

Vektorit paikkatieto-PDF-kartalle GDAL-ohjelmalla

Jukka Rahkonen, <http://latuviitta.org>

Lisenssi [Creative Commons Attribution-Share Alike 3.0 Unported](http://creativecommons.org/licenses/by-sa/3.0/)

Acrobat Reader X:n käyttöliittymäkuvat: Copyright © 1984-2011 Adobe Systems Incorporated

Viimeksi muokattu 6. marraskuuta 2012

Tiivistelmä

GDAL-versiolla 2.0 voidaan lisätä PDF-tiedostoon rasterikartan päälle vektoriaiaineistoja mistä tahansa GDAL:in tukemasta vektoritiedolähteestä. PDF-tiedostoon voidaan kirjoittaa myös vektorikohteiden ominaisuustiedot ja alla oleva kuva näyttää lopputuloksen.

The screenshot shows the Adobe Reader interface displaying a PDF document. The main content is a map of a city area with several vector overlays: blue circles, green circles, and a yellow circle. The map includes labels for 'Hietalahti Sandviken', 'Munkkisaari', and 'Munkholmen'. The left sidebar shows a 'Mallipuu' (Table of Contents) with a list of items, including 'toilets', 'pubs', and 'parkkilippuautomaatit'. A table below the sidebar lists properties for a selected item, such as 'gnl_id', 'tunniste', 'osoite', 'vyöhyke', 'maksullisuus aika', 'max_parkkiaika', 'nimi', and 'OGR_STYLE'. The bottom status bar shows coordinates: Lon: 24 54,939 E, Lat: 60 09,554 N, E: 384,312,21 m, N: 6,670,970,17 m, and a scale of 4 233,3 x 4 233,3 mm.

Tällä kartalla 3 vektoritasoa: toilets (punainen), pubs (vihreä) ja parkkilippu-automaatit (sininen)

Tämä kohde valittu kartalta (keltainen rinkula)

Valitun kohteen ominaisuustiedot

Ominaisuus	Arvo
gnl_id	hki_parkkilippuautomaatit.8
tunniste	11
osoite	HIETALAHDENTORI 2
vyöhyke	2
maksullisuus aika	ma-pe 9-19 la 9-18
max_parkkiaika	4 h
nimi	HIETALAHDENTORI 2
OGR_STYLE	SYMBOL(id:ogr-sym-3,c:#0000FF,s:150px)

Aineistolähteet ja tekijänoikeudet:

- Maanmittauslaitos
- OpenStreetMap
- Helsingin kaupunki

Tiedosto on paikkatieto-PDF, joten se tuntee koordinaatit: sekä WGS-84 että kotimainen ETRS-TM35FIN

Tarvittavat ohjelmistot

GDAL-versiossa 2.0 on merkittäviä parannuksia paikkatieto-PDF-tukeen, ja tämän ohje perustuu sen käyttöön. GDAL 2.0 julkaistaan todennäköisesti vasta joskus alkuvuodesta 2013, joten siihen asti on käytettävä GDAL:in kehitysversiota, jonka saa Windows-koneille kätevimmin sivustolta <http://www.gisinternals.com/sdk/>

Lähtöaineistot

Kartalle tuleva rasterikuva voi olla missä tahansa GDAL:in tukemassa tiedostomuodossa

http://gdal.org/formats_list.html

ja kartalle lisättävät vektoriaineistot voivat olla missä tahansa GDAL/OGR-kirjaston tukemassa tiedostomuodossa

http://gdal.org/ogr/ogr_formats.html

Mallisuorituksen lähtöaineisto WFS-palvelusta – Miksi?

