

Paikkatieto-PDF

Kaksi eri standardia teoriassa ja käytännössä

Jukka Rahkonen <http://latuviitta.org>

Viimeksi muokattu 8. lokakuuta 2012

Ihan tavallinen PDF, mutta vaan vähän enemmän

Paikkatieto-PDF eli geospatiaalinen PDF on PDF-muotoinen kartta, joka tietää, mistä päin maailmaa kartta on. Tämä tarkoittaa käytännössä sitä, että käyttäjä voi osoittaa mitä tahansa kohtaa kartalla ja PDF-katseluohjelma näyttää sen koordinaatit. Vastaavalla tavalla käyttäjä voi syöttää koordinaatit käsin ja PDF-lukuohjelma näyttää kartalta koordinaatteja vastaavan paikan

Yksinkertaistetusti voidaan sanoa, että paikkatieto-PDF:ssä tiedoston sisään tallennetaan kartan nurkkapisteiden koordinaatit sekä tieto siitä, minkä koordinaattijärjestelmän mukaisista koordinaateista on kyse. Näiden tietojen lisääminen suurentaa PDF-tiedoston kokoa mitättömän vähän eikä niistä ole mitään muutakaan haittaa. Ne PDF-lukuohjelmat, jotka eivät osaa käyttää koordinaattitietoja eivät näytä koordinaattitietoja, mutta muuten ne kyllä näyttävät kartan kuvana täysin normaalisti.

Kaksi standardia

Paikkatiedon liittämiseksi PDF-tiedostoon on olemassa kaksi julkistettua ja suhteellisen yleisesti käytettyä tapaa. Näistä toinen on julkaistu Adoben tekemänä laajennoksena PDF-standardiin (ISO32000), toinen taas perustuu TerraGo-yhtiön kehittämään menetelmään, ja se on julkaistu Open Geospatial Consortiumin Parhaat käytännöt -julkaisuna. Kumpaakin menetelmää voidaan käyttää vapaasti ilman lisenssimaksuja tai muita rajoituksia, mutta TerraGo on rekisteröinyt itselleen tuotemerkin ”GeoPDF”, ja sitä saa käyttää vain TerraGo-yhtiön ohjelmistoilla tehdyistä tiedostoista.

Jatkossa näistä kahdesta menetelmästä käytetään nimiä ”ISO-PDF” ja ”OGC-PDF”

Paikkatieto-PDF-standardien erot teoriassa

Teoriassa eri paikkatieto-PDF-strandardit eivät tue eivätkä häiritse toisiaan millään tavalla. Molemmilla on mahdollista tehdä samat asiat eikä toista voi väittää toista paremmaksi. Se, minkälaisista toiminnoista loppukäyttäjä pääsee nauttimaan, on täysin PDF-lukuohjelmasta eikä paikkatiedon tallentamiseen valitusta standardista kiinni. Koska pohjimmiltaan kyse on vain PDF:n sisään kirjoitetuista metatiedoista, niin periaatteessa on mahdollista tallentaa vaikka molempien järjestelmien mukaiset metatiedot samaan PDF-asiakirjaan ilman että ne häiritseisivät toisiaan. Käytännössä näin ei tietenkään ole, ja tähän palataan hieman myöhemmin.

Paikkatieto-PDF-standardien erot käytännössä

ISO-PDF

Adobe Acrobat Reader -lukuohjelmalla voi versiosta 9 alkaen lukea ISO-standardin mukaisesti tallennetut paikkatiedot ilman mitään erityisiä toimenpiteitä. Koordinaattinäyttö saadaan käyttöön valikosta ”Muokkaa – Analyysi – Geospaatialinen sijaintityökalu”. Työkalu ei kovin ihmeellisiä toimintoja tarjoa. Se näyttää kartalta osoitetun paikan WGS84-koordinaatit, tai käyttäjä voi antaa syöttöruutuun WGS84-koordinaatit ja Acrobat Reader piirtää kartalle merkin ja kohdistaa kuvan sen kohdalle.

