-sivuille Hipla.fi-palvelussa. Näin Hipla.fi harmonisoi eri lähteistä peräisin olevaa paikkatietoa yhdelle karttapohjalle vertailtavaksi ja paikanimien yksikäsitteistämistä varten säilyttäen silti alkuperäisten paikkatietolähteiden ominaispiirteet.

Hipla.fi-datapalvelu on toteutettu Semanttisen laskennan tutkimusryhmän kehittämän ja ylläpitämän, linkitetyn datan teknologioihin perustuvan Linked Data Finland (LDF.fi) -julkaisualustan [36] avulla. LDF.fi-julkaisualusta toteuttaa Hipla.fi-palvelun käyttäjien paikkaehdotuksien ja karttalehtien metadatan julkaisun ja URI:en resolvoinnin. RDF-tietokantana toimii Apache Jena Fuseki [83], josta kirjoitushetkellä käytössä on versio 2.4.0. Uudelleenohjauksia ja välimuistiin liittyviä toiminnallisuuksia varten käytössä on Varnish Cache [46].

Käyttäjien paikkaehdotukset sijoitettiin LDF.fi-alustan "demo"-nimiseen SPARQL-palvelupisteeseen, koska se on kirjoitushetkellä ainoa palvelupiste, johon on avattu pääsy tutkimusryhmän ulkopuolelta tuleville SPARQL Update -operaatioille, joita tarvitaan paikkaehdotuksien lisäämiseen. Tulevaisuudessa paikkaehdotuksille on tarkoitus kehittää oma mahdollisesti tunnistautumisen vaativa SPARQL-palvelupiste. Historiallisten karttalehtien metatiedot julkaistiin Hipla.fi-palvelun omassa SPARQL-

palvelupisteessä. Karttalehtien metatiedot on esitellään tarkemmin myöhemmin luvussa 6.1 ja kaikki Hipla.fi-palveluun kytketyt paikkatietolähteet luvussa 5.9.

Paikkatietolähde	Helsingin HTML-sivu Hipla.fi-palvelussa
Paikannimirekisteri	<a href="http://hipla.fi/?uri=http://ldf.fi/pnr/P_10890711">http://hipla.fi/?uri=http://ldf.fi/pnr/P_10890711</a>
The Getty Thesaurus of Geographic Names	<a href="http://hipla.fi/?uri=http://vocab.getty.edu/tgn/7009880">http://hipla.fi/?uri=http://vocab.getty.edu/tgn/7009880</a>
Suomen kunnat 1939–1944	<a href="http://hipla.fi/?uri=http://ldf.fi/warsa/places/municipalities/m_place_509">http://hipla.fi/?uri=http://ldf.fi/warsa/places/municipalities/m_place_509</a>
Kaikki paikkatietolähteet	<a href="http://hipla.fi/?query=helsinki">http://hipla.fi/?query=helsinki</a>

Taulukko 1: Hipla.fi-palvelun tarjoamat sivut eri paikkatietolähteiden “Helsinki”-instansseille

Ulkoisiin luettelointijärjestelmiin integrointia varten Hipla.fi-käyttöliittymästä erotettiin tekstihakuleijuke, joka julkaistiin avoimena lähdekoodina [77]. Tekstihakuleijuke voidaan lisätä web-pohjaiseen luettelointijärjestelmään, jossa se tarjoaa automaattisesti täydentyvät paikkaehdotukset. Kun käyttäjä on valinnut hakutuloksista haluamansa paikan, kyseisen paikan URI ja paikannimi lähetetään luettelointijärjestelmälle callback-funktion avulla. Lisäksi callback-funktion avulla luettelointijärjestelmälle välitetään tietoa käytetystä paikkatietolähteestä esimerkiksi SPARQL-palvelupisteen osoitteen avulla.

Leijukkeen lähdekoodia muokkaamalla sitä voidaan käyttää erilaisia käyttötapauksia varten. Esimerkiksi hakutulosriveillä näytettävää informaatiota voidaan muokata, tai hakutuloksiin voidaan lisätä pieni karttaruutu, joka näyttää kohteen sijainnin. Ajatuksena on myös, että tätä hakuleijuketta voidaan käyttää myös muiden kuin paikkaontologioiden käsitteiden valintaan ja yksikäsitteistämiseen. Esimerkiksi konfiguroimalla leijuke toimimaan eri lähteissä sijaitsevien toimijaontologioiden kanssa voidaan siirtyä annotoimaan aineistoja henkilöillä.

## 5.9 Datalähteet

Hipla.fi-palveluun on kytketty kirjoitushetkellä kahdeksan eri paikkatiedon lähdetä, jotka sisältävät yhteensä noin 3 000 000 paikkaa ja 460 karttaa. Paikkatietolähteiden perustiedot on esitetty taulukossa 2, ja seuraavasta listasta löytyy kullekin lähteelle lyhyt kuvaus ja tieto siitä, missä luvussa lähde on esitelty tarkemmin.

1. **Suomen kunnat 1939–1944.** Kansallisarkiston ”Suomi, sotavangit ja ihmisluovutukset 1939–55”-tutkimushankkeen karttasovelluksesta ladatut sodan ajan kunnat, joille tässä työssä lisättiin kuntien rajat polygoneina ja keskipisteet. Esitelty tarkemmin luvussa 6.1.

2. **Karjalan paikat 1922–1944.** Jyrki Tiittasen kokoama luovutetun Karjalan alueen paikannimien aineisto. Esitelty tarkemmin luvussa 6.1.
3. **Maanmittauslaitoksen paikannimirekisteri.** Nykyhetkellä käytössä oleva nimistö Suomessa käytetyillä kielillä. Esitelty tarkemmin luvussa 3.
4. **Suomen ajallinen paikkaontologia.** Paikkatiedon ajallista ulottuvuutta mallintava ontologia, joka kuvaa viimeisen noin 150 vuoden aikana tapahtuneet muutokset suomalaisissa kunnissa. Esitelty tarkemmin luvussa 3.
5. **The Getty Thesaurus of Geographic Names.** Tunnetuin kansainvälinen historiallisten paikkojen sanasto. Esitelty tarkemmin luvussa 3.
6. **Senaatin kartasto.** Venäjän armeijan topografisten joukkojen 1800-luvun lopulla ja 1900-luvun alussa mittakaavassa 1:21 000 laatimia Etelä-Suomen karttoja. Historiallisten karttojen georektifiointia on esitelty tarkemmin luvussa 5.8 ja karttojen metatietoja luvussa 6.1 .
7. **Karjalan kartat.** Maanmittaushallitus ja Topografikunta valmistivat neliväristä topografista karttaa 1:100 000 vuosina 1928–51. Historiallisten karttojen georektifiointia on esitelty tarkemmin luvussa 5.8 ja karttojen metatietoja luvussa 6.1.
8. **Paikkaehdotukset.** Hipla.fi-palvelun käyttäjien tekemät ehdotukset palveluun kytketyistä paikkatietolähteistä puuttuvista paikoista. Paikkojen yhteisöllistä ehdottamista on esitelty tarkemmin luvussa 5.7.

Nimi	Lähde	Paikkatyypit	Koko	SPARQL-palvelupisteen URL ja graafin URI
Suomen kunnat 1939–1944	Kansallisarkisto	kunta	612	<a href="http://ldf.fi/warsa/sparql">http://ldf.fi/warsa/sparql</a> <a href="http://ldf.fi/warsa/places/municipalities">http://ldf.fi/warsa/places/municipalities</a>
Karjalan paikat	Jyrki Tiittanen	kylä, maastokohde, jne.	34 938	<a href="http://ldf.fi/warsa/sparql">http://ldf.fi/warsa/sparql</a> <a href="http://ldf.fi/warsa/places/karelian_places">http://ldf.fi/warsa/places/karelian_places</a>
Suomen ajallinen paikkaontologia	Semanttisen laskennan tutkimusryhmä	kunta	1 261	<a href="http://ldf.fi/hipla/sparql">http://ldf.fi/hipla/sparql</a> <a href="http://www.yso.fi/onto/sapo">http://www.yso.fi/onto/sapo</a>
Paikannimirekisteri	Maanmittauslaitos	61 paikkatyyppiä	797 668	<a href="http://ldf.fi/pnr/sparql">http://ldf.fi/pnr/sparql</a> <a href="http://ldf.fi/pnr">http://ldf.fi/pnr</a>
The Getty Thesaurus of Geographic Names	Getty-säätiö	1800 paikkatyyppiä	2 156 896	<a href="http://vocab.getty.edu/sparql.json">http://vocab.getty.edu/sparql.json</a> <a href="http://vocab.getty.edu/dataset/tgn">http://vocab.getty.edu/dataset/tgn</a>
Senaatin kartasto	Kansallisarkisto	kartta	414	<a href="http://ldf.fi/hipla/sparql">http://ldf.fi/hipla/sparql</a> <a href="http://ldf.fi/hipla/maps">http://ldf.fi/hipla/maps</a>
Karjalan kartat	Maanmittauslaitos	kartta	47	<a href="http://ldf.fi/hipla/sparql">http://ldf.fi/hipla/sparql</a> <a href="http://ldf.fi/hipla/maps">http://ldf.fi/hipla/maps</a>
Paikkaehdotukset	Hipla.fi-käyttäjät			<a href="http://ldf.fi/demo/sparql">http://ldf.fi/demo/sparql</a> <a href="http://ldf.fi/demo/hipla-suggestions/">http://ldf.fi/demo/hipla-suggestions/</a>

Taulukko 2: Hipla.fi-palvelun datalähteet

## 6 Hipla.fi-palvelun pilotointi Sotasampo-projektissa

Tässä luvussa kuvataan Hipla.fi-palvelun ensimmäistä pilotointia osana Sotasampo-projektia. Aluksi esitellään lyhyesti Sotasampo-projektia sekä historiallisen paikkatiedon roolia Sotasampoon liittyvissä aineistoissa. Tämän jälkeen keskitytään siihen, kuinka Sotasampo varten koottiin eri lähteistä talvi- ja jatkosodan paikkaontologia, ja lopuksi esitellään Sotasampo-sovellukseen luotu paikkanäkymä.