Tässä ohjeessa käytetyt vektoriaineistot ovat peräisin OpenStreetMap-projektista sekä Helsingin kaupungilta Helsinki Region Infoshare -sivuston kautta. Molemmat aineistot haetaan kuitenkin WFS-palveluista. Tähän on kaksi syytä: Ensiksikin se oli ohjeen kirjoittajalle helpointa, ja toiseksi ohjeen kirjoittaja on vahvasti sitä mieltä, että kaikki avoimet vektoritiedot täytyisi ensisijassa laittaa saataville WFS-palveluiden kautta. WFS:ssä on hankaluutensa, mutta jos maltetaan pysyä vanhoissa versioissa 1.0.0 ja 1.1.0, niin hankaluudet alkavat olla jo hallinnassa. Sitä paitsi, mitään WFS:ää yleisemmin tuettua standardia ole käytössä eikä edes näköpiirissä, vaan vaihtoehtona on pikemminkin ”Ikioma API joka aineistolle” -lähestymistapa. Tämä voi ohjelmointikykyisten palveluntuottajien mielestä olla ihan hyvä juttu, mutta latuviitta.org:in kohderyhmä käyttää valmiita ohjelmistoja eikä koodaa riviäkään, ja näille käyttäjille räätälöidyt API:t merkitsevät teknistä estettä tietojen saatavuudelle. Joten ensin WFS ja sitten API:t, jos saa pyytää. WFS on halpaa ja helppoa ja onnistuu ilmaisilla tai vaikka maksullisillakin valmisohjelmilla ilman rivinkään koodaamista.

Kuten edellisessä kappaleessa jo mainittiin, tämä ohje toki toimii pienellä soveltamisella kaikilla GDAL/OGR:n tukemilla tietolähteillä eikä pelkästään WFS-palveluiden kanssa. Päivitettyyn versioon tulee esimerkkejä muutamien muiden tietolähteiden käyttämisestä, mukaan lukien OSM-tietojen haku suoraan OSM:n omista rajapinnoista (XAPI, Overpass API).

Mallisuoritus ja miten se tehdään

Mallisuorituksessa tehdään seuraavaa:

- Otetaan yksi Maanmittauslaitoksen peruskarttalehti ja tehdään siitä paikkatieto-PDF, joka käyttää georeferointiin GDAL:in oletuksena olevaa ISO-menetelmää. Tätä aihetta käsittelee aikaisempi latuviitta.org:in ohje http://latuviitta.org/documents/Paikkatieto-PDF_ja_sen_kaksi_standardia.pdf sekä GDAL:in ohjesivu http://www.gdal.org/frmt_pdf.html

- Lisätään kartalle OpenStreetmapin Suomen aineistosta edellisen peruskarttalehden alueelta julkiset WC:t sekä pubit. Nämä tunnistetaan OSM:n avain-arvopareista ”amenity=toilets” ja ”amenity=pub”.
- Lisätään kartalle vielä Helsinki Region Infoshare:n välittämästä Helsingin kaupungin aineistosta parkkilippuautomaattien tiedot.
- Kehotetaan lukijoita huomaamaan tässä vaiheessa, että nämä aineistot ovat vain esimerkkejä, ja niiden sijasta tai niiden lisäksi voitaisiin käyttää mitä tahansa muita aineistoa. Aineistojen ei myöskään tarvitse olla pelkkiä pisteitä, vaan yhtä hyvin ne voisivat olla viivoja tai alueitakin.

Näin se tehdään

GDAL-ohjelma asettaa toistaiseksi tiettyjä rajoituksia paikkatieto-PDF:en tekemiseen. Ensiksikin GDAL:lla voidaan muuntaa vain yksi rasteriaineisto PDF-muotoon, joten valmiissa PDF:ssä voi olla vain yksi rasterikarttataso. Tämän päälle voidaan lisätä yksi vektoritiedolähde. Tämä ei kuitenkaan onneksi tarkoita sitä, että PDF-tiedostossa voisi olla vain yksi vektoritaso, sillä vektoritiedolähteenä voidaan käyttää GDAL:in virtuaali-OGR-tietolähdettä, johon voi kuulua rajaton määrä karttatasoja.

Vaihe 1. Rasterikartan valitseminen

Esimerkissä käytetään Maanmittauslaitoksen rasterikarttatiedostoa ”UL4133L_RK1_1.tif” jonka alueelle mahtuu suurin osa Helsingiä.

