Where is everywhere: bringing location to the web

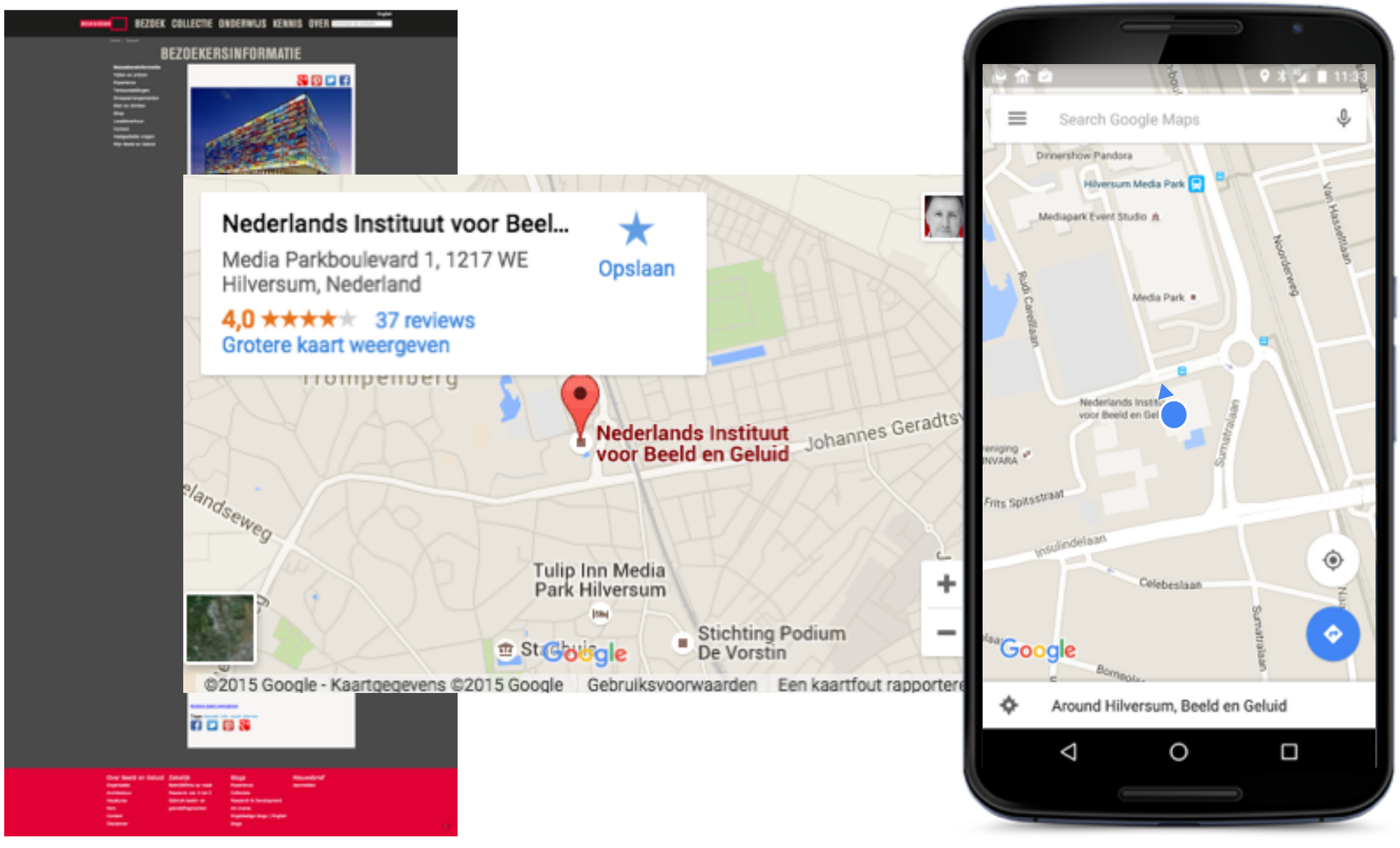
Spatial data on the web WG

Jeremy Tandy
