Why are traditional SDIs not enough?

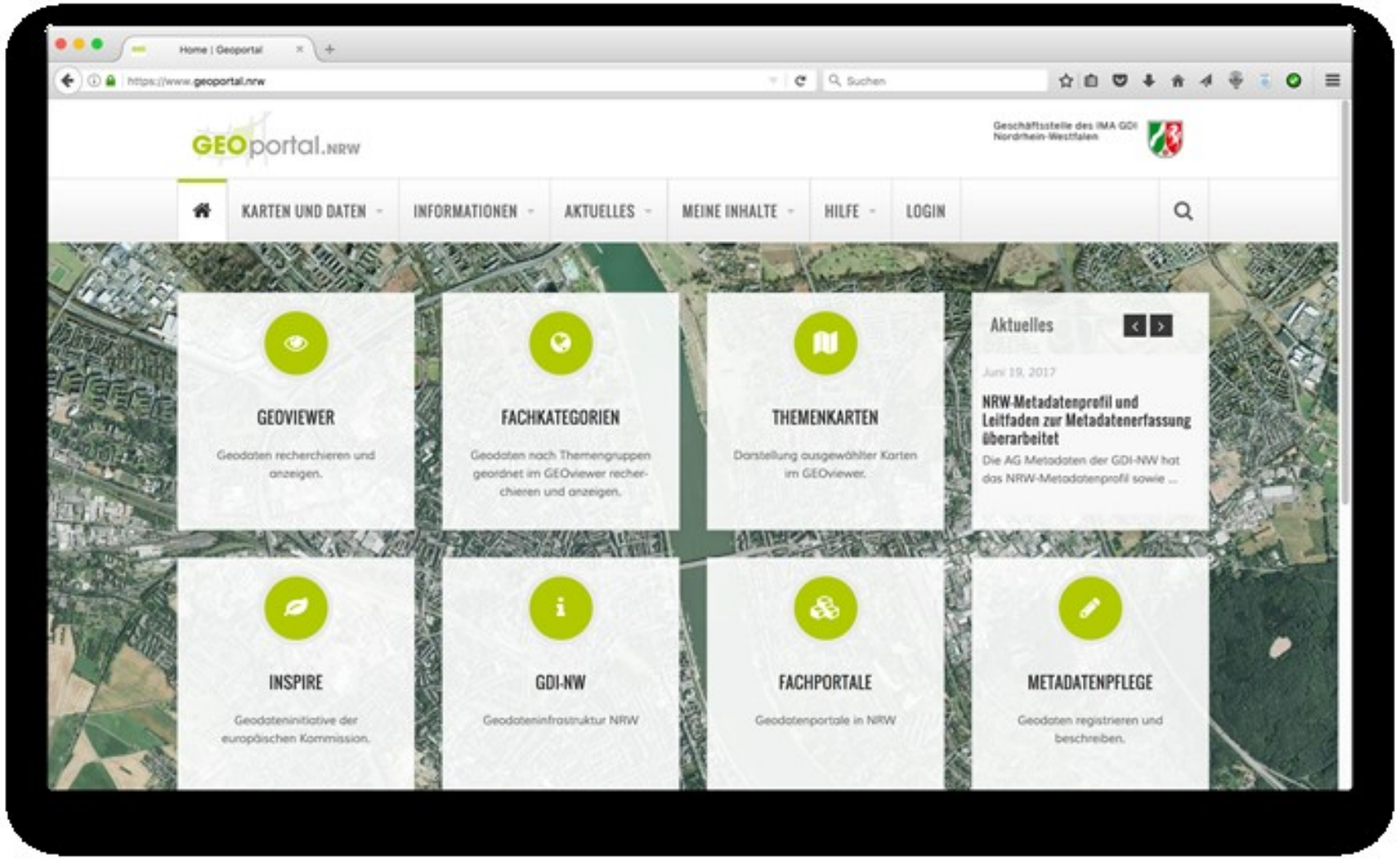
Jeremy Tandy
Met Office
Clemens Portele
interactive instruments