Sotasampo-projektin [34] tavoitteena on avata Suomen toiseen maailmansotaan liittyvää tietoa datapalveluna sitä tarvitseville tutkijoille, historian harrastajille, sovellusten kehittäjille ja toisille verkkopalveluille. Ydinajatuksena on, että jo olemassa olevia aineistoja muunnetaan RDF-muotoon ja aineistojen välille luodaan linkityksiä automaattisesti. Aineistot julkaistaan linkitetyn datan periaatteiden mukaisesti, jolloin tuloksena syntyy yhden SPARQL-palvelupisteen kautta käytettävä datapalvelu. Lisäksi projektissa kehitetään Sotasammon datapalveluun perustuvaa semanttista portaalia (Sotasampo.fi), joka koostuu HTML5-sovelluksina toteutetuista erilaisista näkymistä Sotasammon aineistoihin.

Sotasampo on mielekäs pilotointikohde Hipla.fi-palvelulle, koska useimmat Sotasammon dataseteistä (esimerkiksi tapahtumat, valokuvat ja henkilöt) sisältävät tekstimuotoisia viittauksia paikannimiin. Näin ollen paikkatietoa voidaan hyödyntää aineistojen keskinäisessä automaattisessa linkityksessä ja aineistojen visualisoinnissa kartoille.

### 6.1 Talvi- ja jatkosodan paikkaontologia

Paikkatiedon hyödyntäminen Sotasammon aineistojen keskinäisessä linkityksessä ja aineistojen visualisoinnissa kartoilla luo tarpeen talvi- ja jatkosotaan liittyviä alueita kuvaavalle paikkaontologialle. Pelkästään nykyistä paikannimirekisteriä ei ole mahdollista käyttää, koska sotavuosien aikana Suomen ja Neuvostoliiton välillä tapahtui useita alueluovutuksia ja valtauksia. Esimerkiksi aineistoissa käytettyjä nykyisen Venäjän alueella sijaitsevia suomalaisia paikannimiä ei löydy enää nykyisestä paikannimirekisteristä. Näiden seikkojen ja valmiiksi saatavilla olevien aineistojen puutteellisuuden vuoksi oli välttämätöntä koota Sotasammon ja mahdollisesti muidenkin vastaavaa aihetta käsittelevien tahojen käyttöön talvi- ja jatkosodan paikkaontologia. Paikkaontologia koottiin seuraavista, sekä tässä luvussa että luvussa 5.9 esitellyistä paikkatietolähteistä: ”Suomen kunnat 1939–1944”, ”Karjalan paikat”, Maanmittauslaitoksen paikannimirekisteri, ”Senaatin kartasto” ja ”Karjalan kartat”.

## Tietomalli

Kun talvi- ja jatkosodan paikkaontologian lähdeaineistot olivat selvillä, Semanttisen laskennan tutkimusryhmässä kehitetystä historiallisen paikkatiedon mallintamiseen tarkoitettusta tietomallista, HIPLA-skeemasta, otettiin käyttöön joukko luokkia ja ominaisuuksia, joiden avulla lähdeaineistot muunnettiin RDF-muotoon. Taulukossa 3 on listattu HIPLA-skeemassa käytetyt nimiavaruudet.

Lyhenne	Nimiavaruus
rdf	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>
schema	<a href="http://schema.org/">http://schema.org/</a>
gs	<a href="http://www.opengis.net/ont/geosparql#">http://www.opengis.net/ont/geosparql#</a>
skos	<a href="http://www.w3.org/2004/02/skos/core#">http://www.w3.org/2004/02/skos/core#</a>
dc	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>
hipla	<a href="http://ldf.fi/schema/hipla/">http://ldf.fi/schema/hipla/</a>
wgs84	<a href="http://www.w3.org/2003/01/geo/wgs84_pos#">http://www.w3.org/2003/01/geo/wgs84_pos#</a>
xsd	<a href="http://www.w3.org/2001/XMLSchema#">http://www.w3.org/2001/XMLSchema#</a>

Taulukko 3: HIPLA-skeeman dokumentaatiossa käytetyt nimiavaruudet

Historiallisten paikkojen ja karttojen mallinnusta varten otettiin käyttöön neljä luokkaa:

1. paikka (luokka *hipla:Place*, esitelty taulukossa 4)
2. atlas (luokka *hipla:Atlas*, esitelty taulukossa 5)
3. karttasarja (luokka *hipla:MapSeries*, esitelty taulukossa 6)
4. karttalehti (luokka *hipla:Map*, esitelty taulukossa 7)

Paikkatiedon mallinnuksessa huomioitiin, että osalla paikoista on pistemäiset koordinaatit, osalla aluetta kuvaava polygoni ja joidenkin paikkojen koordinaatit puuttuvat toistaiseksi kokonaan, mutta ne ovat silti tarpeellisia aineistojen välisen linkityksen kannalta. Karttojen mallinnuksessa atlas on karttakokoelma, joka voi sisältää yhden tai useamman karttasarjan. Karttasarja taas on pohjimmiltaan kartta, joka on julkaistu usean karttalehden avulla. Tämän vuoksi samaan karttasarjaan kuuluvilla karttalehdillä on sama mittakaava ja kartografiset ominaisuudet.



Ominaisuus	RDF-tunniste	Kardinaliteetti	Arvojoukko
ensisijainen nimi	skos:prefLabel	1	literaali
paikkatyyppe	rdf:type	1	URI (valtio, lääni, kunta, kylä, rakennettu kohde, vesimuodostuma, maastokohde, kirkonkylä-kaupunki tai symboli)
pituusaste	wgs84:lat	0..1	xsd:float
leveysaste	wgs84:long	0..1	xsd:float
polygoni	schema:polygon	0..1	literaali
ensisijainen laajempi alue	gs:sfWithin	0..1	URI (valtio, lääni tai kunta)
laajempi alue	hipla:topology	0..n	URI (säilyneet alueet, luovutetut alueet tai valloitetut alueet)

Taulukko 4: hipla:Place-luokan ominaisuudet

Ominaisuus	RDF-tunniste	Kardinaliteetti	Arvojoukko
ensisijainen nimi	skos:prefLabel	1	literaali
kuvaus	dc:description	0..1	literaali
tekijä	dc:creator	0..n	literaali
julkaisija	dc:publisher	0..n	literaali
säilyttäjä	hipla:keeper	0..n	literaali

Taulukko 5: hipla:Atlas-luokan ominaisuudet

Ominaisuus	RDF-tunniste	Kardinaliteetti	Arvojoukko
ensisijainen nimi	skos:prefLabel	1	literaali
kuvaus	dc:description	0..1	literaali
julkaisupäiväys	dc:issued	0..1	xsd:date
mittakaava	hipla:scale	0..1	literaali
karttatyyppe	hipla:mapType	0..1	literaali
atlas	hipla:atlas	0..1	URI
karttalehtien lkm	hipla:mapCount	0..1	xsd:int

Taulukko 6: hipla:MapSeries-luokan ominaisuudet

Ominaisuus	RDF-tunniste	Kardinaliteetti	Arvojoukko
ensisijainen nimi	skos:prefLabel	1	literaali
tunniste	hipla:mapId	1	literaali
karttasarja	hipla:series	0..1	URI

Taulukko 7: hipla:Map-luokan ominaisuudet

## Suomen toisen maailmansodan ajan läänit ja kunnat

Toisen maailmansodan suomalaisen paikkaontologian rungoksi oli luonnollista valita läänit ja kunnat. Nykyistä kuntatietoa ei voitu kuitenkaan lukuisten kuntamuutoksien takia käyttää. Läänien ja kuntalistauksen lähteenä päätettiin käyttää Kansallisarkiston ”Suomi, sotavangit ja ihmisluovutukset 1939–55”-tutkimushankkeessa [47] tuotettua karttasovellusta, jossa on listattuna Suomen toisen maailmansodan aikaiset läänit ja kunnat lukuun ottamatta Ahvenanmaan lääniä.

Koska Kansallisarkiston karttasovelluksen dataa ei ollut saatavilla rakenteisessa muodossa, kuntien tiedot kopioitiin käsin karttasovelluksen HTML-sivuilta taulukoon. Kunnan tiedoksi saatiin nimi, karttasovelluksen tunniste (hyödyllinen, jos karttasovelluksen karttoja käytetään), lääni ja läänin voimassaoloaika. Lisäksi karttasovelluksen tietojen perusteella kunnat jaettiin kolmeen luokkaan:

- säilyneet alueet - kunnat, jotka ovat säilyneet Suomessa ainakin osittain
- luovutetut alueet - kunnat, jotka luovutettiin Neuvostoliitolle osittain tai kokonaan vuonna 1944
- valloitetut alueet - uudet kunnat, jotka Suomi valloitti Neuvostoliitolta tilapäisesti hyökkäysvaiheessa

Seuraavaksi CSV-muotoinen kuntataulukko muunnettiin Semanttisen laskennan tutkimusryhmässä kehitetyllä CSV2Whatever-työkalulla RDF-muotoon. Työkalu lukee CSV-muodossa olevaa taulukkoa rivi kerrallaan, ja muuntaa taulukon sarakkeiden arvot parametrina annettuun muotoon. Ihmisluottevuuden vuoksi RDF-syntaksiksi valittiin Turtle [8]. Kuntien mallintamiseen käytettiin HIPLA-skeeman luokkia ja ominaisuuksia. Lääni ilmaistaan *gs:sfWithin*-ominaisuudella (primaarinen kuuluvuussuhde), ja jako säilyneisiin, luovutettuihin ja valloitetuihin alueisiin kuvattiin *hipla:topology*-ominaisuudella.

Kansallisarkiston karttasovellus ei sisältänyt kuntien koordinaattitietoa. Koordinaattitiedon lähteeksi valittiin Suomen ajallinen paikkaontologia (SAPO) [38], joka sisältää aikasarjan Suomen kunnista noin 150 vuoden ajalta. Listaus toisen maailmansodan aikaisista kunnista oli jo saatu Kansallisarkiston karttasovelluksesta, joten SAPO:sta saatava tärkein tieto oli mainitun aikakauden kuntien digitoidut rajat eli polygonit. Polygonien hakua varten Kansallisarkiston karttasovelluksen kuntien nimet muunnettiin samaan muotoon kuin SAPO:ssa, esimerkiksi:

- *Viipurin mlk.* → *Viipurin mlk*
- *Viipuri, kaup.* → *Viipuri*

Tämän jälkeen voitiin ajaa SPARQL Update -operaatio, joka hakee SAPO:sta vuoden 1939 polygonit Kansallisarkiston karttasovelluksen kunnille (kuntia oli yhteensä 612, joista 506 kunnalle löytyi polygoni). Lopuksi kirjoitettiin Java-ohjelma, joka laski vielä keskipisteen kullekin polygonille ja lisäsi sen RDF-dataan. Kuvassa 14 on esitetty lopputuloksena saadun RDF-datajoukon Viipurin kuntaa mallintavan paikkainstanssin HIPLA-skeeman mukaiset tiedot SAHA-editorissa.