<jeremy.tandy@metoffice.gov.uk>
@JeremyTandy

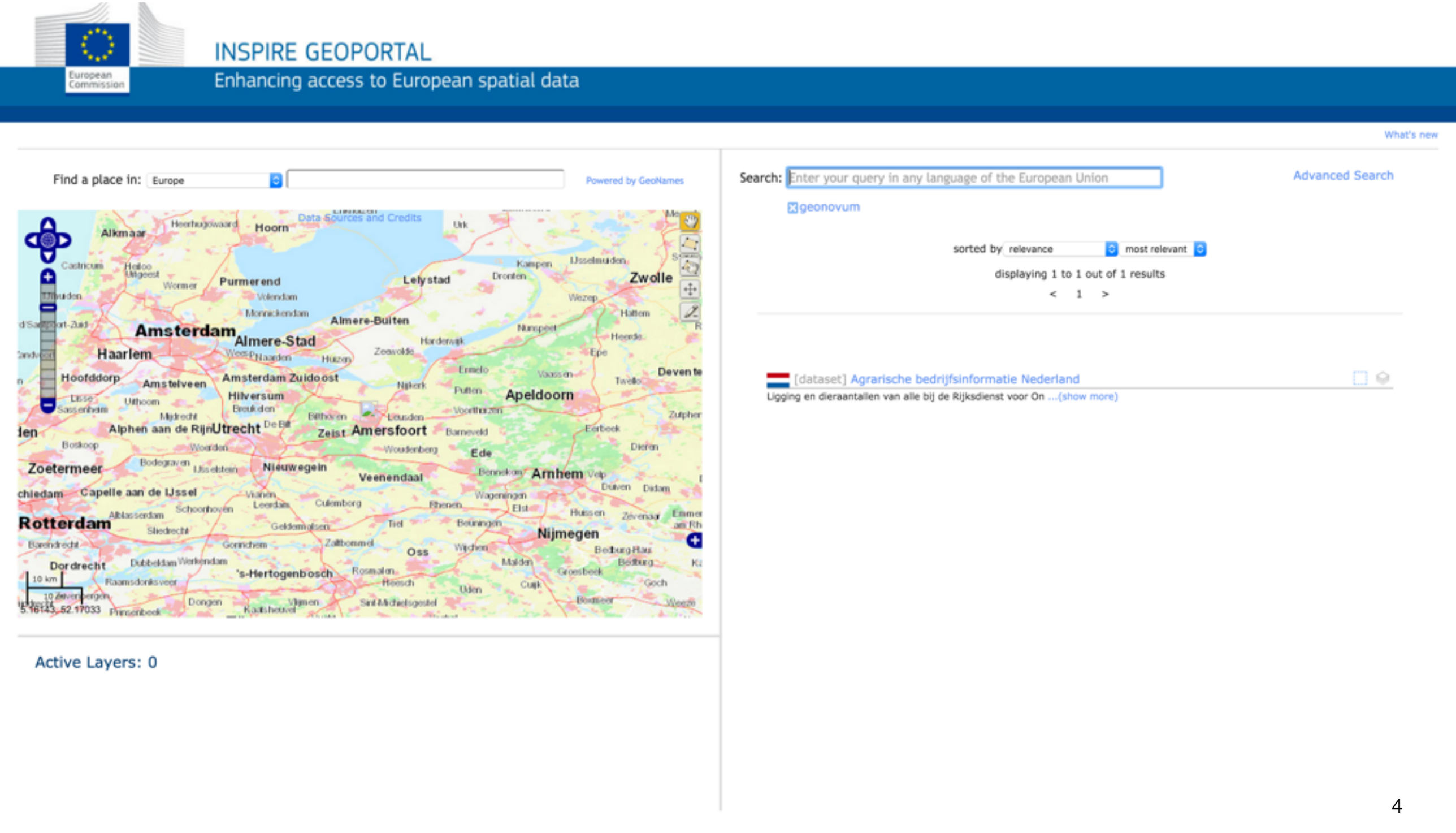
Ed Parsons
<eparsons@google.com>
@edparsons

What’s the problem?