Spatial Data Infrastructure (SDI) implemented ✅

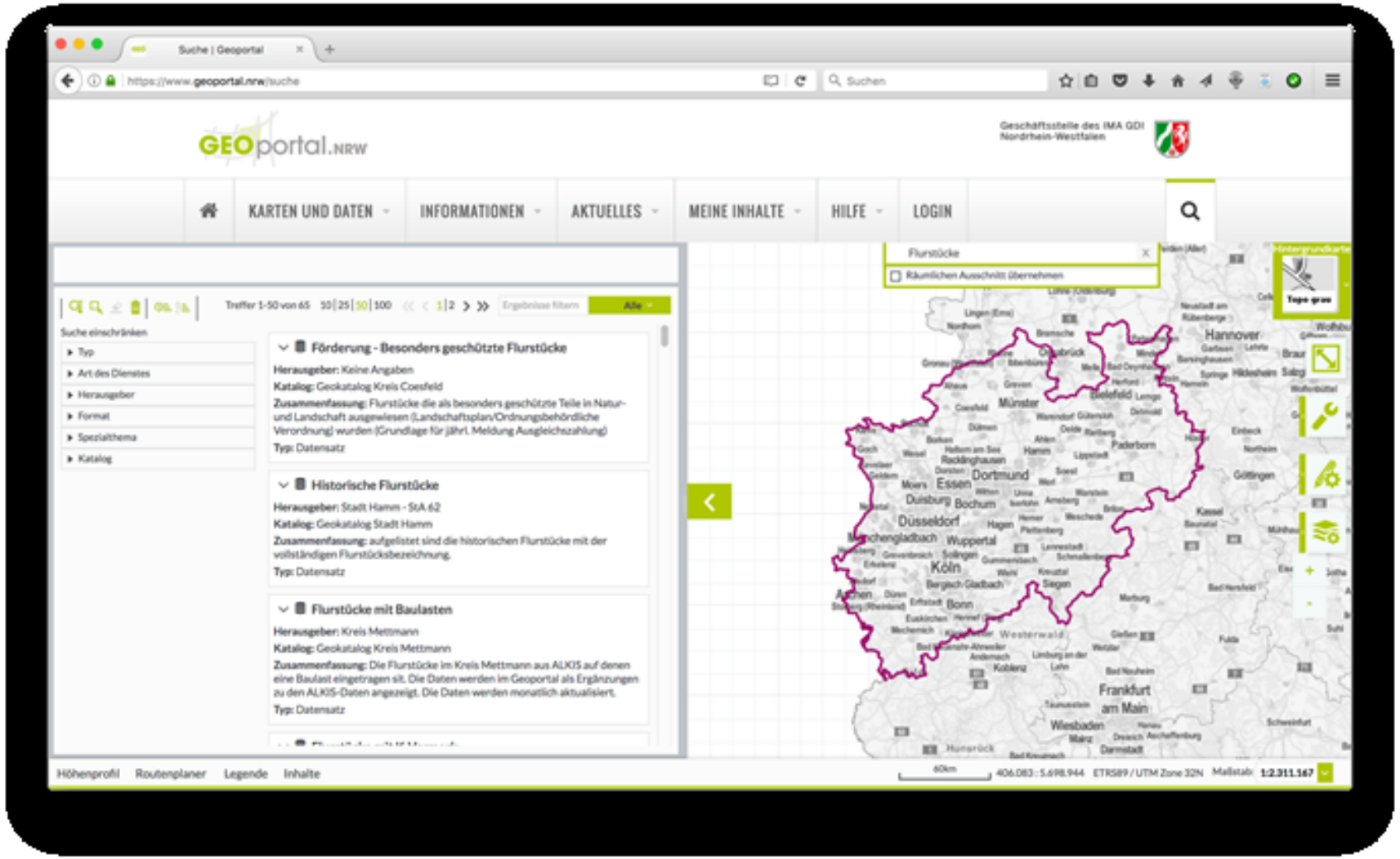
Spatial data management 👍

Spatial data usage 🤷

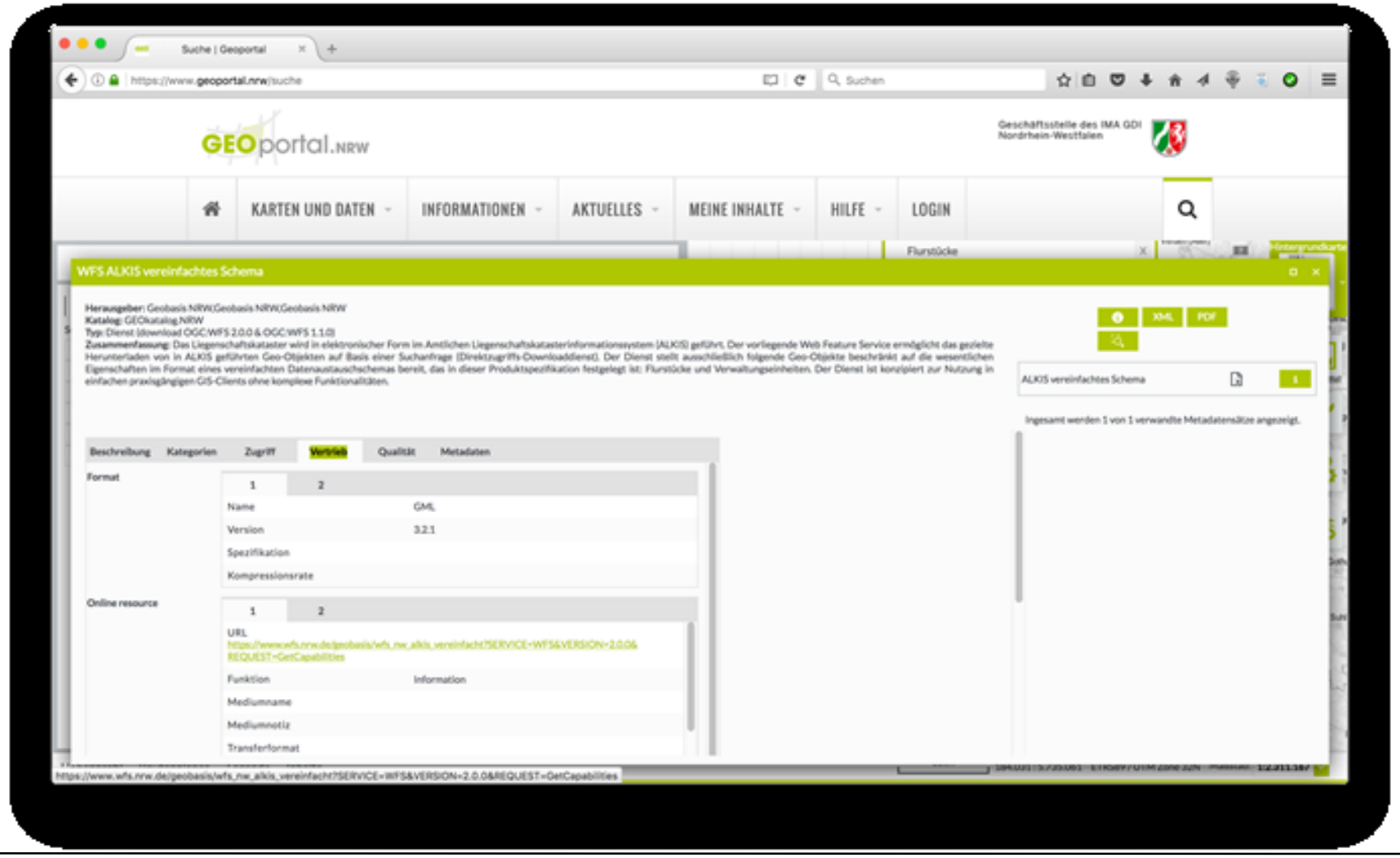
… let’s take a look at the typical journey to find and use data in an SDI →