Vaihe 2. WFS-tasojen valitseminen

OpenStreetMap-aineistot haetaan latuviitta.org:in MapServer WFS -palvelusta. Tämä palvelu ei ole minkäänlaisessa valvonnassa, mutta sitä käytetään siitä syystä, että se sattuu tukemaan OSM-kohteiden valitsemista kaikkien aineistossa esiintyvien avain-arvoparien perusteella.

Helsingin kaupungin parkkilippuautomaatit haetaan latuviitta.org:in TinyOWS WFS -palvelusta, joka on paremmalla huolenpidolla ja jossa on myös paljon enemmän sisältöä (tällä hetkellä 102 WFS-tasoa, joista 100 kotimaisista aineistoista).

GDAL leikkaa vektoriaineistoista automaattisesti rasterikarttatason peittämän alueen, joten käyttäjän ei normaalisti tarvitse tehdä leikkausta itse. Automaattileikkaus ei kuitenkaan mahdollisesti toimi, jos rasterikuva ja vektoriaineistot käyttävät eri koordinaattijärjestelmiä.

Vaihe 3. OGR-virtuaalitiedoston kirjoittaminen

Luodaan tekstinkäsittelyohjelmalla seuraava tekstitiedosto ja tallennetaan se nimellä ”toilets_pubs_parkki.vrt”.

```

<OGRVRTDataSource>
  <OGRVRTLayer name="toilets">
    <SrcDataSource>WFS:http://hip.latuviitta.org/cgi-bin/ms_ows?
sql=(tags@>'amenity=>toilets')</SrcDataSource>
    <SrcSQL>SELECT *,osm_id as nimi, 'SYMBOL(id:ogr-sym-
3,c:#FF0000,s:125px)' as OGR_STYLE from osm_pisteet</SrcSQL>
  </OGRVRTLayer>
  <OGRVRTLayer name="pubs">
    <SrcDataSource>WFS:http://hip.latuviitta.org/cgi-bin/ms_ows?
sql=(tags@>'amenity=>pub')</SrcDataSource>
    <SrcSQL>SELECT *,osm_id as nimi, 'SYMBOL(id:ogr-sym-
3,c:#00ff00,s:225px)' as OGR_STYLE from osm_pisteet</SrcSQL>
  </OGRVRTLayer>
  <OGRVRTLayer name="hki_parkkilippuautomaatit">
    <SrcDataSource>WFS:http://hip.latuviitta.org/cgi-
bin/tinyows</SrcDataSource>
    <SrcSQL>SELECT *,osoite as nimi, 'SYMBOL(id:ogr-sym-
3,c:#0000FF,s:150px)' as OGR_STYLE from "lv:hki_parkkilippuautomaatit"</SrcSQL>
  </OGRVRTLayer>
</OGRVRTDataSource>

```

Huomautuksia .vrt-tiedostosta

- ensimmäinen WFS-palvelu on osoitteessa http://hip.latuviitta.org/cgi-bin/ms_ows
- tämä palvelu on määritelty niin, että käyttäjä voi tehdä valintoja OpenStreetMap:in avainarvoparien perusteella lisäämällä pyyntöihin sql-parametrin, esim.
sql=(tags@>'amenity=>toilets') tai sql=(tags@>'amenity=>pub')
- SQL-lauseella SELECT 'SYMBOL(id:ogr-sym-3,c:#FF0000,s:125px)' as OGR_STYLE määrätään PDF-kartalle piirrettävien kohteiden tyylit tämän määrittelyn mukaisesti: http://www.gdal.org/ogr/ogr_feature_style.html
Katso myös tyylien käytön rajoitukset sivulta http://www.gdal.org/ogr/drv_pdf.html
- toinen WFS-palvelu on osoitteessa <http://hip.latuviitta.org/cgi-bin/tinyows>
- pienenä mutta tärkeänä vihjeenä: tässä palvelussa on välttämätöntä sulkea valittavan WFS-tason nimi lainausmerkkeihin; se olisi hyvä tapa silloinkin, kun se ei ole pakollista "lv:hki_parkkilippuautomaatit"