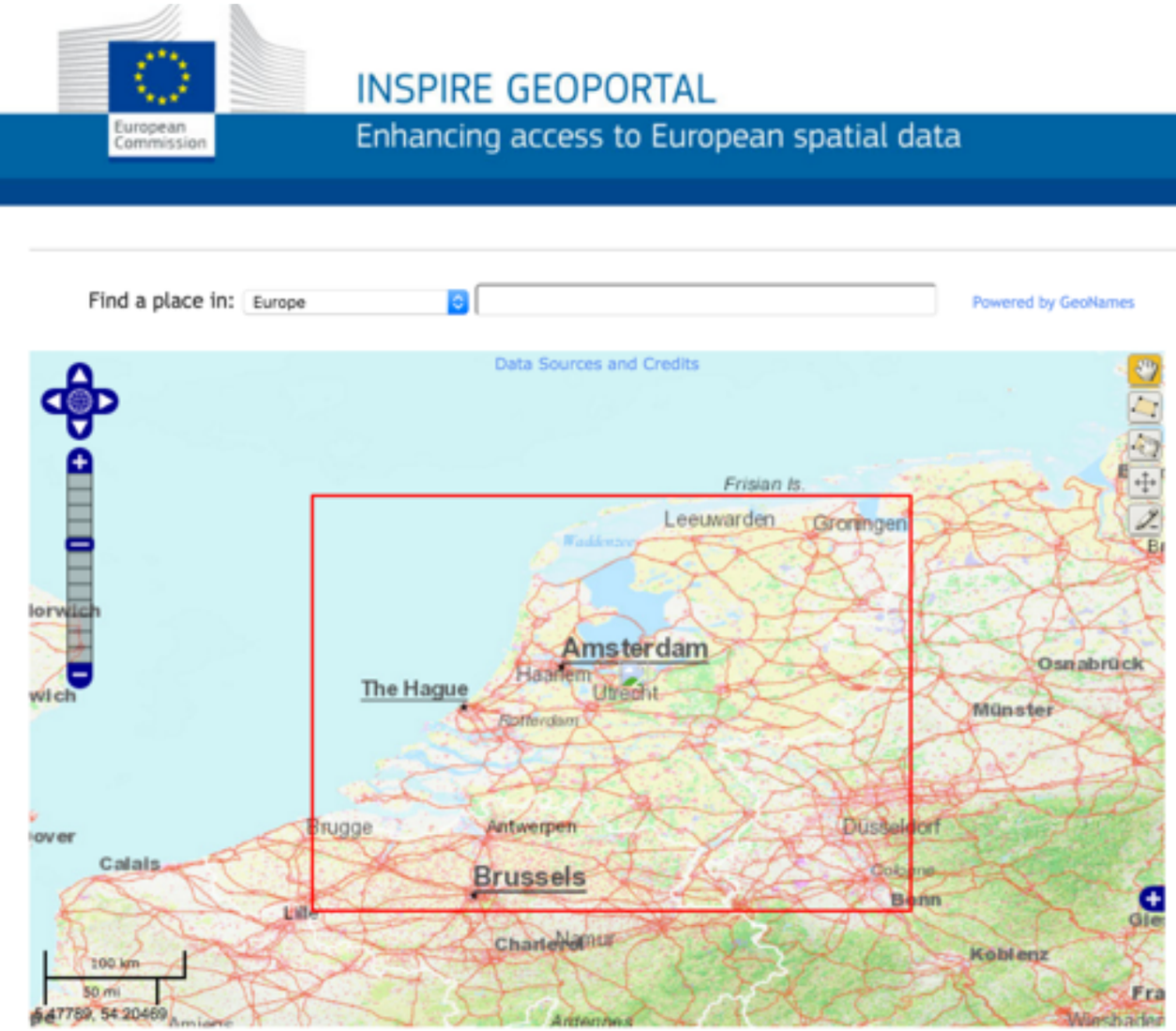
What we expect