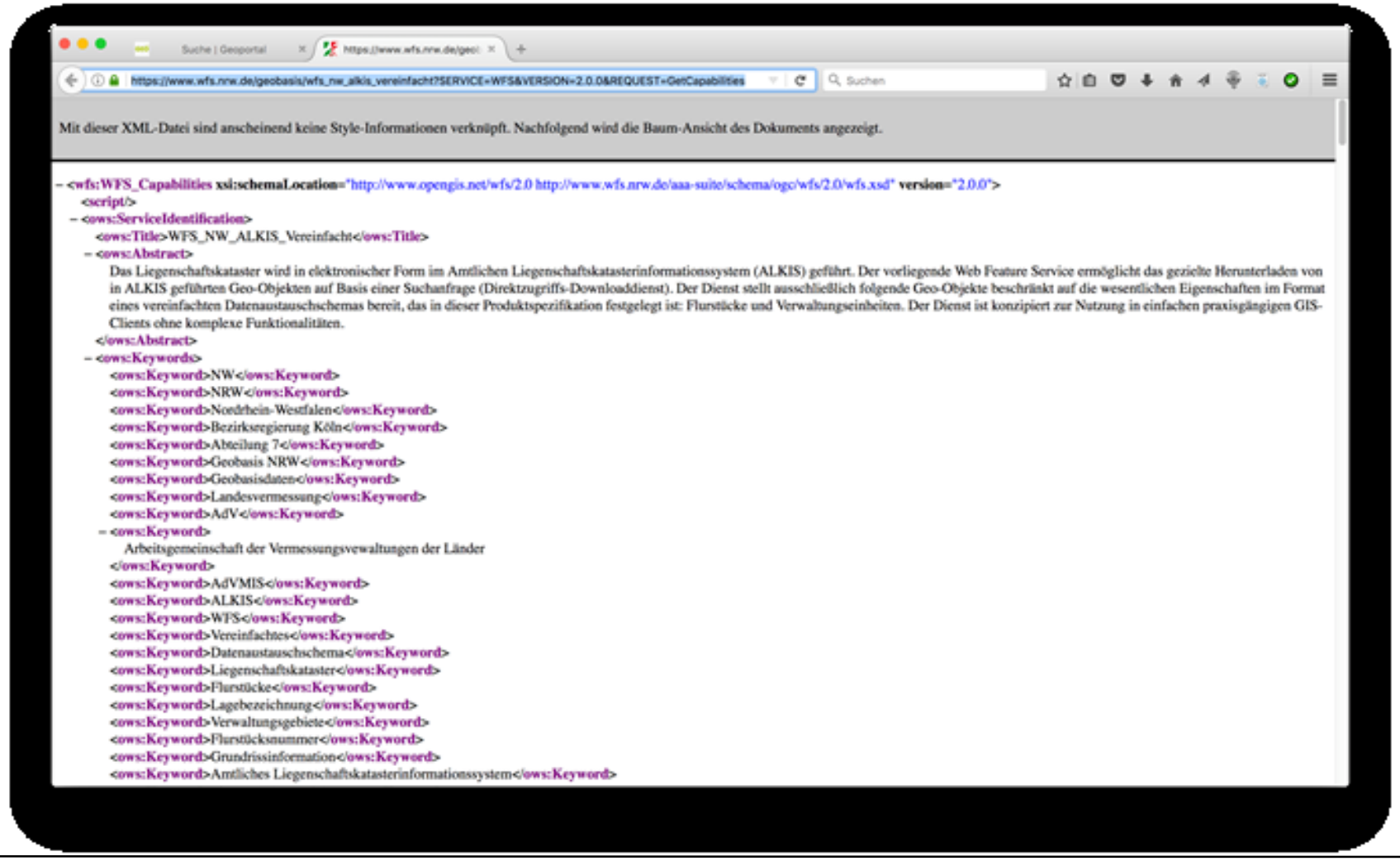
1. Open the geoportal in the browser



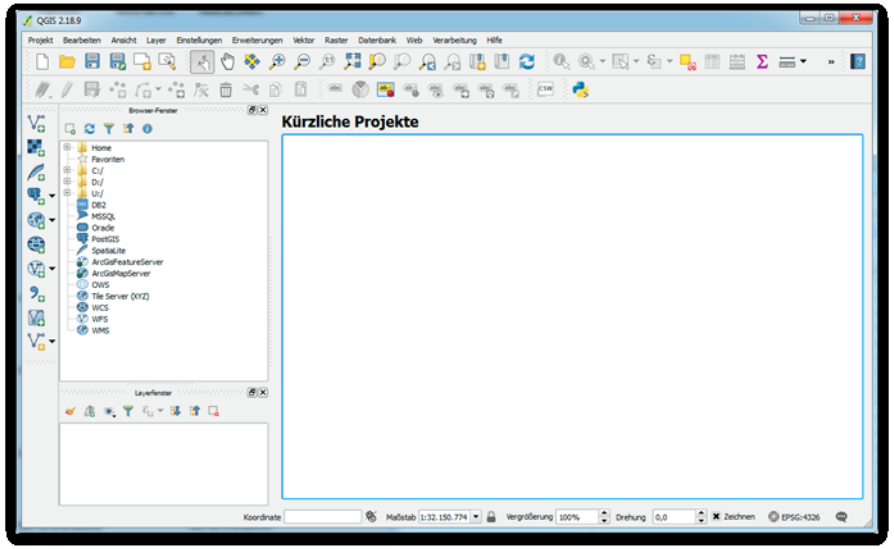
2. Navigate to search for geodata
3. Enter your search text – optionally using structured search criteria (e.g. format)
4. Browse through the results and select a dataset to look at more closely