[http://ldf.fi/warsa/places/municipalities/m\\_place\\_614](http://ldf.fi/warsa/places/municipalities/m_place_614)

name  
message

kunta: Viipuri

[\[edit\]](#)

60.71045, 28.757046

<b>Arkistolaitoksen karttasovelluksen id</b>	vi_61
<b><a href="http://ldf.fi/schema/hipla/topology">http://ldf.fi/schema/hipla/topology</a></b>	<a href="http://ldf.fi/warsa/places/municipalities/m_topo_614">http://ldf.fi/warsa/places/municipalities/m_topo_614</a>
<b><a href="http://schema.org/polygon">http://schema.org/polygon</a></b>	28.7404911344953,60.6964696984546 28.7378128862901,60.7043583055741 28.7171339856755,60.7168499430259 28.7401633729244,60.7319920578853 28.7704945040565,60.7277875566405 28.7876049185302,60.7213658450487 28.7788380343287,60.7152877469702 28.782072465832,60.7084601098807 28.7876947347616,60.7013055588669 28.7764957554262,60.6946697918488 28.7563967382465,60.6902149856449 28.7404911344953,60.6964696984546
<b><a href="http://www.opengis.net/ont/geosparql#sfWithin">http://www.opengis.net/ont/geosparql#sfWithin</a></b>	Viipurin lääni
<b><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></b>	kunta
<b><a href="http://www.w3.org/2003/01/geo/wgs84_pos#lat">http://www.w3.org/2003/01/geo/wgs84_pos#lat</a></b>	60.71045
<b><a href="http://www.w3.org/2003/01/geo/wgs84_pos#long">http://www.w3.org/2003/01/geo/wgs84_pos#long</a></b>	28.757046
<b><a href="http://www.w3.org/2004/02/skos/core#prefLabel">http://www.w3.org/2004/02/skos/core#prefLabel</a></b>	Viipuri

Kuva 14: Viipurin kunnan paikkainstanssin tiedot SAHA-editorissa

## Karjalan karttanimet vuosilta 1922-44

Koska Sotasammon aineistoissa viitattiin usein Suomen itärajan seutujen paikannimiin, Jyrki Tiittasen kokoama luovutetun Karjalan alueen paikannimien aineisto osoittautui tärkeäksi lähteeksi koottaessa sodan ajan paikkaontologiaa. Tiittanen poimi luovutetun Karjalan 1:20 000 topografisilta karttalehdiltä yhteensä noin 40 000 paikannimeä ja lisäsi kullekin paikannimelle koordinaatit ja paikkatyypin.

1920–30-luvuilla valmistettujen 1:20 000 topografisien karttojen karttaruudukon koko on 1 kilometri, ja Tiittanen syötti KKKJ-koordinaattitiedon esitäytettyyn lomakkeeseen ilman digitointialustaa tai karttanäyttöä. Tämä luonnollisesti vähensi koordinaattien tarkkuutta, koska samassa 1 x 1 kilometrin kokoisessa karttaruudussa sijaitsevat paikannimet saivat samat koordinaatit. Joka tapauksessa kyseessä on lajissaan ainutlaatuinen aineisto. Paikkatyypiluokituksessa Tiittanen käytti kuutta eri tyyppiä, joiden jakauma on esitetty taulukossa 8. Lisäksi taulukosta 8 selviää, että paikannimiaineisto sisältää merkittävän määrän paikkoja, joilla on sama nimi ja paikkatyyppi, mutta eri koordinaatit.

Jotta Tiittasen taulukkomuotoinen aineisto saatiin käyttöön Hipla.fi-palvelussa, aineisto täytyi muuntaa RDF-muotoon. Muunnos tehtiin käyttäen CSV2Whatever-työkalua, ja paikkojen mallintamiseen käytettiin HIPLA-skeemaa. Paikkatyypeille luotiin omat luokat ja niille paikkatyypeille, joille löytyi vastaavuus KOKO-ontologiasta [19], tehtiin siltaus KOKO-ontologiaan.

Muun muassa automaattisen linkityksen vaatimusten vuoksi Tiittasen aineistosta haluttiin karsia pois mahdollisimman paljon samannimisiä paikkoja, joilla on lisäksi sama paikkatyyppi. Samannimisten paikkojen karsinta yritettiin ensin toteuttaa

Paikkatyyppi	Lukumäärä	Uniikkien paikannimien osuus
kirkonkylä, kaupunki	50	100 %
kylä	1544	72 %
maastokohde	10 864	64 %
vesimuodostuma	5553	50 %
rakennettu kohde	14 362	31 %
symboli	29	0 %

Taulukko 8: Luovutetun Karjalan karttanimien paikkatyytit ja uniikkien paikannimien osuus

SPARQL Update -operaatiolla käyttäen Jena Spatial -indeksin *nearby*-funktioita. Operaatiosta tuli kuitenkin liian raskas, joten päädyttiin kirjoittamaan Java-ohjelma, joka Apache Jena -kirjaston avulla käy pareittain läpi kaikki aineiston paikannimet ja poistaa aina toisen paikan, jos kummallakin paikalla on sama nimi, sama paikkatyyppi ja jos paikat ovat alle 10 kilometrin etäisyydellä toisistaan. Paikannimet eivät ole uniikkeja, joten aineistoon jäi edelleen runsaasti yli 10 kilometrin päässä toisistaan olevia paikkoja, joilla on sama nimi ja paikkatyyppi. Lopputuloksena alun perin 40 000 paikannimen aineistosta poistettiin yhteensä 7 000 paikannimeä, joiden katsottiin olevan käytetyn heuristiikan mukaan päällekkäisiä.

Samannimisten paikkojen karsinnan jälkeen kaikille Tiittasen aineiston paikoille lisättiin tieto siitä, missä kunnassa ne sijaitsevat. Jo olemassa olleen paikkatyyppin lisänä kuntatieto tarjoaa huomattavasti parantuneen mahdollisuuden samannimisten paikkojen yksikäsitteistämiseen. Kuntien polygonit olivat saatavilla “Suomen kunnat 1939–1944”-paikkatietolähteessä, mikä mahdollisti kuuluvuussuhteen lisäämisen. Operaatiossa syntyi tarve laskea onko annettu piste polygonin sisällä, mutta Jena Spatial -indeksin funktiokirjastosta ei löytynyt tähän soveltuvaa funktiota. Tämän vuoksi Semanttisen laskennan tutkimusryhmässä päädyttiin tekemään oma SPARQL-laajennus (*seco-spatial:withinPolygon*), joka kertoo onko piste polygonin sisällä.

### Senaatin kartasto ja Karjalan kartat

Senaatin kartastot on atlas, joka koostuu kahdesta erilaisesta karttasarjasta Venäjän armeijan topografisten joukkojen 1800-luvun lopulla ja 1900-luvun alussa laatimia Etelä-Suomen karttoja, joiden mittakaava on 1:21 000. Sarjojen nimet ovat Senaatin kartasto ja Venäläiset topografiset kartat 1:21 000. Tämän työn aineistoksi saatiin digitoitu Senaatin kartasto, joka sisältää nykyisin 414 karttalehteä. Alkuperäinen Senaatin kartasto kattoi myös talvi- ja jatkosodan kannalta olennaista Karjalan aluetta, mutta jatkosodan jälkeen uuden rajan taakse jäänyttä Karjalaa koskevat

kartat jouduttiin luovuttamaan Neuvostoliitolle. [48]

Karjalan aluetta kattavia karttoja on saatavilla Maanmittauslaitoksen Karjalan Kartat -verkkopalvelussa [60], joka sisältää Karjalan Kannakselta ja Laatokan Karjalasta Maanmittauslaitoksen ja Topografikunnan ennen vuotta 1939 laatimat kartat. Näistä digitoiduista kartta-aineistoista valittiin 1:100 000 topografinen karttasarja, joka sisältää 47 karttalehteä.

Senaatin kartasto ja Karjalan kartat, eli yhteensä 461 digitoitua karttalehteä, georektifioitiin Map Warper -palvelun avulla luvussa 5.8 kuvatulla tavalla. Georektifointi suoritettiin yhteistyössä Wikimedia Suomi ry:n “Avoin kulttuuridata hyötykäyttöön”-projektin [92] kanssa. Senaatin kartaston karttalehtien metatiedot saatiin taulukkomuodossa Kansallisarkistolta. Karjalan kartoille ei ollut saatavilla valmiita metatietoja, joten niitä varten tehtiin oma taulukko, johon tallennettiin karttalehden nimi ja tunniste. Lisäksi kaikkien karttalehtien metatietoihin lisättiin tunniste, jonka avulla karttalehden kuva voidaan ladata Map Warper -palvelusta eri formaateissa.

Jotta historiallisten karttojen metatiedot olisivat helposti kaikkien saatavilla, niitä ei tallennettu Map Warper -palveluun, vaan ne julkaistiin avoimena linkitettynä datana. Tämä tapahtui muuntamalla karttojen metatiedot CSV2Whatever-työkalun avulla HIPLA-skeeman mukaiseen RDF-muotoon, ja julkaisemalla metatiedot luvussa 5.9 kuvatulla tavalla. Kuvassa 15 on esitetty Senaatin kartasto -karttasarjan metatiedot SAHA-editorissa. SAHA-editori listaa lisäksi kaikki tähän karttasarjaan kuuluvat karttalehdet, joiden tietoihin voi siirtyä karttalehden nimen kautta.