OGC-PDF

Acrobat Reader ei löydä OGC-menetelmällä tallennettua koordinaattitietoa ollenkaan ennen kuin ohjelmaan on asennettu TerraGo-yhtiön tekemä lisäosa. Lisäosan voi ladata ilmaiseksi TerraGo:n sivustolta <http://terragotech.com/> mutta ennen lataamista käyttäjän on rekisteröidyttävä. Ohjelma on saatavissa vain Windowsille. Lisäosan asentamisen jälkeen Acrobat Reader -ohjelmaan ilmestyy uusia toimintoja, jotka saadaan käyttöön oman TerraGo-valikon kautta. Ominaisuuksiin kuuluu koordinaattien näyttö, joka on toteutettu monipuolisemmin kuin Acrobat Reader:in omassa toiminnossa. Koordinaatit saadaan aina näkyviin kahden oletusjärjestelmän mukaisesti, eli WGS84-maantieteellisinä koordinaatteina sekä USA:n armeijan käyttämän MGRS-järjestelmän koordinaatteina, mutta näiden lisäksi voidaan valita koordinaattinäyttö jonkin usean sadan ohjelman mukana tulevan järjestelmän mukaisesti. Muun muassa kaikki Suomen KKJ-kaistat on listalla mukana, samoin kuin ETRS-TM35FIN (EPSG:3067) ja kaikki ETRS-Gauss-Krüger-kaistat. Jos haluttua järjestelmä ei löydy suoraan valikosta, niin käyttäjä voi myös antaa omat projektioparametrisensa. Jokaista tuettua järjestelmää voidaan käyttää myös paikan hakemiseen käytettävien koordinaattien syöttämiseen.

Muista TerraGo-lisäosan mukana tulevista työkaluista voidaan mainita pituuden ja pinta-alan mittaustyökalut, jotka eivät kylläkään ole erityisen hauskoja käyttää, mutta ovat kuitenkin ehdottomasti tyhjää paremmat. Myös tietokoneeseen liitetyn GPS-vastaanottimen ilmoittama paikka voidaan näyttää punaisena rastina kartalla.

Paikkatieto-PDF-tiedostojen luominen avoimen lähdekoodin ohjelmistoilla

Todennäköisesti paras avoimen lähdekoodin ohjelmisto paikkatieto-PDF-tiedostojen luomiseksi tällä hetkellä (lokakuu 2012) on GDAL. Versioon 2.0 tulee merkittäviä parannuksia PDF-ominaisuuksiin, joten kannattaa ehdottomasti käyttää tätä versiota, mikä merkitsee toistaiseksi sitä, että on käytettävä GDAL:in kehitysversiota.

Aiheeseen liittyvät GDAL-ohjeet kannattaa lukea seuraavista linkeistä:

http://gdal.org/frmt_pdf.html

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

Seuraavissa esimerkeissä näkyvissä kuvaruutukaappauksissa käytetään esimerkikuvina Maanmittauslaitoksen maastokarttaa Helsingistä. Kartta ”UL4133L_RK1_1.tif” on ladattu Maanmittauslaitoksen avointen aineistojen latauspalvelusta ja muunnettu paikkatieto-PDF:ksi seuraavilla komennoilla