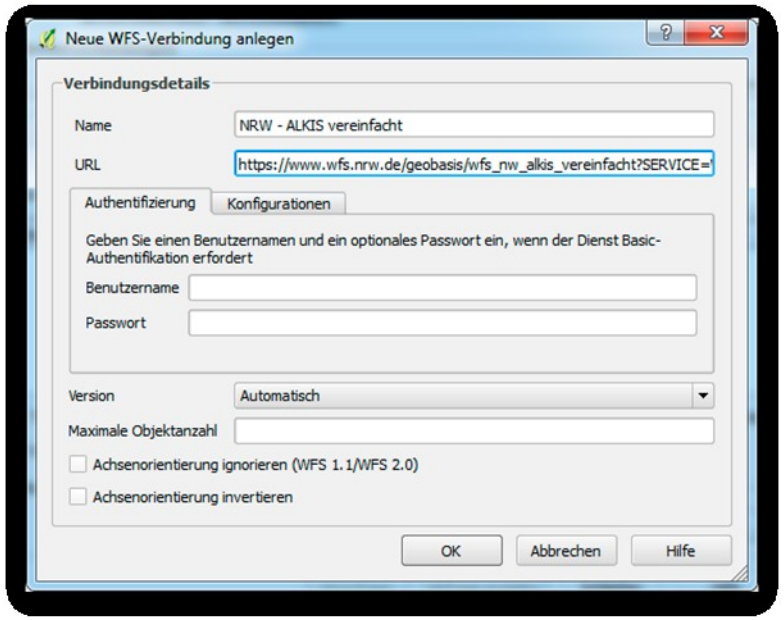


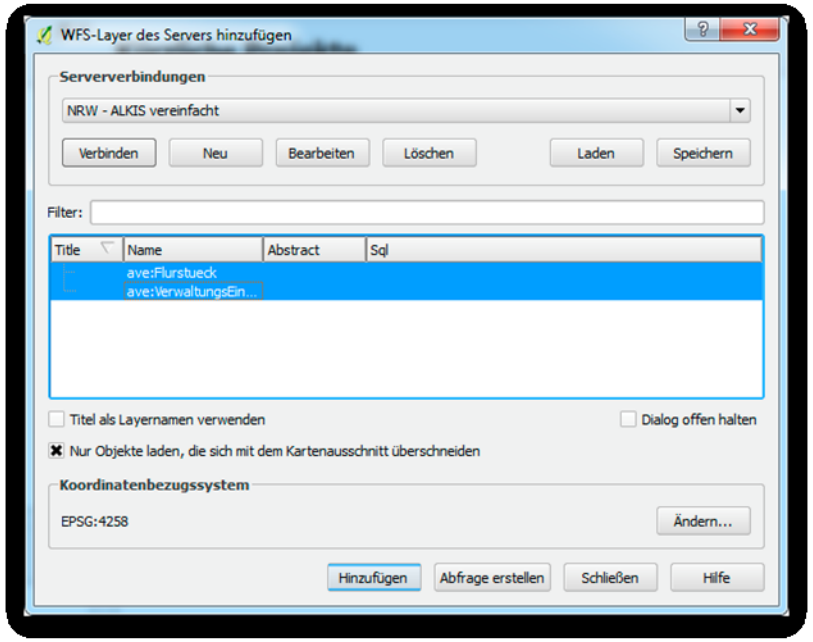
5. View the metadata



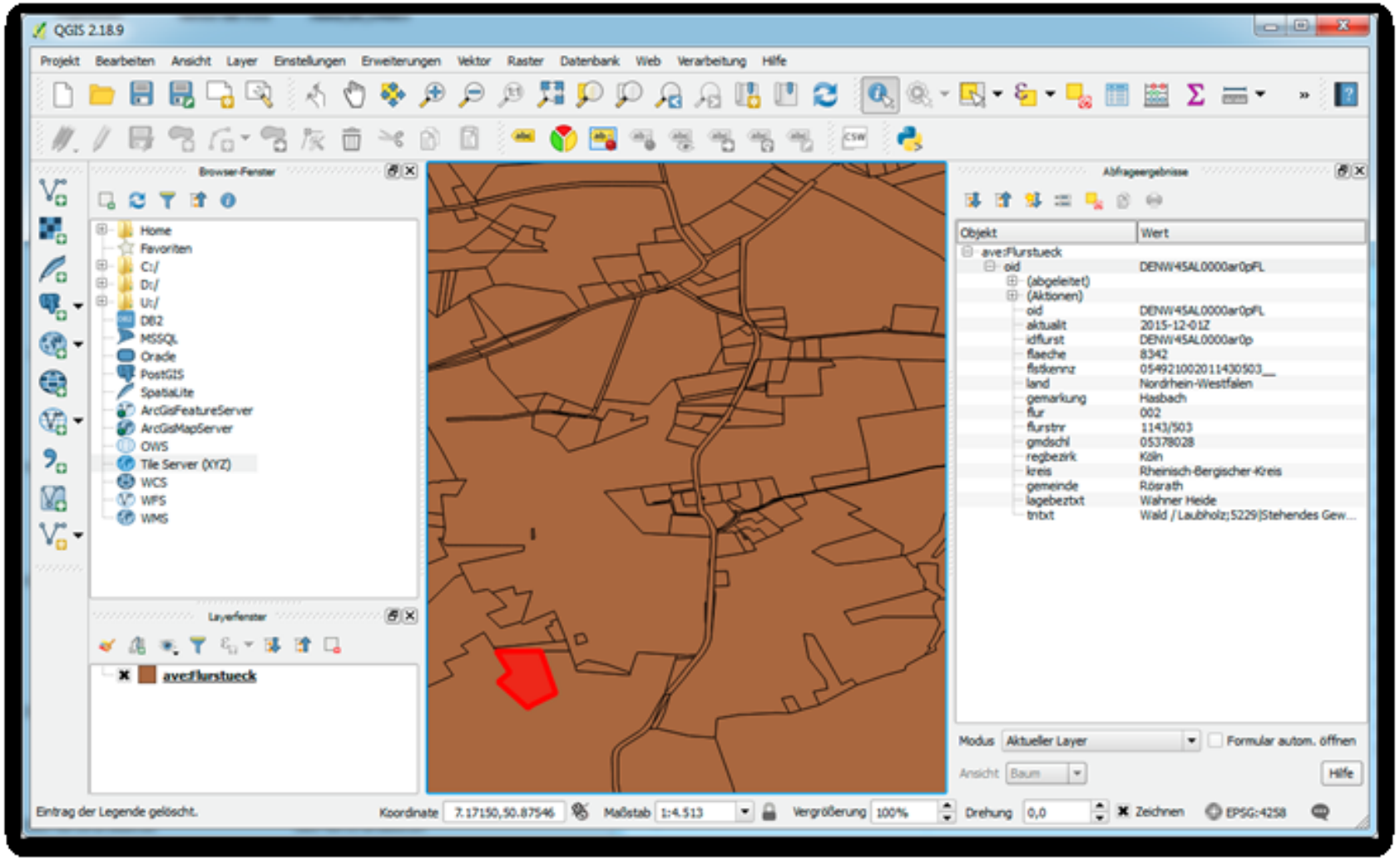
6. Copy the WFS GetCapabilities URL







7. Open a WFS client application and access the data via the WFS service



7. Analyse the record to determine if it contains the information you needed

… *then you can start using the data!*

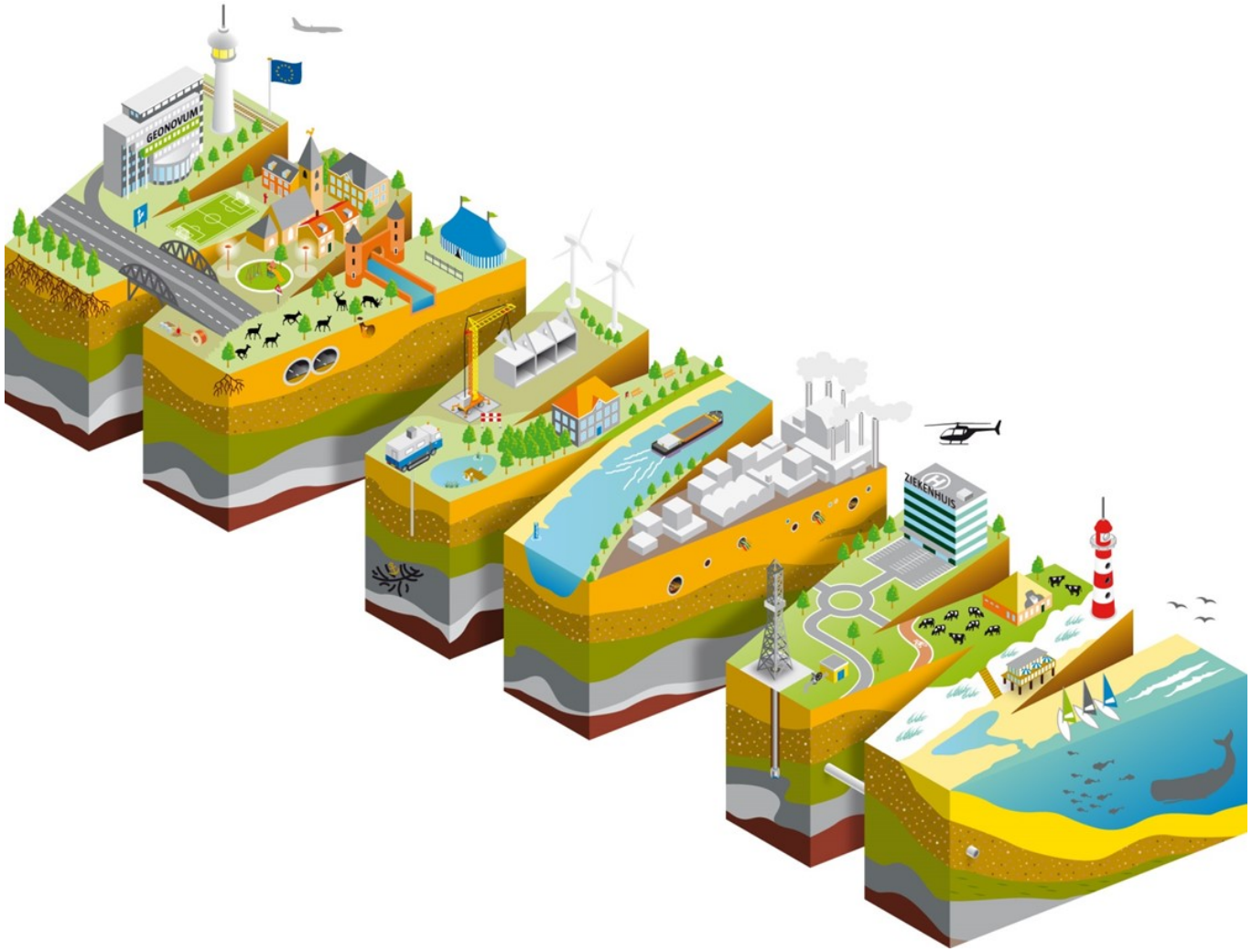
Finding, accessing and using data published though SDIs based on OGC Web services is difficult for non-expert users.

1. prior knowledge about geoportals is needed – most people start with a search engine
2. geoportals and their catalogues, by design, only provide access to metadata – not the data itself
3. geographic metadata (ISO 19115) is for GIS specialists – it is hard for many to understand

(cont.)

(cont.)

4. OGC capabilities documents are hard to interpret – what do you do with this *huge* XML document?
5. there are no links to the data itself in the capabilities document – so you need to query the Web service
6. OGC Web service standards are complex – you need expert knowledge to build your own query, or use a special application to do it for you which many users will not have
7. the data itself is often difficult for non-expert users to understand and use – domain specific complexities (data models, coordinate reference systems, formats etc.) need further explanation, but links to documentation are often unavailable



But there is another way:

… let’s repeat the “find and use” journey using the Webby approach →

1. enter search criteria for the data using a browser and search engine
2. browse through the results and check if one seems to provide your desired data or refers to it
3. click through to browse the data and determine if it contains the required information – don’t forget to check the license!
4. download the data for you to use (the whole dataset or just the part that interests you) – the data is provided in formats that can be parsed in mainstream applications and tools
5. complex data may be presented via an API – links to documentation describing how to use it are provided from the service end-point
6. an expert Web user may want to build an application that uses the data – APIs follow well-known, Web-centric patterns (e.g. REST) allowing developers to build fast using their standard tools

Use of the Web platform’s standard tools:

- search engines
- browsers
- HTTP (and HTTPS)
- hypermedia / Web links
- delegation to applications via media types
- openAPI metadata (Swagger)

W3C Working Group Note

Spatial Data on the Web Best Practices

OGC® Making location count.