Vaihe 4. Rastereiden ja vektoreiden yhdistäminen PDF-tiedostoksi

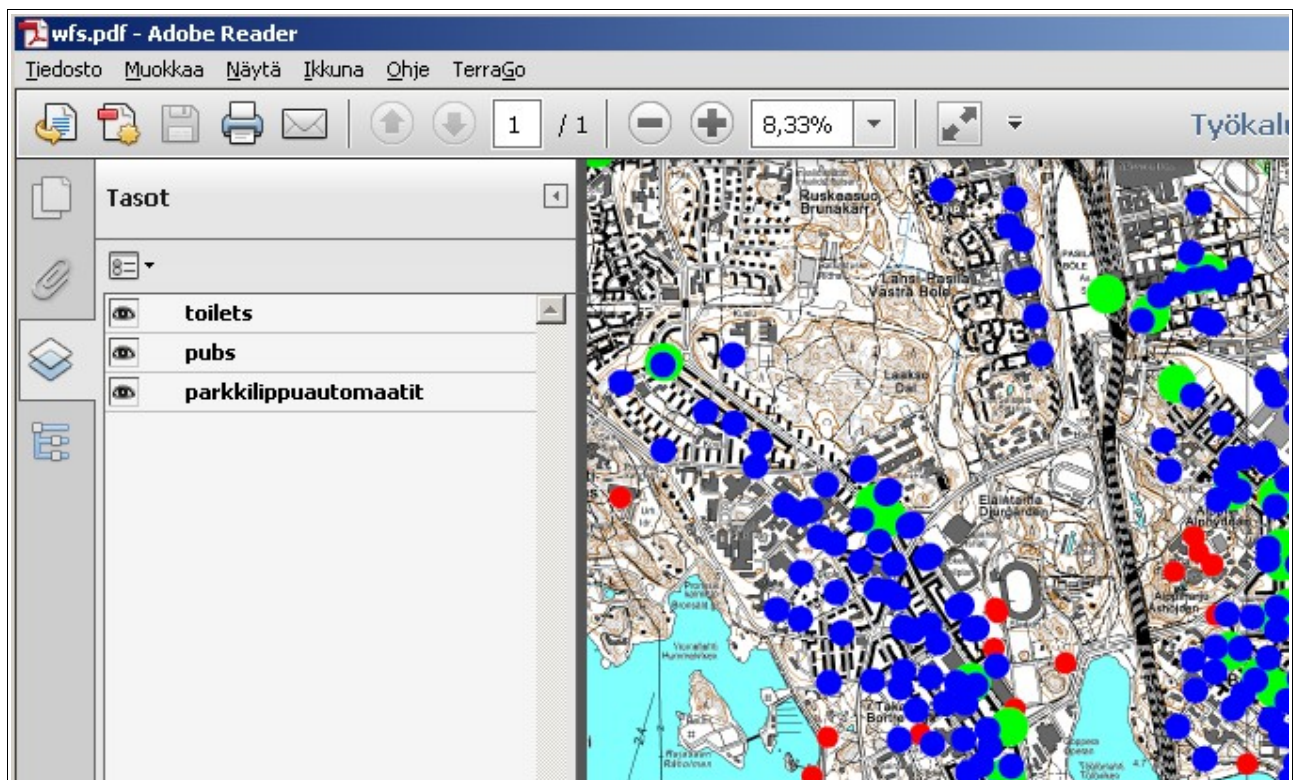
Lopuksi pitää enää suorittaa seuraava komento (jos käytät kopio-liitä-komentoja, niin muista poistaa tähän luettavuuden parantamiseksi lisätyt rivinvaihdot):

```
gdal_translate
-of PDF
-co OGR_DATASOURCE=toilets_pubs_parkki.vrt
UL4133L_RK1_1.tif GDAL_rasteri_ja_vektori.pdf
-a_srs epsg:3067
-co OGR_DISPLAY_LAYER_NAMES="toilets,pubs,parkkilippuautomaatit"
-co OGR_DISPLAY_FIELD="nimi"
```