ISO-PDF -tiedoston luominen tiff-muotoisesta kartasta

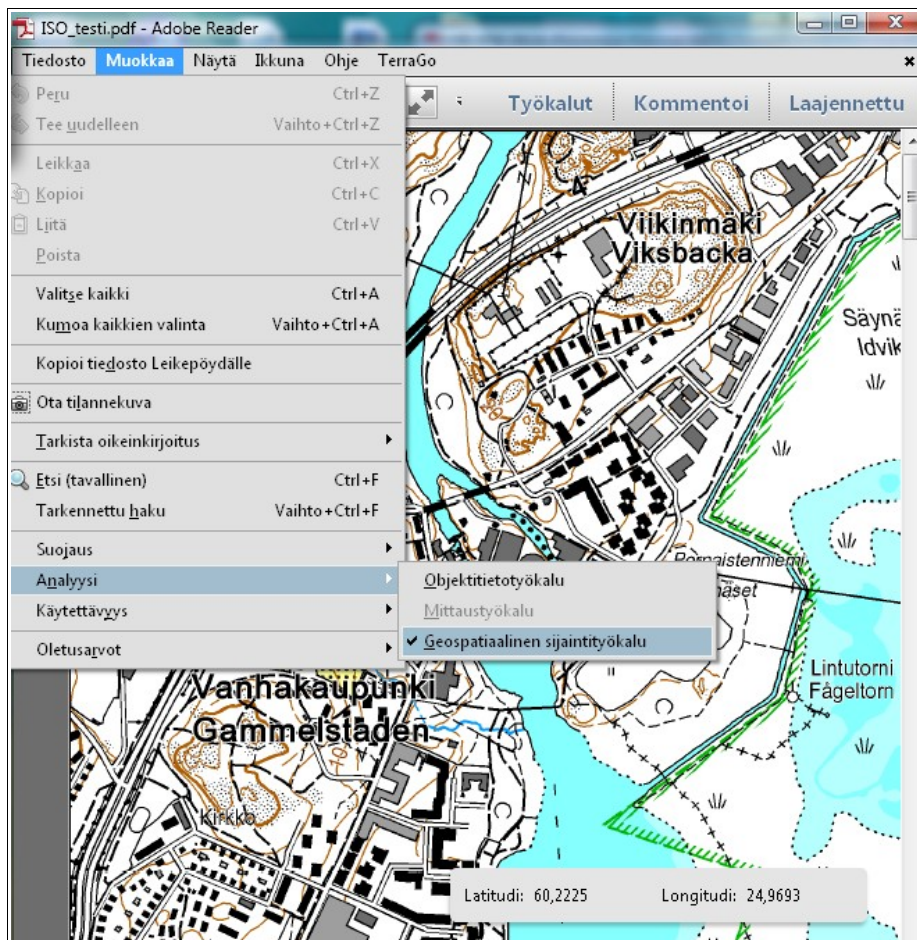
```
gdal_translate -of PDF -co GEO_ENCODING=ISO32000 -a_srs epsg:3067  
UL4133L_RK1_1.tif ISO.pdf
```

OGC-PDF -tiedoston luominen tiff-muotoisesta kartasta

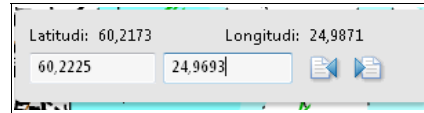
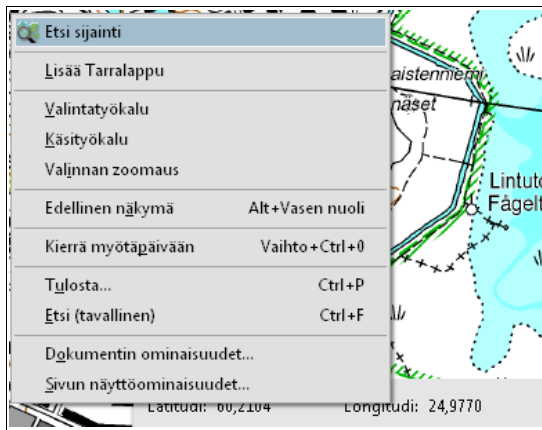
```
gdal_translate -of PDF -co GEO_ENCODING=OGC_BP -a_srs epsg:3067  
UL4133L_RK1_1.tif OGC.pdf
```

Muunnos on nopea, 1 x 3 GHz prosessorilla varustettu tiedokone vuodelta 2004 selviytyy siitä alle 20 sekunnissa.