W3C

W3C Working Group Note 28 September 2017

This version:

https://www.w3.org/TR/2017/NOTE-sdw-bp-20170928/

Latest published version:

https://www.w3.org/TR/sdw-bp/

Latest editor's draft:

https://w3c.github.io/sdw/bp/

Previous version:

https://www.w3.org/TR/2017/NOTE-sdw-bp-20170511/

Editors:

Jeremy Tandy, Met Office
Linda van den Brink, Geonovum
Payam Barnaghi, University of Surrey

Introducing a couple of essential concepts for spatial data on the Web →

Spatial Thing: “Anything with spatial extent, i
shape, or position, e.g. people, places, bowli
as well as abstract areas like cubes” [W3C BASIC GEO

… or even a 5 metre tall orange statue of a m
on the telephone

(Orange Man at Cité Centre de Congrès de Ly

Feature: similar – but is the digital representa
instead of the actual entity

Linked data is an approach to publishing data that puts linking at the core of data representation and uses Web linking to “weave data into a global graph”

By identifying spatial things and other resources with URLs we can link data describing those spatial things just the same as Web-pages are linked using hyperlinks

We (both humans and software) can follow those links to find out more information and build an increasingly complete picture of the world around us

https://www.wikidata.org/wiki/Q57783921

We think that the concept of Linked Data is fundamental to the publishing of spatial data Web.

(come to the Spatial Data on the Web master you'd like to learn more)

This Linked Data approach is well described by the WEB-DATA 5-star scheme:

★ **Linkable**: use stable and discoverable global identifiers

★★ **Parseable**: use standardized data metamodels such as CSV, XML, RDF, or JSON.

★★★ **Understandable**: use well-known or at least well-documented vocabularies/schemas

★★★★ **Linked**: link to other resources whenever possible

★★★★★ **Usable**: label your document with a license

If you have geodata in an SDI today then you should read:

- Best Practice 1: Use globally unique persistent HTTP URIs for Spatial Things
- Best Practice 2: Make your spatial data indexable by search engines
- Best Practice 3: Link resources together to create the Web of data
- Best Practice 12: Expose spatial data through ‘convenience APIs’

OGC Web Feature Service 3.0: Part 1 - Core

Open Geospatial Consortium

Submission Date: <yyyy-mm-dd>

Approval Date: <yyyy-mm-dd>

Publication Date: <yyyy-mm-dd>

External identifier of this OGC® document: http://www.opengis.net/doc/IS/wfs-1/3.0

Internal reference number of this OGC® document: 17-069

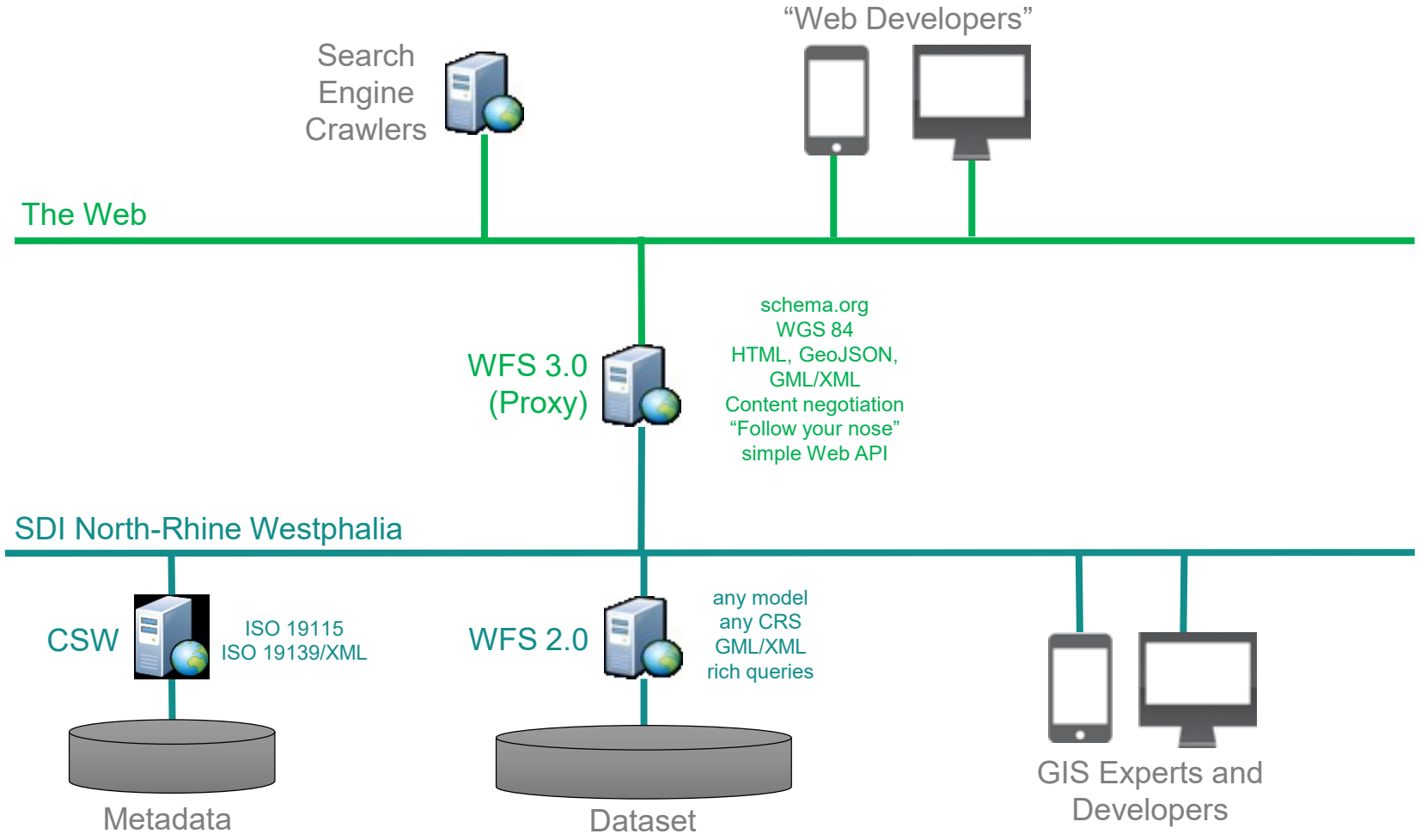
Version: 3.0.0-draft.1 (2018-04-07)