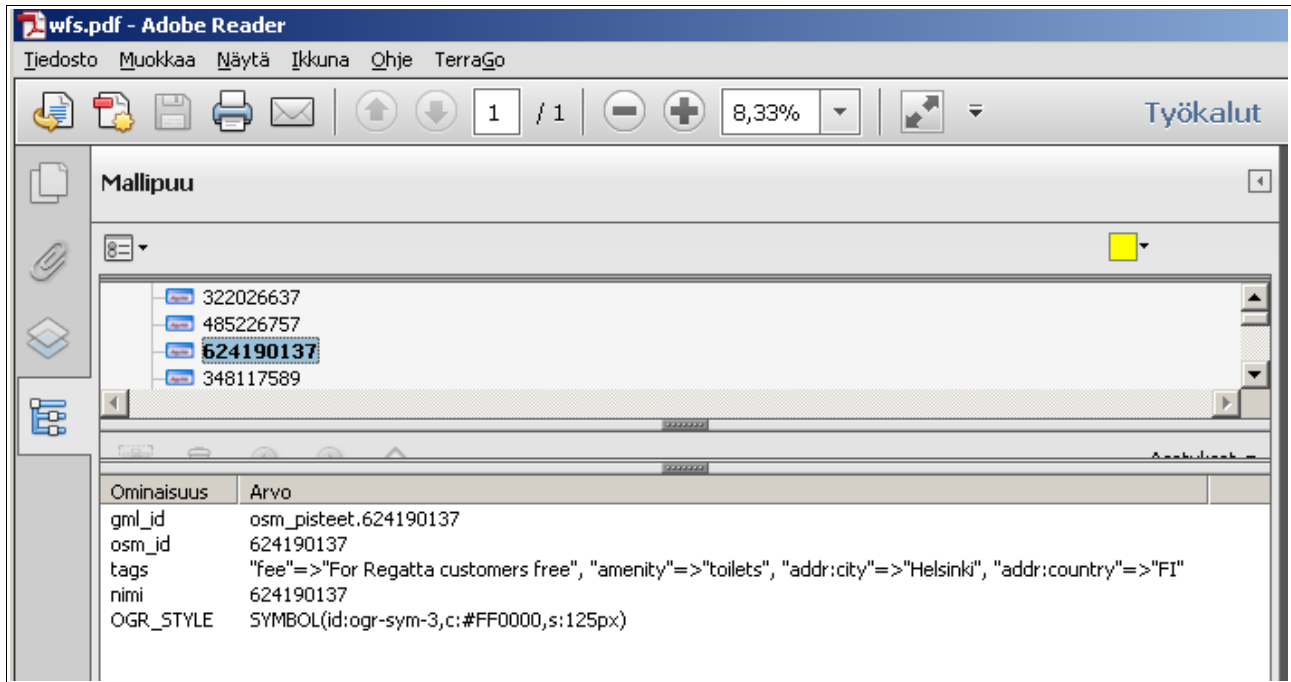
Huomautuksia GDAL-komennoista

- gdal_translate on ohjelma, joka tekee kaiken työn
- -of PDF määrää, että tulokseksi tulee PDF
- -co OGR_DATASOURCE=toilets_pubs_parkki.vrt valitsee vektoreiden lähteen; tämä voi siis olla mitä tahansa, mitä OGR tukee
- seuraavalla rivillä on muunnettavan rasterikuvan nimi ja tulostiedoston nimi
- -a_srs epsg:3067 liittää PDF:ään tiedon siitä, että koordinaattijärjestelmä on ETRS-TM35FIN
- -co OGR_DISPLAY_LAYER_NAMES="toilets,pubs,parkkilippuautomaatit" määrää, minkänimisinä OGR-tasot tulevat näkymään PDF-tiedoston tasojenhallinnassa
- -co OGR_DISPLAY_FIELD="nimi" valitsee, mistä vektoritiedoston ominaisuustietokentästä valitaan PDF-mallipuussa näkyvän kohdeluettelon otsikot