[http://ldf.fi/hipla/maps/senate\\_maps/topographic\\_20000](http://ldf.fi/hipla/maps/senate_maps/topographic_20000)

name  
message

<http://ldf.fi/schema/hipla/Series: Senaatin kartasto>

[[edit](#)]

<a href="http://ldf.fi/schema/hipla/atlas">http://ldf.fi/schema/hipla/atlas</a>	<a href="#">Senaatin kartastot</a>
<a href="http://ldf.fi/schema/hipla/mapCount">http://ldf.fi/schema/hipla/mapCount</a>	414
<a href="http://ldf.fi/schema/hipla/mapType">http://ldf.fi/schema/hipla/mapType</a>	topografinen kartta, topographic map
<a href="http://ldf.fi/schema/hipla/scale">http://ldf.fi/schema/hipla/scale</a>	1:20 000
<a href="http://purl.org/dc/elements/1.1/description">http://purl.org/dc/elements/1.1/description</a>	Senaatin kartasto on toinen Senaatin kartastojen sarjoista. Maanmittausylihallituksessa alun perin ollut sarja on sijoitettu osaksi Rauhankadun toimipisteessä säilytettävää maanmittauslaitoksen historiallista kartta-arkistoa Senaatin kartaston nimellä, ja se sisältää 414 karttalehteä.
<a href="http://purl.org/dc/elements/1.1/issued">http://purl.org/dc/elements/1.1/issued</a>	1870-1907
<a href="http://www.w3.org/2004/02/skos/core#prefLabel">http://www.w3.org/2004/02/skos/core#prefLabel</a>	Senaatin kartasto, Senate atlas
Tyyppi	<a href="http://ldf.fi/schema/hipla/Series">http://ldf.fi/schema/hipla/Series</a>

#### References (411)

[Jokioinen](#) (<http://ldf.fi/schema/hipla/Map>) ← <http://ldf.fi/schema/hipla/series>  
[Sippola](#) (<http://ldf.fi/schema/hipla/Map>) ← <http://ldf.fi/schema/hipla/series>  
[Sipoo](#) (<http://ldf.fi/schema/hipla/Map>) ← <http://ldf.fi/schema/hipla/series>  
[Perniö](#) (<http://ldf.fi/schema/hipla/Map>) ← <http://ldf.fi/schema/hipla/series>  
[Nousiainen](#) (<http://ldf.fi/schema/hipla/Map>) ← <http://ldf.fi/schema/hipla/series>

Kuva 15: Senaatin kartasto -karttasarjan metatiedot SAHA-editorissa

## 6.2 Sotasammon aineistojen linkittäminen talvi- ja jatkosodan paikkaontologiaan

Historiallisiin paikkoihin liittyvän kontekstitiedon tarjoaminen on yksi diplomityön keskeisistä tutkimushaasteista. Jotta kontekstitietoa olisi saatavilla, täytyy paikkatietoaineistoja linkittää muihin aineistoihin joko käsin tai automaattisesti. Hipla.fi-prototyyppi tarjoaa käyttöliittymän, joka helpottaa käsin tehtävää linkitystä ja annotointia. Kontekstitietoa sisältävien muiden aineistojen automaattinen linkitys paikkatietoaineistoihin vaatii puolestaan automaattisia tekstintunnistuksen menetelmiä, jotka löytävät aineistoissa esiintyviä tekstimuotoisia paikkaviittauksia.

Hipla.fi-palvelussa julkaistu talvi- ja jatkosodan paikkaontologia oli keskeisessä osassa Sotasampo-projektissa tehdyissä aineistojen keskinäisissä linkityksissä. Sotasammon eri aineistojen parissa työskennelleet tutkijat käyttivät automaattiseen linkitykseen automaattista annotointipalvelu ARPA:a [67]. Talvi- ja jatkosodan paikkaontologiaa kehitettiin yhteistyössä automaattisten linkitysten aikana, ja siihen lisättiin puuttuneita, mutta muissa aineistoissa esiintyneitä paikkoja. Koska automaattinen linkitys tuottaa vääjäämättä lukuisia virheellisiä linkkejä, Hipla.fi-käyttöliittymää käytettiin paikkojen yksikäsitteistämiseen ja manuaaliseen virheellisten linkkien korjailuun. Ensimmäisessä pilotointivaiheessa seuraavat aineistot linkitettiin talvi- ja jatkosodan paikkaontologiaan:

1. Talvi- ja jatkosodan tapahtumat: yhteensä 181 000 tapahtumaa sisältäen muun muassa sotatoimet, syntymät, haavoittumiset, kuolemat ja joukko-osastojen nimeämiset
2. Puolustusvoimien kuvakeskuksen SA-kuva-arkisto: 160 000 valokuvaa
3. Kansa Taisteli -lehtien artikkelit: 3400 Sotamuisto-yhdistyksen julkaisemaa talvisotaa ja jatkosotaa käsittelevää artikkelia

Linkityksien jälkeen historiallisiin paikkoihin liittyvää kontekstitietoa voitiin näyttää sekä Hipla.fi-käyttöliittymässä, että Sotasammon paikkanäkymässä.

## 6.3 Sotasammon paikkanäkymä

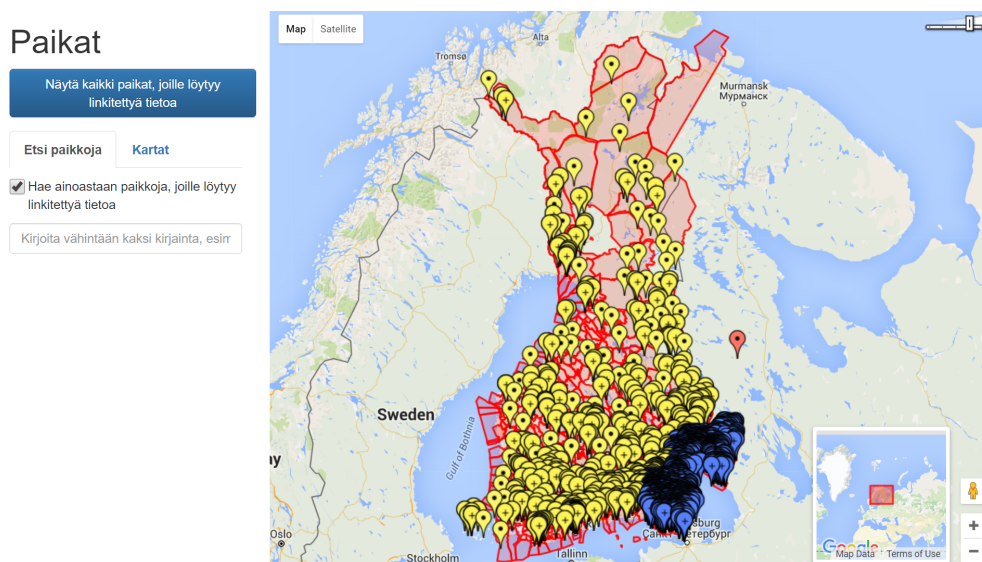
Sotasammon paikkanäkymän tarkoituksena on mahdollistaa talvi- ja jatkosodan paikkaontologian selaus- ja hakutoiminnot. Tämän lisäksi paikkanäkymä tarjoaa mahdollisuuden tutkia paikkaontologiaan linkitettyjä muita aineistoja, kuten esimerkiksi valokuvia tai tapahtumia sekä modernilla että historiallisella kartalla.

Paikkanäkymä toteutettiin samanlaisena HTML5-sovelluksena kuin luvussa 5.1 kuvattu Hipla.fi-käyttöliittymä, mikä mahdollisti Hipla.fi-käyttöliittymän haku- ja karttakomponenttien hyödyntämisen. Kuvassa 16 on esitetty paikkanäkymän etusivu, jossa käyttäjälle näytetään paikkojen selailua varten Googlen karttapohjalla kaikki ne talvi- ja jatkosodan paikkaontologian paikat, joille löytyy linkityksiä kontekstitietoon, kuten esimerkiksi historiallisiin tapahtumiin tai valokuviin. Kirjoitushetkellä näitä paikkoja on yhteensä 2 360, ja jos Sotasammon dataan lisätään viittauksia paikkoihin,

joita ei ole aiemmin näytetty kartalla, tulevat nämä paikat automaattisesti näkyviin. Paikkatietolähteistä ”Suomen kunnat 1939–1944” on esitetty kartalla punaisina polygoneina, ”Karjalan paikat” sinisinä karttamerkkeinä ja Maanmittauslaitoksen paikannimirekisterin paikat keltaisina karttamerkkeinä. Kaikkia talvi- ja jatkosodan paikkaontologian paikkoja ei ole mielekästä näyttää kerralla, koska tällöin SPARQL-kyselyn tulosjoukosta tulee niin suuri, että Google Maps API ei kykene näyttämään paikkoja sujuvasti kartalla.

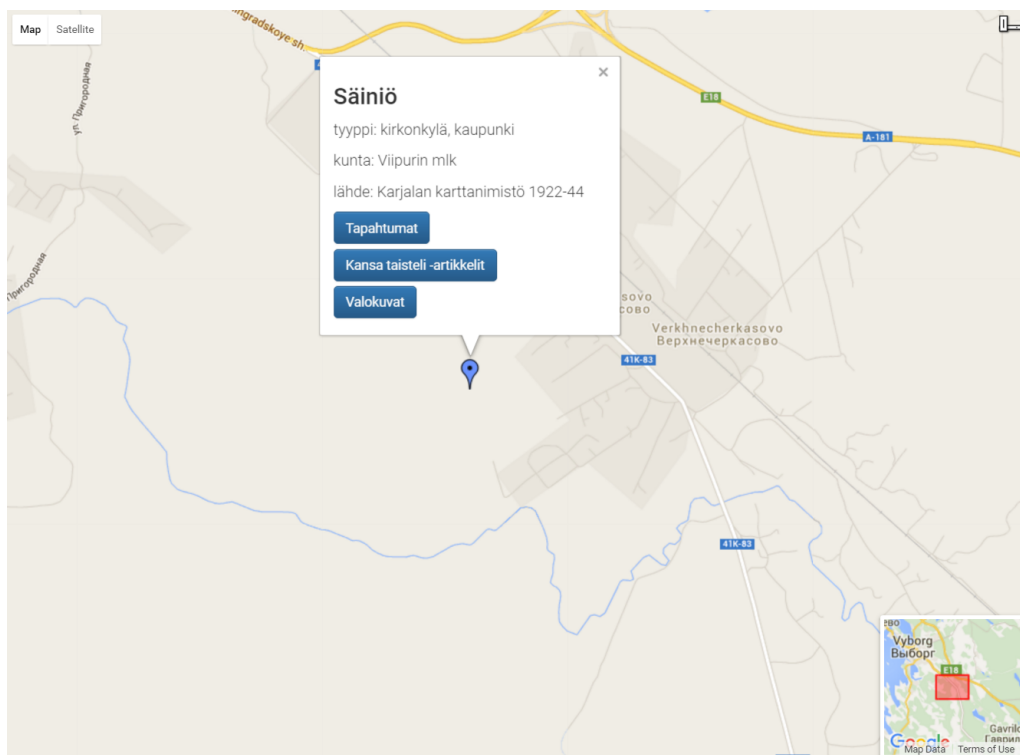
Kuvan 16 karttanäkymä on tarkoitettu paikkojen ja niihin liittyvän kontekstitiedon selailua varten. Lisäksi käyttöliittymän vasemmasta alalaidasta löytyy haku-toiminto, joka on toteutettu muuten samalla tavalla kuin Hipla.fi-käyttöliittymässä, mutta yksinkertaisuuden vuoksi käyttäjä ei voi valita käytettäviä paikkatietolähteitä, vaan lähteenä on aina talvi- ja jatkosodan paikkaontologia. Käyttäjä voi myös valita kohdistuuko haku ainoastaan paikkoihin, joille löytyy hakuhetkellä kontekstitietoa, vai kaikkiin paikkoihin. Tämä toiminto on mahdollinen, koska kontekstitieto haetaan dynaamisesti SPARQL-kyselyillä. Hakutulokset esitetään tekstilistauksena eriteltynä talvi- ja jatkosodan paikkaontologian kolmen osan mukaisesti, sekä visualisoituna kartalle. Käyttäjä voi valita paikan joko hakutulosrivin tai karttamerkin avulla.