Category: OGC® Implementation Specification

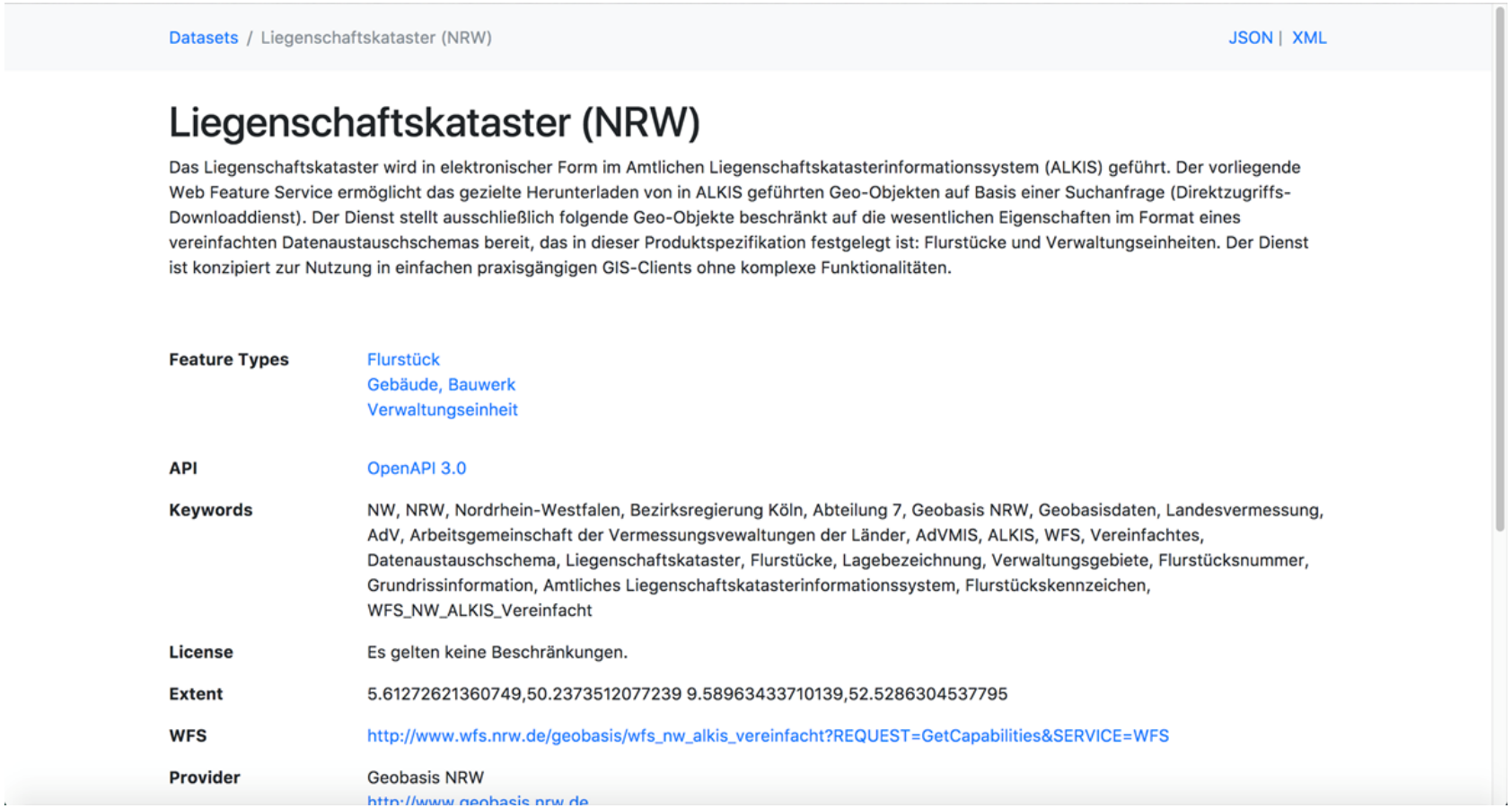
WFS 3.0 (currently in development) is strongly influenced by the Best Practices and will make it much easier for you to publish spatial data on the Web

SDIs are a key component of the broader spatial data ecosystem – providing policies, workflows and tools related to the management of spatial datasets, plus the rich set of capabilities required by the expert community.

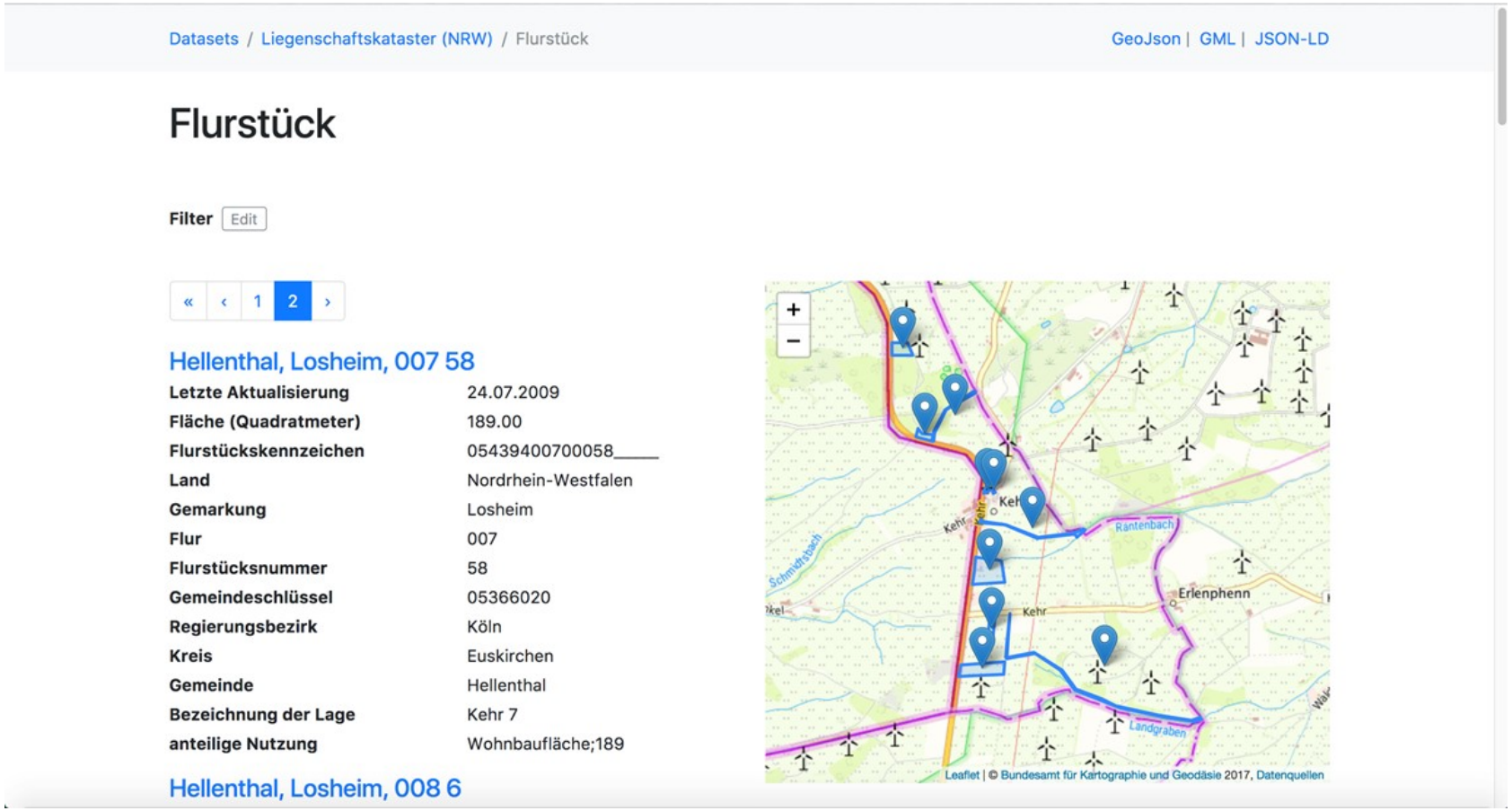
You don’t need to start again – you can leverage your existing investment in SDI …



