Role of spatial catalogs in a search engine aware INSPIRE landscape

Workshop on SEO

Paul van Genuchten

Linda vd Brink

GEOCAT

OGC API Features (aka wfs3)

- Based on spatial data on the web best practices
- Crawlable by search engines
- [Core model in public review](#)
- Implementation was part of the standardisation process, therefore various implementations exist

GEOCAT

Discussion paper on INSPIRE IR's for Network Services mapping with WFS 3.0

Authors: Michael Lutz, Jari Reini, Alexander Kotsev
Version: 1.2 (2019-03-27)

Table of Contents

OAPI Features and Schema.org

Implementations can combine OAPI Features with schema.org (or DCAT)

Search engines will be able to distinguish datasets and offer dataset search to their users

What will then be the role of Spatial Catalogues and discovery services?

ldproxy wfs

5 results found

Liegenschaftskataster (NRW)

www.ldproxy.nrw.de

INSPIRE Download Service Adressen Gebaeudereferenzen...

www.ldproxy.nrw.de

INSPIRE Download Service Adressen Gebaeudereferenzen...

www.ldproxy.nrw.de

INSPIRE Download Service Adressen Gebaeudereferenzen

Explore at www.ldproxy.nrw.de

Dataset provided by

Bezirksregierung Koeln, Abteilung Geobasis NRW

License

Die Geobasisdaten des amtlichen Vermessungswesens werden als oeffentliche Aufgabe ge gebuehrenfrei nach Open Data-Prinzipien ueber online-Verfahren bereitgestellt. Nutzungsbec koeln.nrw.de/brk_internet/geobasis/lizenzbedingungen_geobasis_nrw.pdf

Available download formats from providers

gml-sf2";version=3.2, geo+json

Time period covered

01.01.1970 - 21.06.2019

GEOCAT

Some thoughts...

- Currently catalogues are the facilitators of search engines, wfs2 itself is not crawlable.
- Resources that are not (publicly) available need to be registered anyway.
- Search engines offer discovery, no assessment options.
- How can users distinguish authoritative content?
- Can government delegate the discovery task to the popular search platforms?

GEOCAT

Session discussion results

Group 1

CSW is useful for specialized apps (qgis metasearch). It offers capability to add a layer to the qgis project from metadata info only.

The topic of archiving and versioning requires that metadata be closely linked to datasets. Unless the metadata describes more the provenance of the dataset. A challenge here is if the standard itself changes, how does a catalog manage records which are still in the old version of the standard. Should records be updated or leave untouched for archiving purposes.

Group 2

- We should make sure not to delegate the full use case to the search engines.
- New standards developed should be inclusive for external communities.
- CSW offers a niche search experience.
- Spatial portals provide specific tools that analysts need in their work.
- For an any-search you don't need csw, this can well be managed by search engines.
- In spatial catalogues we have more than datasets, there is codelists, applications, sensors, processes.
- The pallet of OGC standards needs also a catalogue standard to complete the SDI architecture. Other OGC standards depend on this.
- Maybe the question should be: Given we have a catalog standard, what operations should it support?

Group 3

- OGC has a common query language in their O*S standards, it is useful to also use that for catalogue queries.
- Spatial community needs a specific set of queryables, such as time/space, scale, crs. Compare it to a used car website, it is more capable to find a relevant car then search engine, due to specific queryables.
- You can use google for initial search, and use the catalogue for refinement (similar datasets, facets).
- Catalogues provide an option to claim authoritative content. This is often mixed with expecting a high ranking in search engines. Ranking is based on relevance. Search engines may actually be more capable of estimating relevance then catalogues (based on their big data algotythms).
- OGC/Earth observation has a typical challenge with granularity. Many quite similar datasets need very specific queryables.
- Last comment: if it works, don't touch it.