Kuvassa 17 on valittu Viipurin maalaiskunnassa sijaitseva Säiniön kylä, jolloin karttamerkin yläpuolelle ilmestyy kontekstitietoa sisältävä infoikkuna, josta käyttäjä voi selata Säiniöön liittyviä tapahtumia, valokuvia ja Kansa Taisteli -lehtien artikkeleita. Lisäksi käyttäjä voi halutessaan siirtyä tutkimaan kyseistä kontekstitietoa Sotasammon muissa näkymissä. Paikat näytetään oletuksena Googlen karttapohjalla, ja kuten kuvasta 17 huomataan, Säiniön kylä sijaitsee nykyisen Venäjän alueella, mikä tarkoittaa sitä, että suomalaisia paikannimiä ole enää merkitty nykyiseen karttapohjaan. Käyttämällä paikkanäkymän ”kartat”-välilehteä päästään kuitenkin tutkimaan Säiniön kylää paremmin alkuperäisessä kontekstissaan, eli ”Karjalan kartat 1:100 000 topografinen”-sarjan karttalehdellä (kuva 18). Kuvien 17 ja 18 vertailu havainnollistaa georektifioitujen historiallisten karttalehtien tuomaa lisäarvoa.

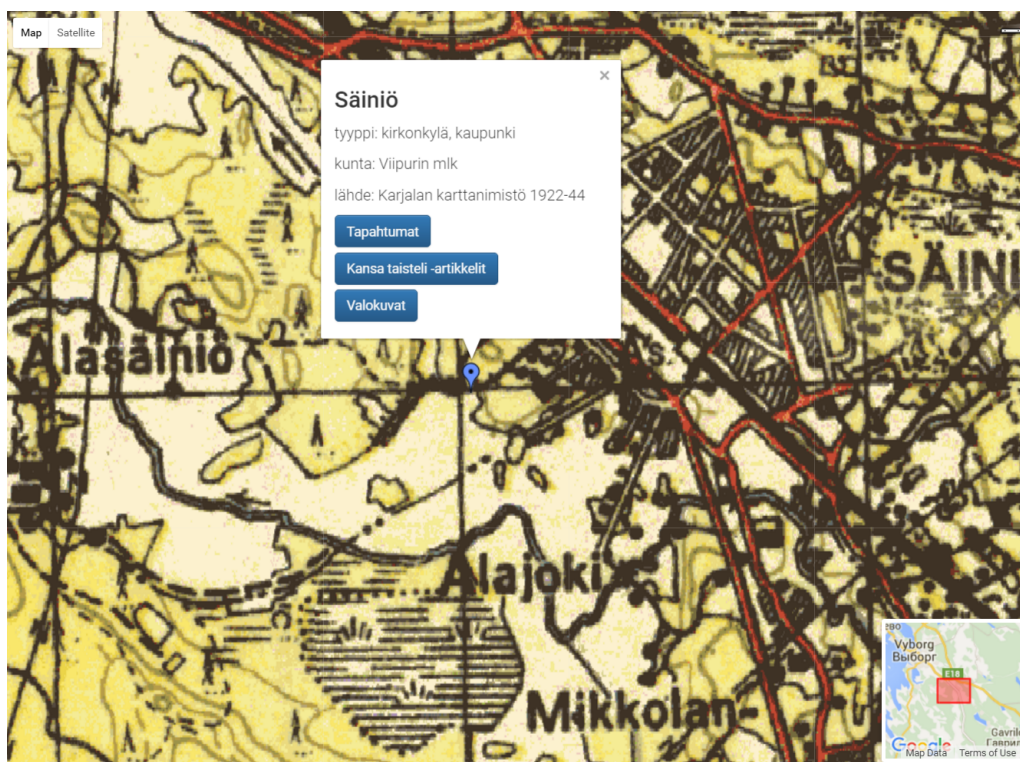


Kuva 16: Sotasammon paikkanäkymän etusivu





Kuva 17: Viipurin maalaiskunnan Säiniön kylä Googlen karttapohjalla



Kuva 18: Viipurin maalaiskunnan Säiniön kylä Karjalan 1:100 000 topografisella kartalla

## 7 Tulosten arviointi

Tässä luvussa arvioidaan miten historiallisten paikkojen ja karttojen palvelumalli HIPLA sekä sen prototyypitoteutus Hipla.fi vastaavat johdannossa esitettyihin tutkimuskysymyksiin. Samalla tehdään vertailua Hipla.fi-palvelun ja muiden historiallista paikkatietoa käsittelevien palveluiden välillä. Lopuksi esitetään yhteenveto siitä, miten eri historiallisen paikkatiedon palvelut vastaavat HIPLA-palvelumallin suunnittelun lähtökohtana käytettyihin vaatimuksiin.

### 7.1 Hakutyökaluja ja kontekstietoa georeferoinnin avuksi

Tutkimuskysymys 1: “Minkälaisia toiminnallisuuksia historiallisen aineiston tutkija tarvitsee, jotta hän voi löytää kohtaamalleen tekstimuotoiselle paikkaviittaukselle koordinaatit ja yksilöivän tunnusteen?” toimi HIPLA-palvelumallin sekä Hipla.fi-prototyypin käyttöliittymän suunnittelun lähtökohtana. Luvussa 4 esitellyt historiallisten paikkojen ja karttojen palvelun vaatimukset analysoitiin ja niiden perusteella muodostettiin joukko käyttötapauksia. Osa näistä käyttötapauksista otettiin mukaan Hipla.fi-käyttöliittymän ensimmäiseen versioon, joka tarjoaa muun muassa toimintoja koordinaattien ja yksilöivän tunnusteen hakuun. Tärkeimpiä hakutoimintoja ovat 1) federoitu automaattinen tekstihaku, 2) karttaperustainen paikkojen haku ja 3) paikkakäsitteillä tapahtuva semanttinen haku.

Lisäksi HIPLA-palvelumallin ajatuksena on tarjota kontekstietoa historiallisista paikoista aineistojen yhdistelemistä sekä tekstimuotoisten paikkaviittauksien yksikäsitteistämistä ja ymmärtämistä varten. Historiallisten nimistöluetteloiden tietoja ei ole tyypillisesti linkitetty paikkoihin liittyviin ulkoisiin tietolähteisiin. HIPLA-palvelumalli pyrkii linkitetyn datan teknologioiden avulla tuomaan eri lähteissä sijaisevien nimistöluetteloiden tietoja yhtenäiseen näkymään, jossa SPARQL-kyselyjä muokkaamalla voidaan näyttää paikkainstanssien infoikkunoissa paikkoihin mahdollisesti liittyvää kontekstietoa. Esimerkkinä kontekstiedosta on kuvassa 19 esitetty Sotasammon aineistojen paikkaviittausten lukumääriä. Linkkien määrä on huomattavasti suurempi kuin paikkojen määrä, mikä johtuu osin siitä että aineistoissa usein mainitut isot kaupungit saavat usein tuhansia linkityksiä.

### 7.2 Historiallisten karttojen tarve

Toista tutkimuskysymystä “Miten nykyisillä koordinaattijärjestelmillä esitettyyn paikkatietoon voidaan yhdistää historiallisia karttalehtiä?” varten HIPLA-palvelumalliin on integroitu ulkoinen historiallisten karttalehtien georektifointipalvelu. Hipla.fi-prototyypissä karttojen georektifointipalveluksi valittiin muun muassa New York Public Libraryn käytössä oleva avoimen lähdekoodin Map Warper -palvelu. Toinen suosittu historiallisten karttalehtien georektifointipalvelu on Georeferencer [51], joka asemoi karttoja samalla periaatteella kuin Map Warper. Erona on se, että Georeferencer-palvelun käyttö on maksullista ja sitä ei voi asentaa omalle palvelimelle.

Historiallisten karttalehtien hakuun on verkossa saatavilla useita palveluita, mutta

suuri osa näistä palveluista, esimerkiksi OldMapsOnline [52] ei tarjoa georektifioituja karttalehtiä eikä paikannimillä hakua. HIPLA-mallin mukainen palvelu tarjoaa historiallisten karttalehtien metadatan linkitettynä avoimena tietona ja integroi historiallisten karttalehtien palvelun nimistöpalveluun, mikä parantaa historiallisten karttalehtien löydettävyyttä huomattavasti.