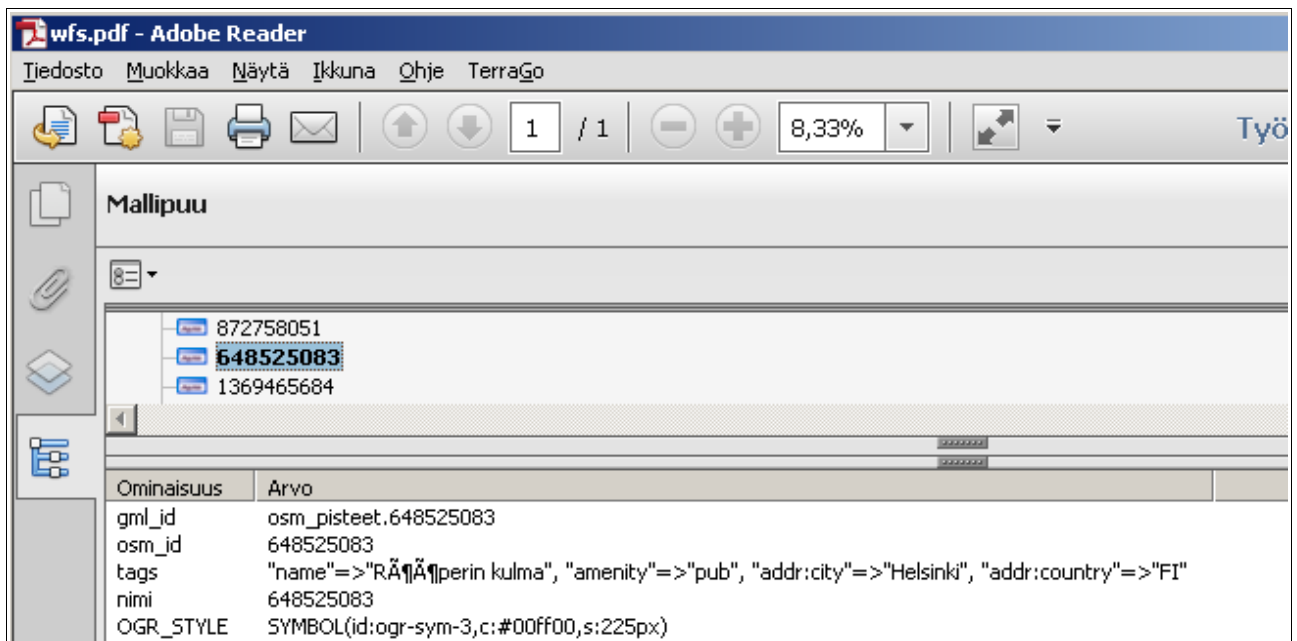
Lopputuloksen esittelyä kuvaruutukaappauksin



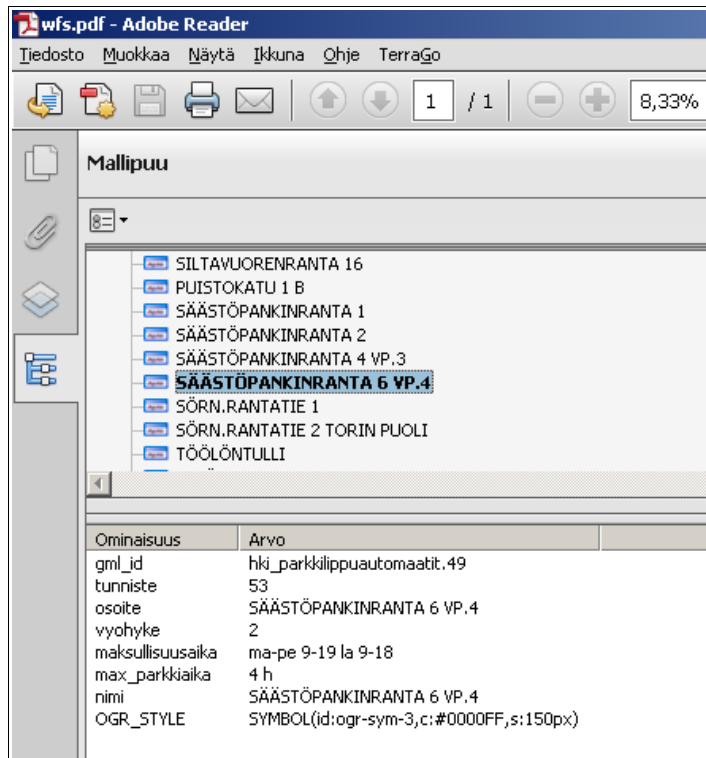
Tasojen nimet niin kuin ne haluttiin näyttää. Tasojen näkyvyyttä voidaan hallita Acrobat Readerissa.



