

Paikkatiedot LIDossa

Tälle sivulle on koottu LIDossa tulevien paikkatietojen nykyisiä ongelmia Finnan näkökulmasta ja hahmoteltu ehdotus siitä, miten LIDOn paikkatietoa voisi parantaa erityisesti geokoodauksen näkökulmasta. Paikkatietoa hyödynnetään Finnassa muun muassa mobiilikäyttäjän sijaintiin perustuvassa hakutoiminnossa, Finna Streetissä (<https://finna.fi/Search/Street>). Alla kuvattu tapa merkitä paikkatietoa edesauttaa aineistojen löytymistä sijainin perusteella. Geokoodaukseen Finnassa käytetään tällä hetkellä OpenStreetMap Nominatim -palvelua. Yksittäisen paikkatiedon toimivuutta voi kokeilla osoitteessa nominatim.openstreetmap.org.



HUOM! Tässä ohjeessa ei oteta kantaa museoiden luettelointiohjeeseen, vaan vain siihen, miten kokoelmanhallintajärjestelmien tulisi tuottaa LIDO-formaattia. Järjestelmien tulisi mahdollistaa sisällönkuvailu niin, että siitä saadaan muodostettua korkealaatuista paikkatietoa.

Ongelmat

Nykyisessä kuvailussa on useita ongelmakohtia, jotka vaikeuttavat aineiston geokoodausta. Ennen kaikkea paikkatieto on koodattu erittäin epäyhtenäisesti ja LIDO-standardin vastaisesti.

Yleisesti näyttää olevan niin, että LIDOn partOfPlace-elementti on ymmärretty väärin päin. Standardin mukaan ja esimerkeissä (ks. http://www.lido-schema.org/documents/examples/LIDO-Example_FMobj00154983-LaPrimavera.xml) rakenne menee niin, että tarkin tieto, esim. katuosoite, esitetään päätasolla, ja siitä sitten lähdetään ylempille tasoille partOfPlace-elementeillä.

1. Paikkatieto on jäsenitelemätöntä - vain näytettävä pötkö, jota on vaikea parsia.
2. Paikkatieto on jäsennetty osiin, mutta osien tyyppiä ei ole merkitty.
3. On käytetty lyhenteitä esim. "Kaikuk."
4. Samaan paikkatietoon on lisätty useita osoitteita erilaisilla merkinnöillä.
5. Paikkatieto sisältää tietoa, jota ei voida käyttää geokoodauksessa, tai joka on voinut muuttua, esim. kaupunginosa tai kuvaus rakennuksen sijainnista.
6. Paikkatiedosta puuttuu varsinainen pääelementti.
7. Paikkatietoa ei ole kuvattu hierarkkisesti tai hierarkia on väärin päin.

Mitä pitäisi tehdä

Tämä on osittain käänteinen ylläolevan ongelmaliistan kanssa. Koska käytössä olevat järjestelmät asettavat rajoituksia sille, mitä pystytään tekemään, ei kaikkia kohtia voida toteuttaa, mutta silti paikkatiedon käsittely voidaan huomioida esim. välttämällä lyhenteiden käyttöä.



Allaolevassa listassa lähdetään siitä, että järjestelmässä on mahdollista luokitella paikkatiedon osat. Jos tällaista mahdollisuutta ei ole, ja järjestelmä pystyy käytännössä tuottamaan vain LIDOn displayPlace-kentän (yksi tekstikenttä), voidaan silti huomioida geokoodaus. Tässä tapauksessa suositellaan, että paikkatiedot annetaan järjestyksessä suurimmasta paikasta pienimpään, eli esim. "Suomi, Helsinki, Kaikukatu 4". Jos samaan kenttään täytyy luetteloida useita paikkoja, voi ne erotella puolipisteellä, esim. "Helsinki, Kaikukatu 4; Helsinki, Unioninkatu 36". Huom! Molemmissa täydellinen osoite.