ISO-PDF ja normaali Adobe Acrobat Reader kuvaruutukaappauksin



Koordinaatit saadaan näkyviin, kun otetaan geospaatialinen sijaintityökalu käyttöön. Työkalulle ei ole mitään muita asetuksia. Koordinaatit näytetään WGS84-järjestelmässä leveys- ja pituusasteina, yksikkönä on asteet ja asteen desimaalit.

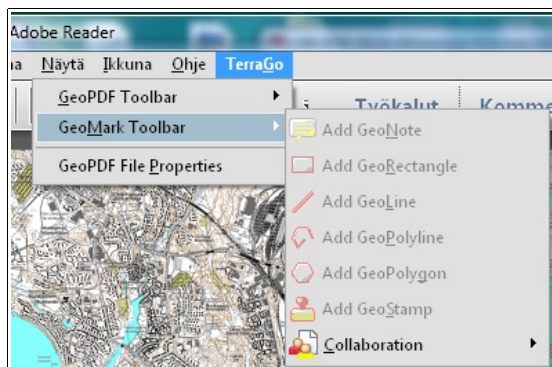


Hiiren oikea näppäin aktivoi sijainnin näyttämisen. Kartta keskittyy käyttäjän syöttämien koordinaattien kohdalle. Tarkennus menee turhankin lähelle tällä kartalla, eikä sitä voi mitenkään säätää.

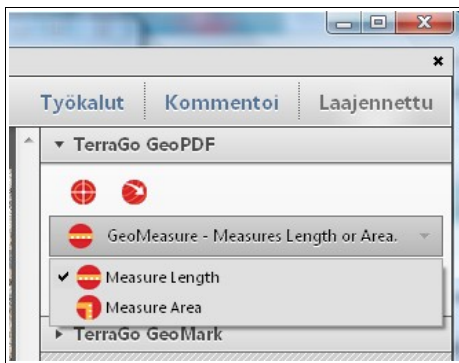
OGC-PDF, Adobe Acrobat Reader ja TerraGo Toolbar -laajennos kuvaruutukaappauksin



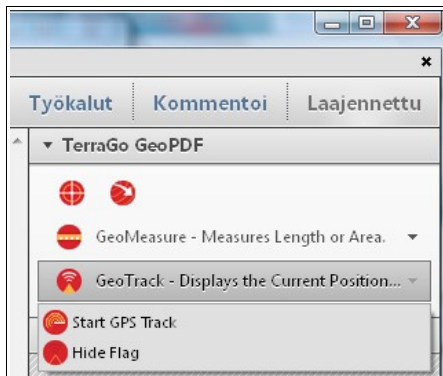
Acrobat Readeriin tulee lisäosan asentamisen jälkeen käyttöön uusi valikko ”TerraGo”



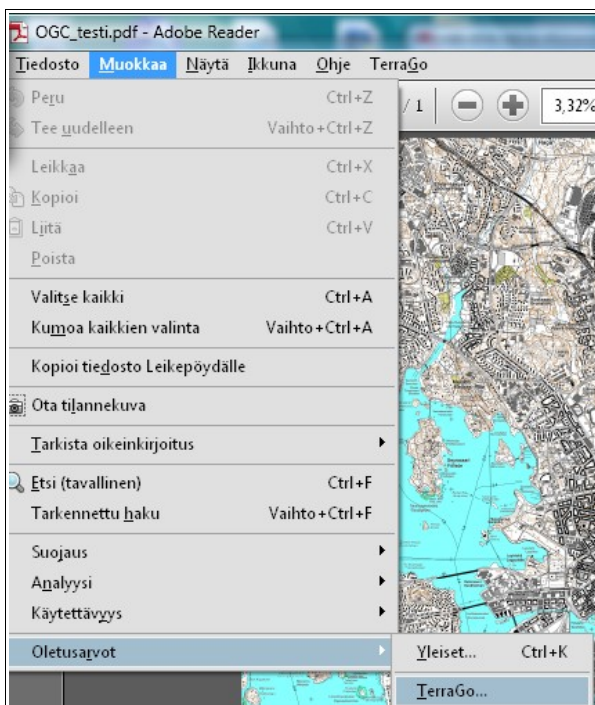
GDAL:lla luoduilla PDF-tiedostoilla ei saada kaikkia TerraGo-lisäosan ominaisuuksia käyttöön, sillä ne vaativat PDF-tiedoston avaamisen kommentointia varten, eikä GDAL pysty tekemään sitä PDF:n suojausmenetelmien takia. Avaamiseen tarvitaan joko maksullinen Adobe Acrobat -ohjelma tai jokin TerraGo-yhtiön valmistamista GeoPDF-ohjelmista.