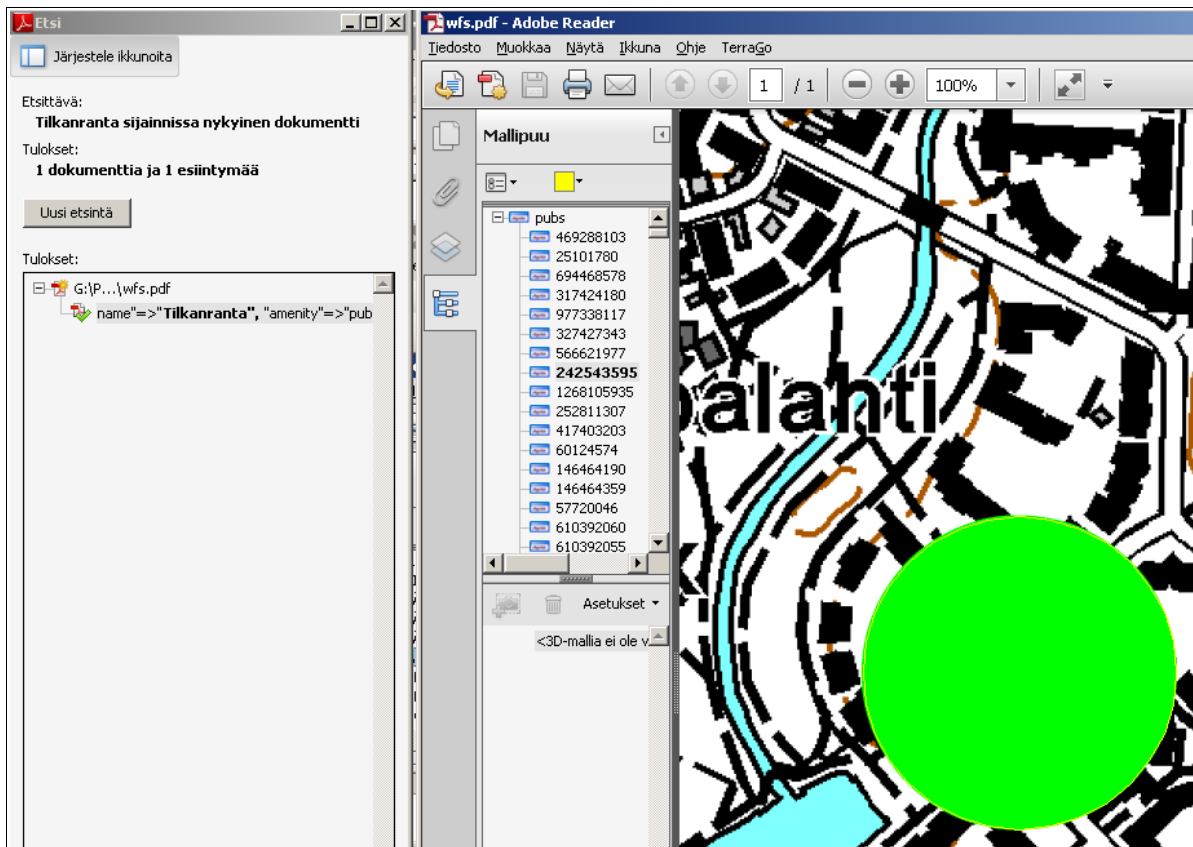
Koko OpenStreetMap:in rikas WC-tietosisältä on säilynyt, vaikka sitä ei esitetäkään kaikkein näteimmässä muodossa.