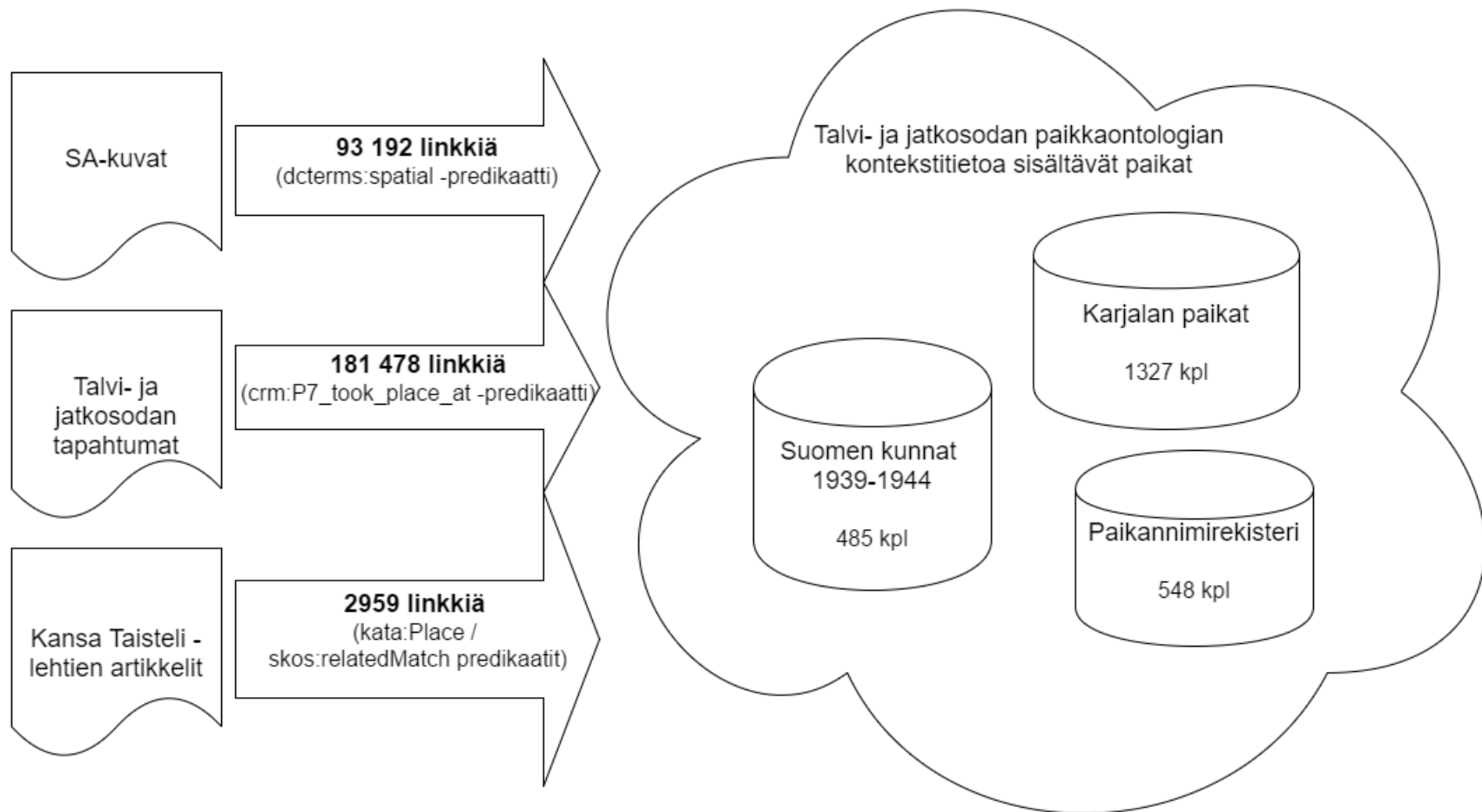
### 7.3 Paikkatietoaineiston hajautettu tuottaminen ja täydentäminen

Kolmas tutkimuskysymys “Miten paikkatietoaineiston jatkuva tuottaminen ja täydentäminen voidaan hajauttaa aineiston käyttäjille reaaliaikaisesti aineiston käyttötilanteessa?” liittyy HIPLA-malliin kuuluvaan kehittyvän ontologian hallinnointiprosessiin, joka kuvattu luvussa 4.6. Ideana on, että historiallisen paikkaontologian kehitystyö joukkoistetaan kulttuuriperintöaineistojen annotoijille, jolloin annotoinnin sivutuloksena syntyy uusia instansseja paikkaontologiaan. Toisin kuin esimerkiksi Pleaides-palvelussa [6], jossa joukkoistetaan ontologian kehitystyötä, HIPLA-mallin visiona on, että myös paikkaontologian käyttäjät tulee ottaa mukaan ontologian kehitystyöhön.

Koska historiallinen paikkatieto on tyypillisesti hajautunut useisiin erillisiin lähteisiin, HIPLA on suunniteltu pohjimmiltaan välityspalveluksi, joka kerää yhteen historiallista paikkatietoa monesta lähteestä. Linkitetyn datan teknologiat mahdollistavat HIPLA-malliin kuuluvan dynaamisen paikkatietolähteiden valinnan. Ehdotettu ratkaisu muistuttaa “normalisoituja ontologiarepositorioita” [89], jossa hajautettuja ontologiapalveluja normalisoidaan SKOS-perustaisten ohjelmointirajapintojen avulla yhdeksi laajemmaksi palveluksi. HIPLA:n tapauksessa normalisointiin käytetään SPARQL-kyselyjä, ja aihealueena on historiallinen paikkatieto. HIPLA:n välityspalvelumalli muistuttaa myös VIAF-mallia (The Virtual International Authority File) [11], jossa hajautettuja auktoriteettitietokantoja sillataan toisiinsa käyttäen *owl:sameAs*-predikaattia ja uusille instansseille annetaan uudet URI-tunnisteet uudessa nimiavaruudessa. HIPLA:n tapauksessa vastaavanlaisia siltauksia ei ylläpidetä ja olemassa oleville paikkatietoaineistoille ei luoda uusia URI-tunnisteita, vaan käyttäjän vastuulle jää päätös siitä, minkä paikkatiedon haltijan tunnisteita käytetään.

### 7.4 HIPLA-palvelumallin toiminnallisuudet muissa palveluissa

Taulukossa 9 on havainnollistettu sitä, miten luvussa 3 esitellyt palvelut sekä Hipla.fi-palvelu toteuttavat historialliseen paikkatietoon liittyviä toiminnallisuuksia. Linkitettyyn dataan ja HIPLA-palvelumalliin perustu Hipla.fi-palvelu sisältää kattavan yhdistelmän toiminnallisuuksia, joita ei löydy muista esitellyistä palveluista.



Kuva 19: Talvi- ja jatkosodan paikkaontologian linkitykset Sotasammon muihin aineistoihin

Palvelu	Hakutavat	Historialliset karttalehdet	Paikkojen kontekstieto	Joukkoistaminen	Ohjelmallinen käyttö
The Historical Gazetteer of England's Place-names	tekstihaku sekä historiallisella että modernilla kirjoitusasulla, ei karttapohjaista hakua	saatavilla, ilman metatietoja	ei saatavilla	ei saatavilla	ei saatavilla
Gazetteer for Scotland	tekstihaku ja karttahaiku	ei saatavilla	dataan kiinnitettyjä linkkejä	ei saatavilla	ei saatavilla
Getty TGN	tekstihaku	ei saatavilla	ei saatavilla	ei saatavilla	SPARQL-rajapinta
Pelagios / Pleiades	tekstihaku	neljä eri karttapohjaa	ei saatavilla	rekisteröityneillä käyttäjillä mahdollisuus ehdottaa täydennyksiä aineistoon	yksittäisen paikan tiedot voi hakea GeoJSON-formaatissa
Hipla.fi	tekstihaku, karttahaiku ja semanttinen haku	saatavilla, metatietojen kanssa	dynaamisesti päivittyvä kontekstieto	dynaamiset paikkaehdotukset	tekstihakuwidgetti ja SPARQL-rajapinta

Taulukko 9: Historiallisen paikkatiedon palveluiden vertailu

## 8 Yhteenveto ja jatkokehitys

Tässä työssä esitettiin kulttuurialan organisaatioiden tarpeeseen kehitetty uusi linkitetyn datan teknologioihin perustuva historiallisten paikkojen ja karttojen palvelumalli, HIPLA. Yksi HIPLA-mallin keskeisistä innovaatioista on yhteisöllinen toimintamalli, jolla historiallisia paikannimiä voidaan alkaa systemaattisesti keräämään erilaisilta historiallilta paikkatietoa hallinnoivilta organisaatioilta. HIPLA-malliin perustuvassa järjestelmässä on mahdollisuus ehdottaa ja rekisteröidä uusia paikannimiä heti, kun aineistojen luettelointityössä syntyy tarve uudelle nimistölle. Näin historiallisten paikkaontologioiden kehitys voidaan joukkoistaa varsinaisille paikkaontologian käyttäjille osana kulttuurilaitosten normaalia toimintaa, paikkaontologian hallinnan säilyessä edelleen siitä vastaavan organisaation hallussa. Lisäksi HIPLA-malli tarjoaa käyttöliittymän, jonka avulla eri organisaatioiden hajautettua historiallista ja nykyhetkeä koskevaa paikkatietoa voidaan hyödyntää yhtenäisessä näkymässä. Sijaintien tutkimiseen ja yksikäsitteistämiseen käyttöliittymä tarjoaa automaattisesti täydentyvän federoidun tekstihaun, karttahaun ja paikkakäsitteillä tapahtuvan semanttisen haun. Lisäksi käyttöliittymä näyttää paikkoihin liittyvää kontekstittietoa, kuten esimerkiksi historiallisia tapahtumia tai valokuvia. Tärkeänä lisäarvona käyttöliittymään on integroitu historiallisia karttalehtiä tarjoava palvelu, jonka avulla HIPLA-mallin mukaiseen palveluun kytkettyjä aineistoja on mahdollista tutkia sekä historiallisilla että nykyisillä karttalehdillä.

Työ toteutettiin soveltuvuusselvityksenä, jossa yleisen HIPLA-palvelumallin pohjalta kehitettiin ensimmäinen prototyyppitoteutus, avoimia SPARQL-palvelupisteitä ja Linked Data Finland -julkaisualustaa hyödyntävä HTML5-sovellus Hipla.fi. Tutkimusaineistoina käytettiin seuraavia historiallilta ja nykyistä paikkatietoa sisältäviä aineistoja: 1) Kansallisarkiston “Suomi, sotavangit ja ihmisluovutukset 1939–55”-tutkimushankkeen karttasovelluksesta ladatut sodan ajan kunnat ja läänit, 2) Suomen ajallinen paikkaontologia (SAPO), joka kuvaa historiallisten ja nykyisin olemassa olevien suomalaisten paikkojen ajallisen ja spatiaalisen ulottuvuuden 3) Jyrki Tiittasen kokoama luovutetun Karjalan alueen paikannimien aineisto, 4) Maanmittauslaitoksen paikannimirekisteri, 5) The Getty Thesaurus of Geographic Names, joka on yksi tunnetuimmista maailmanlaajuisista historiallisten paikkojen sanastoista, 6) 1900-luvun Etelä-Suomea kattavat Senaatin kartasto ja 7) Karjalan kartat 1930-50-luvuilta. Hipla.fi-käyttöliittymä julkaistiin avoimesti verkossa kokeiltavaksi, ja kaikki tutkimusaineistot joko julkaistiin avoimena linkitettyinä datana tai olivat jo aiemmin avoimesti saatavilla.