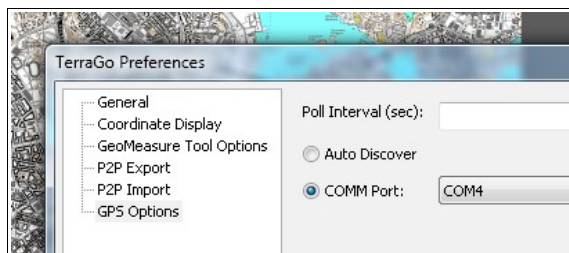
TerraGo-valikon voi avata myös Acrobat Readerin Työkalut-valikon kautta. Pituuden ja pinta-alan mittaus toimii GDAL:lla luoduilla tiedostoilla.



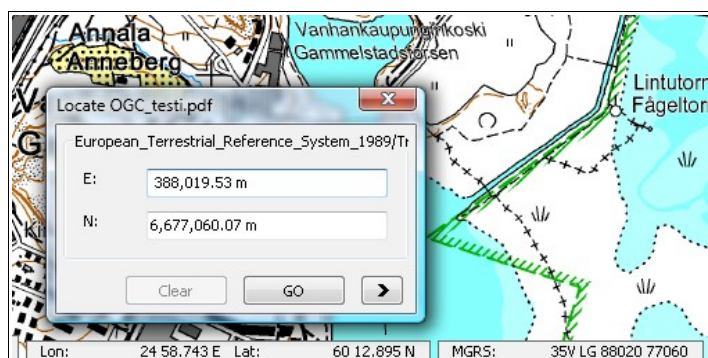
TerraGo Toolbar tarjoaa myös tuen GPS-laitteen käyttämiselle.



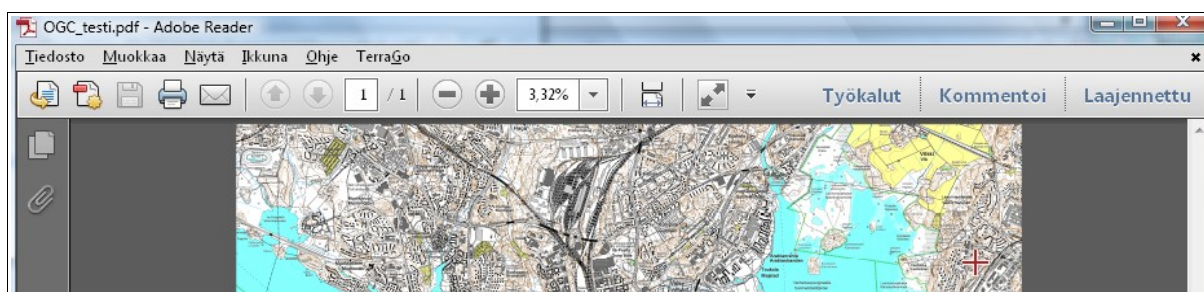
TerraGo Toolbar:in asetuksia voidaan muodata Acrobat Readerin Muokkaa-valikon kautta. Asetusmahdollisuudet ovat toisin paikoin varsin monipuoliset, esimerkiksi WGS84-koordinaattien esitysmuodon osalta.