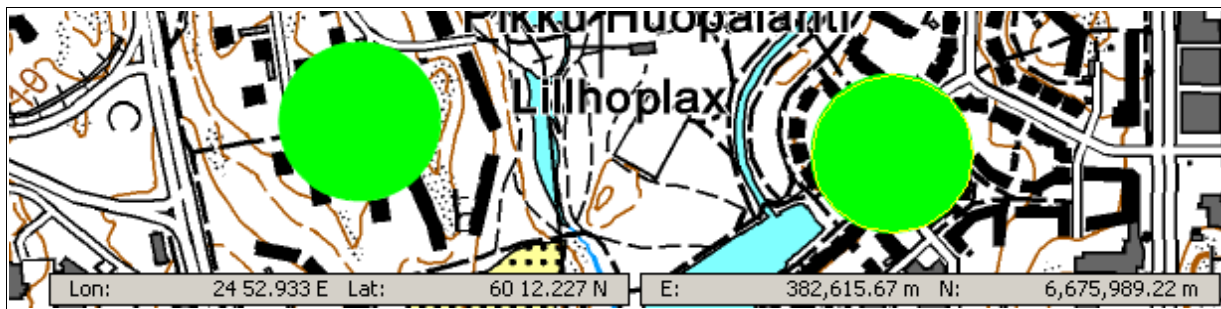
Tämä kuva kertoo paitsi ravintolan tiedot, myös sen, että GDAL:lla on hankaluuksia, jos WFS-palvelin ei käytä UTF-8-merkistökoodausta.



Sen sijaan UTF-8:a merkistökoodauksena käytävillä WFS-palvelimilla tekstit siirtyvät PDF-tiedostoon hienosti.



Vektoreiden ominaisuustiedoista voi tehdä hakuja Acrobat Readerin vakiotyökaluilla..



Se on paikkatieto-PDF ja se tuntee koordinaatit.

Nopeammin, paremmin ja halvemmalla

GDAL:in tuki paikkatieto-PDF:in luomiselle on vasta luonnosvaiheessa. Asiakirjaan voidaan toistaiseksi kirjoittaa (yksinkertaisesti) vain yksi rasteritaso, vektorikohteiden tyylien hallinta ei ole kovin monipuolista ja SQL:llä valitut tyyli määrittelyt näyttävät siirtyvän myös kohteiden ominaisuustietoihin. Muiden elementtien, kuten logojen, tekijänoikeusmerkintöjen ym. lisäämiseen valmiiseen PDF:ään on joitakin työkaluja olemassa, mutta ne ovat hankalia käyttää. Nykyisellään yhdistettyjen rasteri- ja vektori-PDF:ien tekeminen GDAL:lla sopisi paremmin jonkin palveluntuottajan kuin tavallisen kansalaisen tehtäväksi. Mallisuorituksesta kuitenkin voidaan päätellä, ettei tuollaisen tilauksesta tehtävän paikkatieto-PDF -palvelun tekeminen pitäisi mahdottoman vaikeaa olla. Lähes kaikkia tilausprosessin osiahan tässä on esitelty:

- Valitse alue (ei ole esitelty, mutta on hyvin yksinkertainen toimenpide)
- Valitse vektorihöyrysteet
- Valitse vektorihöyrysteiden piirtotyylit
- Valitse vektorihöyrysteiden ominaisuustiedot
- Valitse logo, otsikko, kompassinuolet ym. kartan koristeluun tarvittavat osat (tämä puuttuu).

Lopuksi vielä yksi käytännössä tehty huomio: Acrobat Reader käy aika hitaaksi, jos vektorikohteita on kymmeniä tuhansia. Muutama sata kohdetta vielä menettelee.