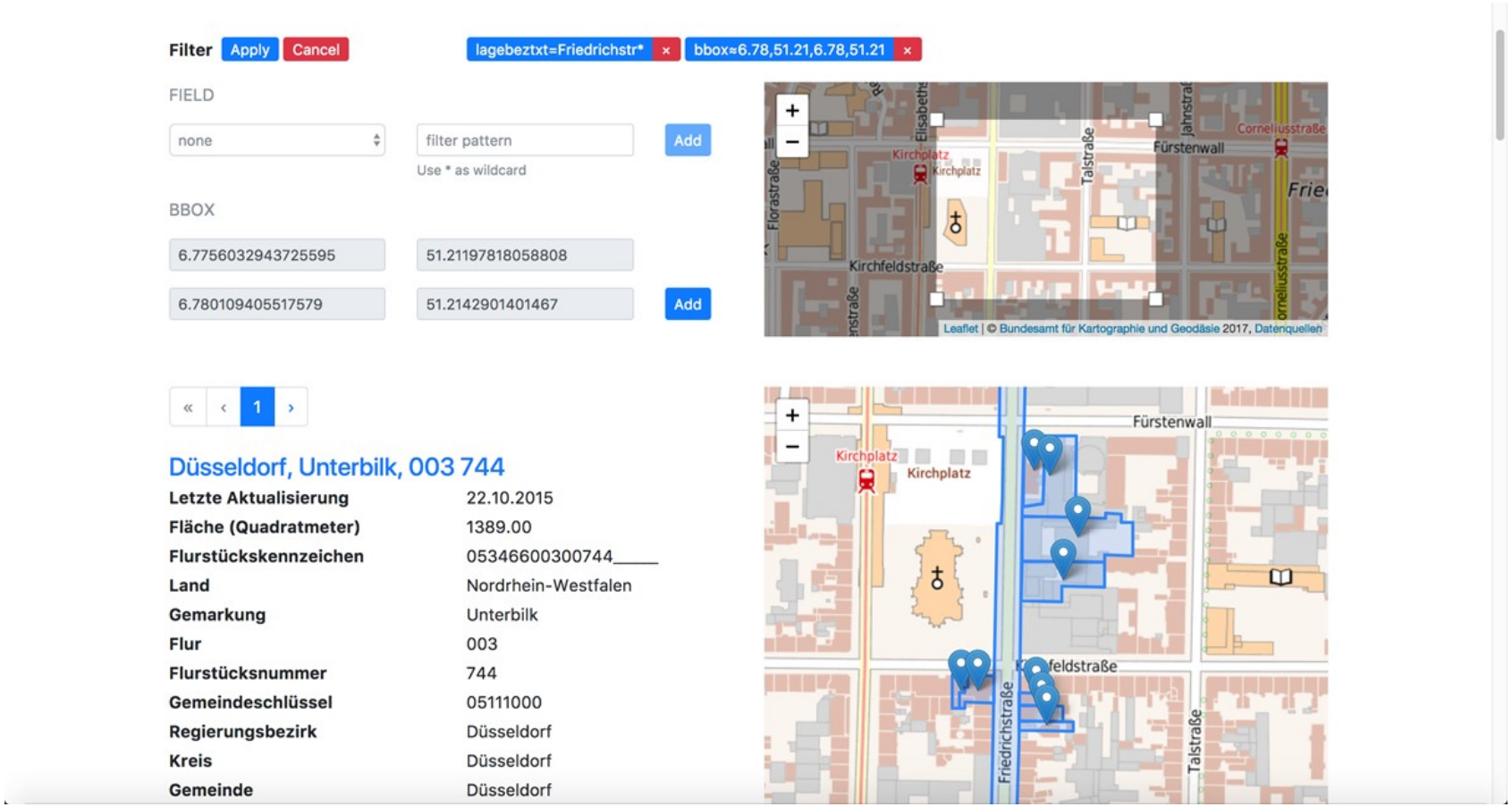
Example: Spatial data on the Web *over* North-Rhine Westphalia SDI platform (implemented by interactive instruments)



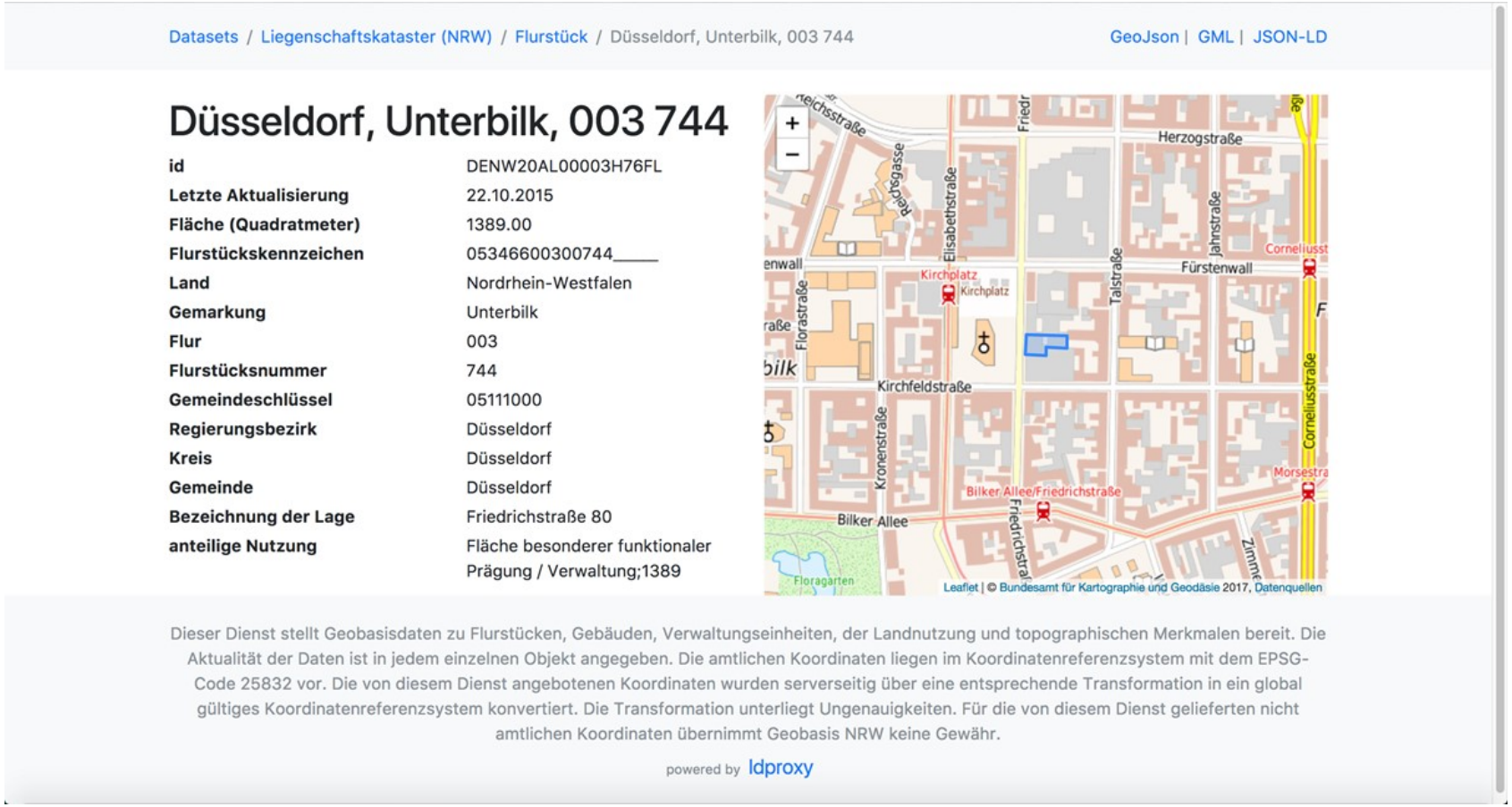
All geodata resources in HTML; browse via hyperlinks