Tässä kohtaa asetuksia valitaan GPS-vastaanottimen käyttämä portti. Nämä asetukset ovat vaatimattomat, melkein kelvottomat, sillä sarjaportin käyttämää nopeutta ei voi valita. Oletuksena on 9600 baudia, ja jos GPS-vastaanotin käyttää esimerkiksi melko tavallista 4800 baudin nopeutta, niin GPS-kytkentä ei toimi.



Koordinaatit näytetään WGS84-koordinaatteina, USA:n armeijan MGRS-järjestelmän mukaan ja käyttäjän valitseman järjestelmän mukaan. Tässä käytössä suomalainen ETRS-TM35FIN.



Punainen rasti näyttää GPS-laitteen ilmoittaman paikan.



PDF-tiedostosta voidaan myös siirtyä katsomaan kartalta osoitettua paikkaa selaimella Google Maps-palvelusta.



Valikon Google Maps -vaihtoehdon valinta avaa selaimen uuden ikkunan Googlen karttapalvelua varten. Kuvan vihreä nuoli näyttää PDF-tiedostosta osoitun paikan. *Tämän kuvan © Google 2012.*

Käytännön kokemuksia ISO-PDF:n ja OGC-PDF:n yhteensopivuudesta

Periaatteessa, mutta vain periaatteessa sekä ISO- että OGC-metatiedot voidaan kirjoittaa samaan PDF-tiedostoon ilman että ne häiritsevät toisiaan. Käytännössä OGC-metatiedot eivät häiritse Acrobat Readerin vaatimatonta geospaatialista sijaintityökalua, mutta ISO-metatietojen mukanaolo estää eräiden TerraGo-toolbarin edistyneiden ominaisuuksien käytön. Näihin kuuluu ainakin GPS-paikannus ja linkitys Google Maps -palveluun. Jos loppukäyttäjä itse tekee muunnoksen paikkatieto-PDF:ksi GDAL-ohjelmalla, niin valinta ISO- tai OGC-menetelmän välillä riippuu vain siitä, haluaako asentaa TerraGo Toolbar -lisäosan vai ei. Jos sen sijaan tehdään paikkatieto-PDF:iä muita käyttäjiä varten, niin valinta on vaikeampi. Ratkaisuun vaikuttavia asioita on ainakin nämä:

- Jos halutaan, että jotain koordinaatteja saadaan näkyviin pelkällä Acrobat Reader-ohjelmalla, niin ISO-metatiedot on kirjoitettava mukaan.
- Jos halutaan, että käyttäjät voivat liittää PDF-karttaan sijainnin näyttämisen GPS-vastaanottimen sijainnin mukaan tai että käyttäjät voivat käyttää linkitystä Google Maps -palveluun, niin ISO-metadatta ei saa kirjoittaa.
- Jos GPS- ja Google Maps -ominaisuuksien puuttuminen ei haittaa, niin kannattaa kirjoittaa sekä ISO- että OGC-metatiedot. Koordinaatit tulevat kaikkien käyttöön, ja TerraGo Toolbar:in asentajat saavat käyttöönsä muutamia lisäominaisuuksia.

Se, että ISO-metadatan toiminta häiritsee TerraGo Toolbar:in toimintaa johtuu joko lisäosan ohjelmointivirheestä tai halusta kannustaa paikkatieto-PDF-karttojen tekijöistä käyttämään OGC-menetelmää. Voi olla, että ohjelmointivirhe korjataan tulevissa versioissa, jolloin sekä ISO- että OGC-yhdistelmästä tulisi paras vaihtoehto, joka tarjoaa kaikille käyttäjille parasta mitä voidaan tarjota. Voi myös olla, että ohjelman toiminta pysyy samanlaisena kuin se on nyt.