1. Esitetään paikkatieto hierarkkisesti:

```

<place>
  <namePlaceSet>
    <appellationValue>Kaikukatu 4</appellationValue>
  </namePlaceSet>
  <placeClassification>
    <term>katuosoite</term>
  </placeClassification>
  <partOfPlace>
    <namePlaceSet>
      <appellationValue>Sörnäinen</appellationValue>
    </namePlaceSet>
    <placeClassification>
      <term>kaupunginosa</term>
    </placeClassification>
    <partOfPlace>
      <namePlaceSet>
        <appellationValue>Helsinki</appellationValue>
      </namePlaceSet>
      <placeClassification>
        <term>kunta/kaupunki</term>
      </placeClassification>
      <partOfPlace>
        <namePlaceSet>
          <appellationValue>Suomi</appellationValue>
        </namePlaceSet>
        <placeClassification>
          <term>maa/valtio</term>
        </placeClassification>
      </partOfPlace>
    </partOfPlace>
  </partOfPlace>
</place>

```

2. Vain yksi paikka per place-elementti
3. Ei lyhenteitä
4. Kaupunginosa vain, jos se on luokiteltu sellaiseksi
5. Pääasia päätasolla
6. Hierarkia käytössä ja oikein päin
7. Käytetään yhteisesti sovittuja luokittelutermejä, esim.
 - katuosoite
 - saari
 - kunta/kaupunki
 - maa/valtio
8. Kunhan ylläolevat ovat yhteisiä, voidaan käyttää muitakin luokituksia, joita ei tarvita geokoodauksessa. Hyvä näidenkin olisi olla yhteisiä, esim.
 - rakennus
 - tila
 - tontti
 - kartano
 - kortteli
 - puisto
 - hautausmaa
 - tori
 - aukio
 - satama
 - kaupunginosa
 - kylä
 - maakunta
 - osavaltio
 - maanosa
 - pitjäjä/entinen kunta
 - löytöpaikka
 - avaruus/taivaankappale
9. Vältetään kuitenkin liian epämääräisiä luokituksia, joiden avulla ei voi päätellä mitään, kuten
 - alue
 - paikka
 - luonnon paikka
 - maantieteellinen alue/paikka
 - hallinnollinen alue (tarkempi tieto hallinnollisen alueen tyypistä olisi hyödyllisempi)
 - nyk. tai ent. hallinnollinen alue
10. Paikkaluokkiin ei laiteta mukaan tietoa esim. siitä, minkä maan kunta tai kaupunki on kyseessä, vaan se laitetaan omaan partOfPlace-elementtiinsä ko. kohdan alle, esim.

```

...
<partOfPlace>
  <namePlaceSet>
    <appellationValue>Korsö</appellationValue>
  </namePlaceSet>
  <placeClassification>
    <term>saari</term>
  </placeClassification>
  <partOfPlace>
    <namePlaceSet>
      <appellationValue>Porvoo</appellationValue>
    </namePlaceSet>
    <placeClassification>
      <term>kunta/kaupunki</term>
    </placeClassification>
    <partOfPlace>
      <namePlaceSet>
        <appellationValue>Suomi</appellationValue>
      </namePlaceSet>
      <placeClassification>
        <term>maa/valtio</term>
      </placeClassification>
    </partOfPlace>
  </partOfPlace>
...

```

ja

```

...
<partOfPlace>
  <namePlaceSet>
    <appellationValue>Korsö</appellationValue>
  </namePlaceSet>
  <placeClassification>
    <term>saari</term>
  </placeClassification>
  <partOfPlace>
    <namePlaceSet>
      <appellationValue>Tukholma</appellationValue>
    </namePlaceSet>
    <placeClassification>
      <term>kunta/kaupunki</term>
    </placeClassification>
    <partOfPlace>
      <namePlaceSet>
        <appellationValue>Ruotsi</appellationValue>
      </namePlaceSet>
      <placeClassification>
        <term>maa/valtio</term>
      </placeClassification>
    </partOfPlace>
  </partOfPlace>
...

```

11. Toki kaikkein parasta on, jos kuvailutiedossa on valmiiksi koordinaatit tai laajemminkin sijaintitietoa (polygonit, linestringit, keskipisteet gml-elementeissä). Finna tukee WGS84 koordinaatistoa.