Hipla.fi-palvelua pilotoitiin Semanttisen laskennan tutkimusryhmän toteuttamassa talvi- ja jatkosodan tietoaineistoja julkaisevassa Sotasampo-projektissa, jota varten diplomityössä koottiin edellä mainittuja tutkimusaineistoja käyttäen talvi- ja jatkosodan paikkaontologia. Talvi- ja jatkosodan paikkaontologiaa käytettiin Sotasampo-projektin tutkijoiden toimesta eri aineistojen, kuten sodan ajan tapahtumien, toimijoiden, valokuvien ja artikkelien keskinäiseen automaattiseen linkitykseen. Lisäksi Sotasampo-portaalia varten kehitettiin Hipla.fi-prototyypin käyttöliittymäkomponentteja hyödyntäen Sotasammon paikkanäkymä, jonka avulla kaikkia Sotasammon paikkoihin linkittyneitä aineistoja voidaan selata ja hakea sekä nykyisillä

että historiallisilla kartoilla.

Työn tuloksena syntynyttä HIPLA-palvelumallia ja Hipla.fi-protyyppiä arvioitiin suhteessa muihin verkossa saatavilla oleviin historiallista paikkatietoa ja karttoja käsitteleviin ja julkaiseviin palveluihin. Arvioinnin tulokset osoittivat, että Hipla.fi-prototyypin sisältää kattavan yhdistelmän historiallisen paikkatiedon kannalta olennaisia toiminnallisuuksia, joita ei löydy muista tässä työssä esitellyistä palveluista.

Jatkossa Hipla.fi-protyyppitoteutukselle olisi tarpeen suorittaa käyttäjätestausta ja lisätä käyttöliittymään uusia toiminnallisuuksia liittyen muun muassa tekstihaun ja historiallisten karttalehtien haun tulosten järjestämiseen ja suodattamiseen sekä kehittyvän ontologian hallinnointiprosessiin. Lisäksi HIPLA-palvelumallia ja Hipla.fi-prototyypitoteutusta on tarkoitus kehittää yhdessä Kotuksen Nimiarkiston digitointihankkeen [54], Maanmittauslaitoksen koordinoiman JulkICT Lab -verkostohankkeen [45] sekä Oxfordin yliopiston johtaman COST-verkostohankkeen [14] kanssa.

## Viitteet

- [1] GeorSS in RDF, [http://www.georss.org/rdf\\_rss1.html](http://www.georss.org/rdf_rss1.html), viitattu 18.04.2016
- [2] NeoGeo Geometry Ontology, <http://geovocab.org/geometry.html>, viitattu 18.04.2016
- [3] Proposed GeoJSON-LD Vocabulary, <http://geojson.org/vocab>, viitattu 18.04.2016
- [4] Tieteen termipankki: toponyymi, <http://tieteentermipankki.fi/wiki/Nimitys:toponyymi>, viitattu 06.05.2016
- [5] Agafonkin, V.: Leaflet – an open-source JavaScript library for mobile-friendly interactive maps, <http://leafletjs.com/>, viitattu 03.06.2016
- [6] Ancient World Mapping Center, Stoa Consortium, Institute for the Study of the Ancient World: Pleiades gazetteer, <http://pleiades.stoa.org/>, viitattu 14.06.2016
- [7] Battle, R., Kolas, D.: Enabling the geospatial semantic web with parliament and GeoSPARQL. *Semantic Web* 3(4), 355–370 (2012)
- [8] Beckett, D., Berners-Lee, T., Prud'hommeaux, E., Carothers, G.: RDF 1.1 Turtle, Terse RDF Triple Language, W3C Recommendation 25 February 2014, <https://www.w3.org/TR/turtle/>, viitattu 05.07.2016
- [9] Berners-Lee, T.: Design issues: Linked data (2006), <http://www.w3.org/DesignIssues/LinkedData.html>, viitattu 16.06.2016
- [10] Bootstrap Core Team: Bootstrap, <http://getbootstrap.com/>, viitattu 14.04.2016
- [11] Bourdon, F., Boulet, V.: VIAF: A hub for a multilingual access to varied collections. In: *World library and information Congress: 78th IFLA general Conference and Assembly* (2011)
- [12] Brickley, D., Guha, R.: RDF Schema 1.1, W3C Recommendation 25 February 2014, <https://www.w3.org/TR/rdf-schema/>, viitattu 14.06.2016
- [13] Butler, H., Daly, M., Doyle, A., Gillies, S., Schaub, T., Schmidt, C.: The GeoJSON Format Specification, <http://geojson.org/geojson-spec.html>, viitattu 18.04.2016
- [14] COST Action IS1310: Reassembling the republic of letters, <http://www.republicofletters.net/>, viitattu 05.07.2016
- [15] Cuno, J.: Getty Thesaurus of Geographic Names Released as Linked Open Data, <http://blogs.getty.edu/iris/getty-thesaurus-of-geographic-names-released-as-linked-open-data/>, viitattu 20.04.2016



- [16] Cyganiak, R., Wood, D., Lanthaler, M.: RDF 1.1 Concepts and Abstract Syntax, W3C Recommendation 25 February 2014, <http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>, viitattu 13.04.2016
- [17] Digital Exposure of English Place-names (DEEP): The Historical Gazetteer of England's Place-names, <http://placenames.org.uk/>, viitattu 18.05.2016
- [18] European Commission Joint Research Centre: INSPIRE Metadata Implementing Rules: Technical Guidelines based on EN ISO 19115 and EN ISO 19119, [http://inspire.ec.europa.eu/documents/Metadata/MD\\_IR\\_and\\_ISO\\_20131029.pdf](http://inspire.ec.europa.eu/documents/Metadata/MD_IR_and_ISO_20131029.pdf), viitattu 11.05.2016
- [19] FinnONTO-konsortio, Kansalliskirjasto, Sanastokeskus TSK, Semanttisen laskennan tutkimusryhmä (SeCo): KOKO-ontologia, <https://finto.fi/koko/fi/>, viitattu 14.06.2016
- [20] Fotheringham, S., Wegener, M.: Spatial models and GIS: New and potential models. Taylor & Francis (1999)
- [21] Frain, B.: Responsive web design with HTML5 and CSS3. Packt Publishing Ltd (2012)
- [22] Gaede, V., Günther, O.: Multidimensional access methods. ACM Computing Surveys (CSUR) 30(2), 170–231 (1998)
- [23] García, R., de Castro Fernández, J.P., Riaza, E.V., Verdú, M.J., Regueras, L.M.: Web map tile services for spatial data infrastructures: Management and optimization. In: Bateira, C. (ed.) Cartography - A Tool for Spatial Analysis, pp. 26–48. Infotech (2012)
- [24] Gittings, B.M.: The Gazetteer for Scotland. <http://www.learningservices.is.ed.ac.uk/wordpress/wp-content/uploads/2015/06/gaz-for-teaching-wshop.pdf> (2008), viitattu 13.04.2016
- [25] Goodchild, M.F., Hill, L.L.: Introduction to digital gazetteer research. International Journal of Geographical Information Science 22(10), 1039–1044 (2008)
- [26] Google: AngularJS, <https://angularjs.org/>, viitattu 14.04.2016
- [27] Google Developers: Google Maps JavaScript API, <https://developers.google.com/maps/documentation/javascript/>, viitattu 13.04.2016
- [28] Gregory, I.N., Geddes, A.: Toward spatial humanities: Historical GIS and spatial history. Indiana University Press (2014)
- [29] Haslhofer, B., Sanderson, R., Simon, R., Van de Sompel, H.: Open annotations on multimedia web resources. Multimedia Tools and Applications 70(2), 847–867 (2014)

- [30] Heath, T., Bizer, C.: *Linked data: Evolving the web into a global data space*. Morgan & Claypool (2011)
- [31] Hill, L.L.: *Georeferencing: The geographic associations of information*. MIT Press (2009)
- [32] Hill, L.L., Zheng, Q.: Indirect geospatial referencing through place names in the digital library: Alexandria digital library experience with the developing and implementing gazetteers: Analysis and preliminary evaluation of the classical digital library model. In: *Proceedings of the Annual Meeting-American Society for Information Science*. vol. 36, pp. 57–69. Information Today (1999)
- [33] Hyvönen, E.: *Publishing and using cultural heritage linked data on the semantic web*. Morgan & Claypool (2012)
- [34] Hyvönen, E., Heino, E., Leskinen, P., Ikkala, E., Koho, M., Tamper, M., Tuominen, J., Mäkelä, E.: WarSampo data service and semantic portal for publishing linked open data about the second world war history. In: *The Semantic Web - Latest Advances and New Domains (ESWC 2016)*. pp. 758–773. Springer–Verlag (2016)
- [35] Hyvönen, E., Ikkala, E., Tuominen, J.: Linked data brokering service for historical places and maps. In: *Proceedings of the 1st Workshop on Humanities in the Semantic Web co-located with 13th ESWC Conference 2016*. vol. 1608, pp. 39–52. CEUR-WS Proceedings (2016)
- [36] Hyvönen, E., Tuominen, J., Alonen, M., Mäkelä, E.: Linked Data Finland: A 7-star model and platform for publishing and re-using linked datasets. In: *The Semantic Web: ESWC 2014 Satellite Events, Revised Selected Papers*. pp. 226–230. Springer–Verlag (2014)
- [37] Hyvönen, E., Tuominen, J., Ikkala, E., Mäkelä, E.: Ontology services based on crowdsourcing: Case national gazetteer of historical places. In: *Proceedings of the ISWC 2015 Posters & Demonstrations Track co-located with the 14th International Semantic Web Conference*. vol. 1486. CEUR-WS Proceedings (2015)
- [38] Hyvönen, E., Tuominen, J., Kauppinen, T., Väätäinen, J.: Representing and utilizing changing historical places as an ontology time series. In: Ashish, N., Sheth, A. (eds.) *Geospatial Semantics and Semantic Web: Foundations, Algorithms, and Applications*, pp. 1–25. Springer–Verlag (2011)
- [39] Ikkala, E., Mäkelä, E., Hyvönen, E.: TourRDF: Representing, enriching, and publishing curated tours based on linked data. In: *Knowledge Engineering and Knowledge Management: EKAW 2014 Satellite Events, VISUAL, EKM1, and ARCOE-Logic, Revised Selected Papers*. pp. 145–149. Springer–Verlag (2015)
- [40] Ikkala, E., Tuominen, J., Hyvönen, E.: Contextualizing historical places in a gazetteer by using historical maps and linked data. In: *Digital Humanities 2016, Conference Abstracts*. pp. 573–577 (2016)