Nopeammin, paremmin ja halvemmalla eli paikkatieto-PDF:n ja GDAL:n PDF-ajurin arviointi

Paikkatieto-PDF on karttojen esittämiseen ehdottomasti parempi kuin ei-paikkatieto-PDF. Siinä ei ole mitään vähemmän, mutta jotain lisää. Kovin hyviä paikkatieto-ominaisuuksia PDF ei kuitenkaan tarjoa etenäkään tietojen muokkaamiseen. Standardien avoimuudesta huolimatta paikkatieto-PDF:iä tukevia ohjelmia ei ole kovin paljon, eikä uusien tekeminen ole helppoa, sillä edistyneiden ominaisuuksien toteuttamiseksi olisi hallittava PDF-tiedostojen suojaamiseen käytettävät menetelmät.

GDAL muuntaa rasterikuvan PDF-muotoon nopeasti (12000 x 12000 pikselin kokoisesta tiff-tiedostosta PDF-muotoon 20 sekunnissa vuoden 2004 toimistotietokoneella) ja oletusasetukset tuottavat hyvälaatuisen tuloksen. Lähtöaineistosta riippuen kannattaa kuitenkin kokeilla eri pakkausvaihtoehtoja (-co COMPRESS), sillä oletuksena olevan DEFLATE-menetelmän sijasta voi kannattaa käyttää JPEG- tai JPEG2000-pakkausta esimerkiksi jos ilmakuvia PDF-muotoon, sillä niillä päästään parempaan pakkaussuhteeseen ja pienempään tiedostokokoon. Samoin jossain tapauksissa voi kannattaa tehdä PDF-tiedosto, jossa rasterikartta on jaettu sisäisesti pieniin laattoihin (-co TILED=YES) oletuksena olevan yhden suuren kuvan sijasta.

GDAL ei ainakaan toistaiseksi voi kirjoittaa PDF-tiedostoon kuin yhden rasterikarttatason, vaikka olisikin hauskaa, jos käyttäjä voisi samasta PDF-tiedostosta valita näytettäväksi ilmakuvatason tai maastokarttatason tai vaikkapa historiallisen kartan muutaman kymmenen vuoden takaa. Kaupalliset paikkatieto-PDF-ohjelmat pystyvät tähän, ja siinä voi olla syy, jonka takia kannattaa maksaa jostain muusta ohjelmistosta, vaikka GDAL:lla pystyy tekemään sen mihin se pystyy ilmaiseksi. Kuitenkin myös GDAL:in avulla voidaan lisätä PDF-tiedostoon rajaton määrä ylimääräisiä vektoritasoja, mikä sekin on oikein mukava ominaisuus ja voi tehdä kartasta paremman kuin mitä pelkkä koordinaatteja näyttävä rasterikartta on. Tästä aiheesta kertoo latuviitan ohje edistyneiden paikkatieto-PDF-tiedostojen tekemiseksi GDAL-ohjelmalla (*työn alla*).

Paikkatieto-PDF ei ainakaan toistaiseksi ole käytännössä äärimmäisen avoin tiedostomuoto. Ei liene vaikeaa löytää ihmisiä, joita ärsyttää suunnattomasti se, että laajimpien mahdollisten ominaisuuksien saamiseksi käyttäjän tietokoneessa on oltava Windows-käyttöjärjestelmää, Adoben valmistama Acrobat Readeria ja TerraGo-yhtiön valmistamaa lisäosa, jonka saadakseen täytyy vielä rekisteröityä yhtiön tietokantoihin. Kaiken lisäksi monipuolisimpien PDF-tiedostojen tekemiseen pitää käyttää maksullisia ohjelmistoja. Kuitenkin ihan mukavia tiedostoja voi tehdä avoimen lähdekoodin GDAL:lla ja koska lukuohjelma ja lisäosa ovat maksuttomia, niin myös paikkatieto-PDF:n avulla voidaan välittää loppukäyttäjille hyvää, ilmaista paikkatietoa. Hip.