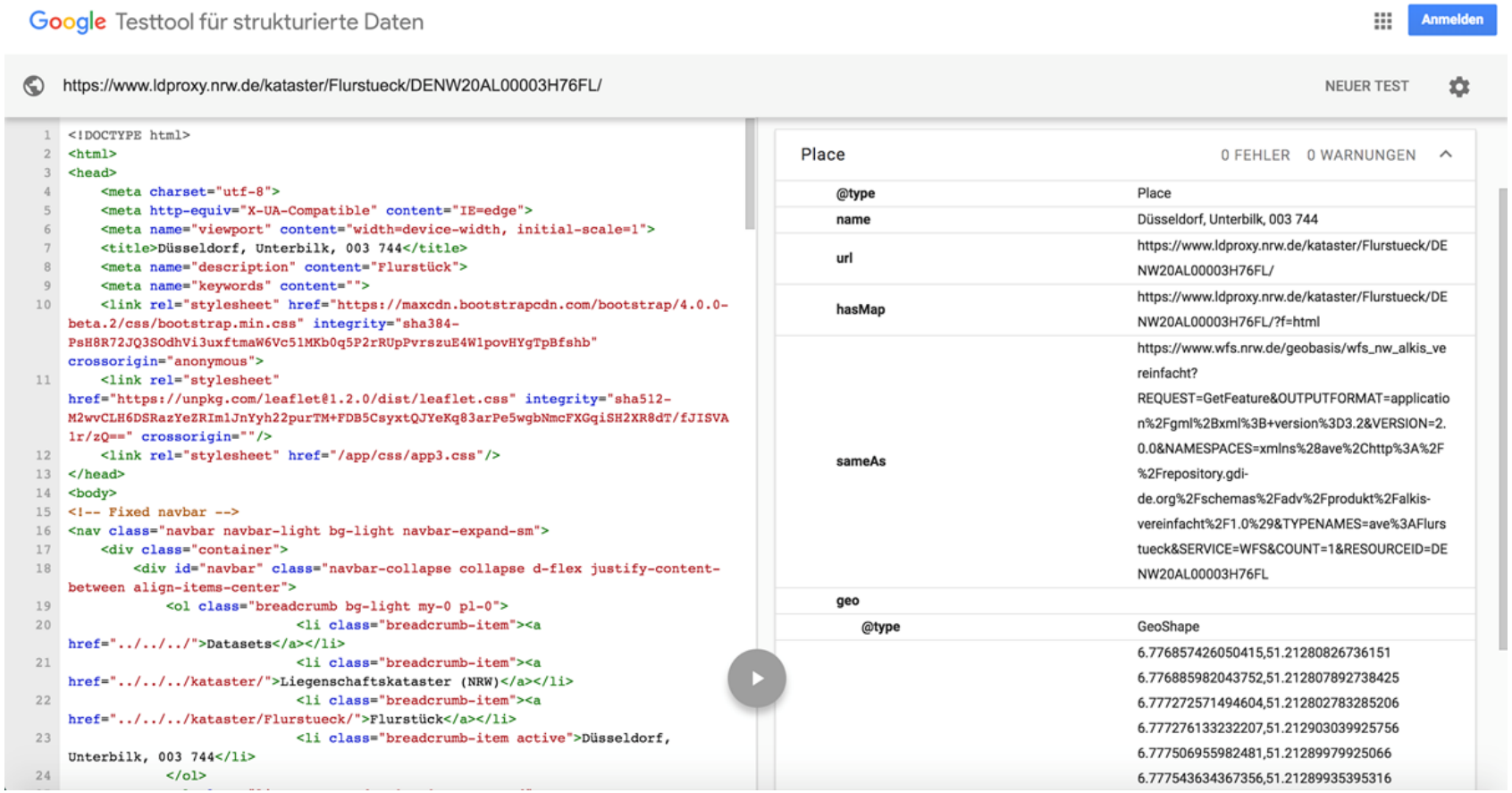
Textual content is complemented by Web maps – again supporting browsing by hyperlinks



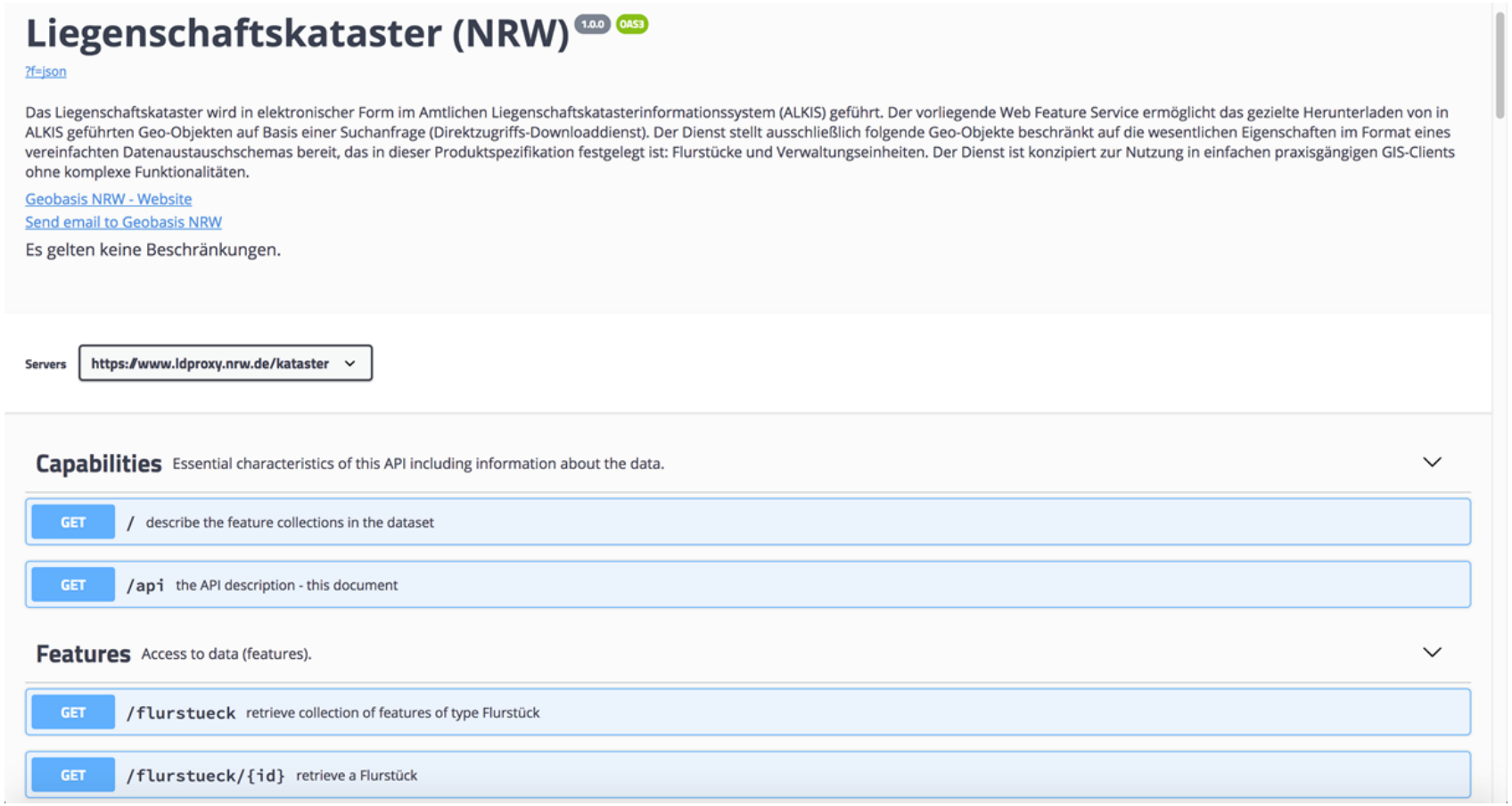
Easy data selection



Permanent URLs for all resources exposed onto the Web



Behind the scenes – HTML with schema.org annotation for search engines



Access the data via an easy to use Web API – documented with openAPI metadata (Swagger)

SDIs are useful – but they're hard for non-experts to use

Release your geodata – embrace the Web and Linked Data:

1. Use persistent URLs for your spatial things
2. Make your spatial data indexable by search engines
3. Link resources together to create the Web of data
4. Expose spatial data through Web APIs

Leverage your SDI investment – build Web-centric access on top

---

Conclusion: SDI ☺ SDW

DANK JE