- [41] International Organization for Standardization: ISO 19112:2003 Geographic Information-Spatial Referencing by Geographic Identifiers (2003)
- [42] Jones, C.B., Abdelmoty, A.I., Fu, G.: Maintaining ontologies for geographical information retrieval on the web. In: *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, pp. 934–951. Springer–Verlag (2003)
- [43] Jones, C.B., Purves, R.S.: Geographical information retrieval. *International Journal of Geographical Information Science* 22(3), 219–228 (2008)
- [44] jQuery Foundation: jQuery API, <http://api.jquery.com/>, viitattu 13.04.2016
- [45] JulkICT Lab: Paikannimipalvelua linkitettyinä tietona, <http://julkictlab.fi/fi/pilottiprojektit/paikannimipalvelua-linkitettyyna-tietona/>, viitattu 14.06.2016
- [46] Kamp, P.H.: Introduction to Varnish, <https://varnish-cache.org/intro/index.html#intro>, viitattu 03.06.2016
- [47] Kansallisarkisto: Suomi, sotavangit ja ihmislouvutukset 1939–55 , <http://kronos.narc.fi/>, viitattu 05.07.2016
- [48] Kansallisarkisto: Senaatin kartastot, [http://wiki.narc.fi/portti/index.php/Senaatin\\_kartastot](http://wiki.narc.fi/portti/index.php/Senaatin_kartastot), viitattu 07.07.2016
- [49] Kauppinen, T., Hyvönen, E.: Modeling and reasoning about changes in ontology time series. In: Sharman, R., Kishore, R., Ramesh, R. (eds.) *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*, pp. 319–338. Springer–Verlag (2007)
- [50] Kemp, K.: *Encyclopedia of Geographic Information Science*. SAGE Publications (2007)
- [51] Klokan Technologies GmbH: Georeferencer, <http://www.georeferencer.com/>, viitattu 03.06.2016
- [52] Klokan Technologies GmbH: Old Maps Online – Discovering the Cartography of the Past, <http://www.oldmapsonline.org/>, viitattu 14.06.2016
- [53] Knutzen, M.A.: Unbinding the atlas: Moving the NYPL map collection beyond digitization. *Journal of Map & Geography Libraries* 9(1-2), 8–24 (2013)
- [54] Kotimaisten kielten keskus: Nimiarkiston digitalisointihanke, <http://www.kotus.fi/nadigi>, viitattu 14.06.2016
- [55] Kraak, M.J., Ormeling, F.: *Cartography: visualization of spatial data*. Guilford Press (2011)

- [56] Kuhn, W., Kauppinen, T., Janowicz, K.: Linked data - a paradigm shift for geographic information science. In: *Geographic Information Science*, pp. 173–186. Springer–Verlag (2014)
- [57] Larson, R.R.: Geographic information retrieval and spatial browsing. In: *Geographic information systems and libraries: patrons, maps, and spatial information*. pp. 81–123. Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign (1996)
- [58] Lindroos, R.: Paikkatiedon ontologiapalvelu. Diplomityö. Teknillinen korkeakoulu (2008)
- [59] Linked Data Finland: Finnish Geographic Names, <http://www.ldf.fi/dataset/pnr/index.html>, viitattu 14.06.2016
- [60] Maanmittauslaitos: Karjalan kartat, <http://www.maanmittauslaitos.fi/aineistot-ja-palvelut/palvelut/karjalan-kartat>, viitattu 07.07.2016
- [61] Maanmittauslaitos: Kartta-aineistoja ja paikannimirekisteri avattiin vapaaseen käyttöön, <http://www.maanmittauslaitos.fi/tiedotteet/2011/01/kartta-aineistoja-paikannimirekisteri-avattiin-vapaaseen-kayttoon>, viitattu 03.06.2016
- [62] Mäkelä, E., Hyvönen, E.: SPARQL SAHA, a configurable linked data editor and browser as a service. In: *Proceedings of the ESWC 2014 demonstration track*. pp. 434–438. Springer–Verlag (2014)
- [63] Manning, C.D., Raghavan, P., Schütze, H.: *Introduction to information retrieval*. Cambridge University Press (2008)
- [64] McGuinness, D.L., van Harmelen, F.: *OWL Web Ontology Language, W3C Recommendation 10 February 2004*, <https://www.w3.org/TR/owl-features/>, viitattu 05.07.2016
- [65] Mikowski, M.S., Powell, J.C.: *Single page web applications*, <http://deals.manningpublications.com/spa.pdf>, viitattu 18.05.2016
- [66] Mozilla Developer Network: *Manipulating the browser history*, [https://developer.mozilla.org/en-US/docs/Web/API/History\\_API](https://developer.mozilla.org/en-US/docs/Web/API/History_API), viitattu 18.05.2016
- [67] Mäkelä, E.: Combining a REST lexical analysis web service with SPARQL for mashup semantic annotation from text. In: *The Semantic Web: ESWC 2014 Satellite Events, Revised Selected Papers*. pp. 424–428. Springer–Verlag (2014)
- [68] Old-Maps: *Old-Maps – Britain’s most comprehensive historical map archive*, <https://www.old-maps.co.uk/>, viitattu 14.06.2016
- [69] Open Geospatial Consortium: *KML*, <http://www.opengeospatial.org/standards/kml>, viitattu 03.06.2016

- [70] Open Geospatial Consortium: W3C and OGC to Collaborate to Integrate Spatial Data on the Web, <http://www.opengeospatial.org/pressroom/pressreleases/2149>, viitattu 18.04.2016
- [71] Open Geospatial Consortium: Web Map Service, <http://www.opengeospatial.org/standards/wms>, viitattu 03.06.2016
- [72] OpenLink Software: Virtuoso Geo Spatial Enhancements, <http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VirtGeoSPARQLEnhancementDocs>, viitattu 03.06.2016
- [73] OpenStreetMap Foundation: OpenStreetMap API, <http://wiki.openstreetmap.org/wiki/API>, viitattu 13.04.2016
- [74] OpenStreetMap Foundation: OpenStreetMap Google Maps example, [http://wiki.openstreetmap.org/wiki/Google\\_Maps\\_Example](http://wiki.openstreetmap.org/wiki/Google_Maps_Example), viitattu 13.04.2016
- [75] Sanastokeskus TSK: Geoinformatiikan sanasto (2014), <http://www.tsk.fi/tiedostot/pdf/GeoinformatiikanSanasto.pdf>, viitattu 16.6.2016
- [76] Scharl, A., Tochtermann, K.: The geospatial web: How geobrowsers, social software and the Web 2.0 are shaping the network society. Springer Science & Business Media (2009)
- [77] Semantic Computing Research Group: Federated SPARQL Search Widget, <https://github.com/SemanticComputing/federated-sparql-search-widget>, viitattu 14.06.2016
- [78] Simon, R., Barker, E., Isaksen, L., de Soto Cañamares, P.: Linking early geospatial documents, one place at a time: annotation of geographic documents with Recogito. *e-Perimtron* 10(2), 49–59 (2015)
- [79] Southall, H., Mostern, R., Berman, M.L.: On historical gazetteers. *International Journal of Humanities and Arts Computing* 5, 127–145 (2011)
- [80] Stuckenschmidt, H., van Harmelen, F.: Information sharing on the semantic web. Springer Science & Business Media (2005)
- [81] Svensson, P.: The landscape of digital humanities. *Digital Humanities* 4(1) (2010), <http://digitalhumanities.org/dhq/vol/4/1/000080/000080.html>
- [82] The J. Paul Getty Trust: Getty Thesaurus of Geographic Names® Online, <http://www.getty.edu/research/tools/vocabularies/tgn/>, viitattu 18.05.2016
- [83] The Apache Software Foundation: Apache Jena Fuseki, <https://jena.apache.org/documentation/fuseki2/>, viitattu 03.06.2016
- [84] The Apache Software Foundation: Spatial searches with SPARQL, <https://jena.apache.org/documentation/query/spatial-query.html>, viitattu 03.06.2016

- [85] Tuan, Y.F.: Place: An experiential perspective. *Geographical Review* 65(2), 151–165 (1975)
- [86] Tuominen, J., Frosterus, M., Viljanen, K., Hyvönen, E.: ONKI SKOS server for publishing and utilizing SKOS vocabularies and ontologies as services. In: *Proceedings of the 6th European Semantic Web Conference (ESWC 2009)*. pp. 768–780. Springer–Verlag (2009)
- [87] Twitter: typeahead.js, <https://github.com/twitter/typeahead.js>, viitattu 14.04.2016
- [88] Väättäin, J.: Suomen hallinnollisten alueiden muutosten kuvaaminen ja hyödyntäminen tiedonhaussa. Diplomityö. Aalto-yliopisto, Sähkötekniikan korkeakoulu (2016)
- [89] Viljanen, K., Tuominen, J., Mäkelä, E., Hyvönen, E.: Normalized access to ontology repositories. In: *Proceedings of the Sixth International Conference on Semantic Computing (IEEE ICSC 2012)*. pp. 109–116. IEEE Press (2012)
- [90] W3C SPARQL Working Group: SPARQL 1.1 Overview, W3C Recommendation 21 March 2013, <http://www.w3.org/TR/sparql11-overview/>, viitattu 05.07.2016
- [91] Waters, T.: Map Warper, <https://github.com/timwaters/mapwarper>, viitattu 03.06.2016
- [92] Wikimedia Suomi ry: Wikiprojekti Avoin kulttuuridata hyötykäyttöön, [https://fi.wikipedia.org/wiki/Wikiprojekti\\_Avoin\\_kulttuuridata\\_hy%C3%B6tyk%C3%A4ytt%C3%B6%C3%B6n](https://fi.wikipedia.org/wiki/Wikiprojekti_Avoin_kulttuuridata_hy%C3%B6tyk%C3%A4ytt%C3%B6%C3%B6n), viitattu 07.07.2016
- [93] World Wide Web Consortium: Basic Geo (WGS84 lat/long) Vocabulary, <https://www.w3.org/2003/01/geo/>, viitattu 18.04.2016
- [94] World Wide Web Consortium: Geo OWL, <https://www.w3.org/2005/Incubator/geo/XGR-geo-20071023/#owl>, viitattu 18.04.2016
- [95] Younis, E.M., Jones, C.B., Tanasescu, V., Abdelmoty, A.I.: Hybrid geo-spatial query methods on the semantic web with a spatially-enhanced index of DBpedia. In: Xiao, N., Kwan, M.P., Goodchild, M.F., Shekhar, S. (eds.) *Geographic Information Science*, vol. 7478, pp. 340–353. Springer Berlin Heidelberg (2012)