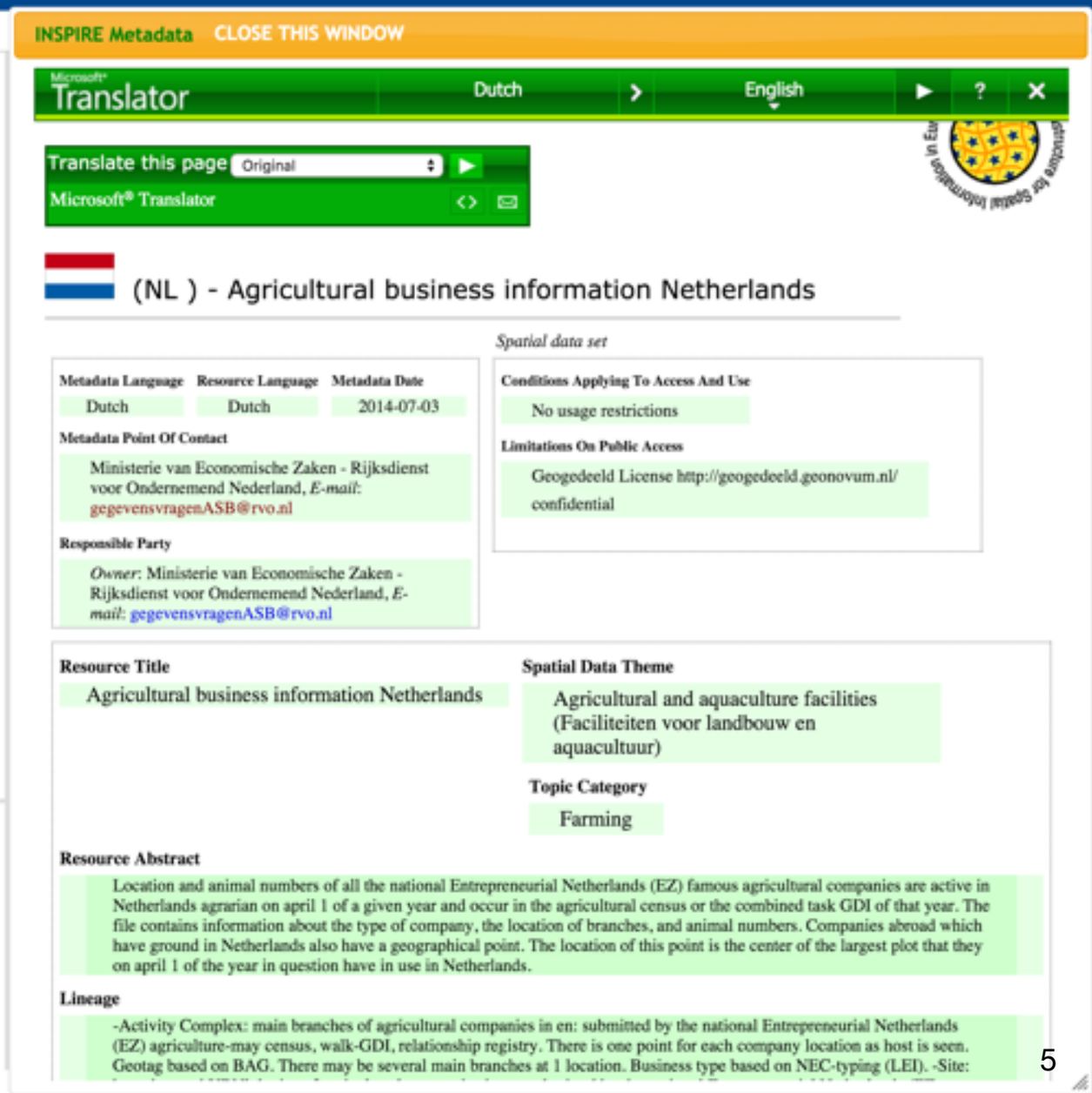


What we get

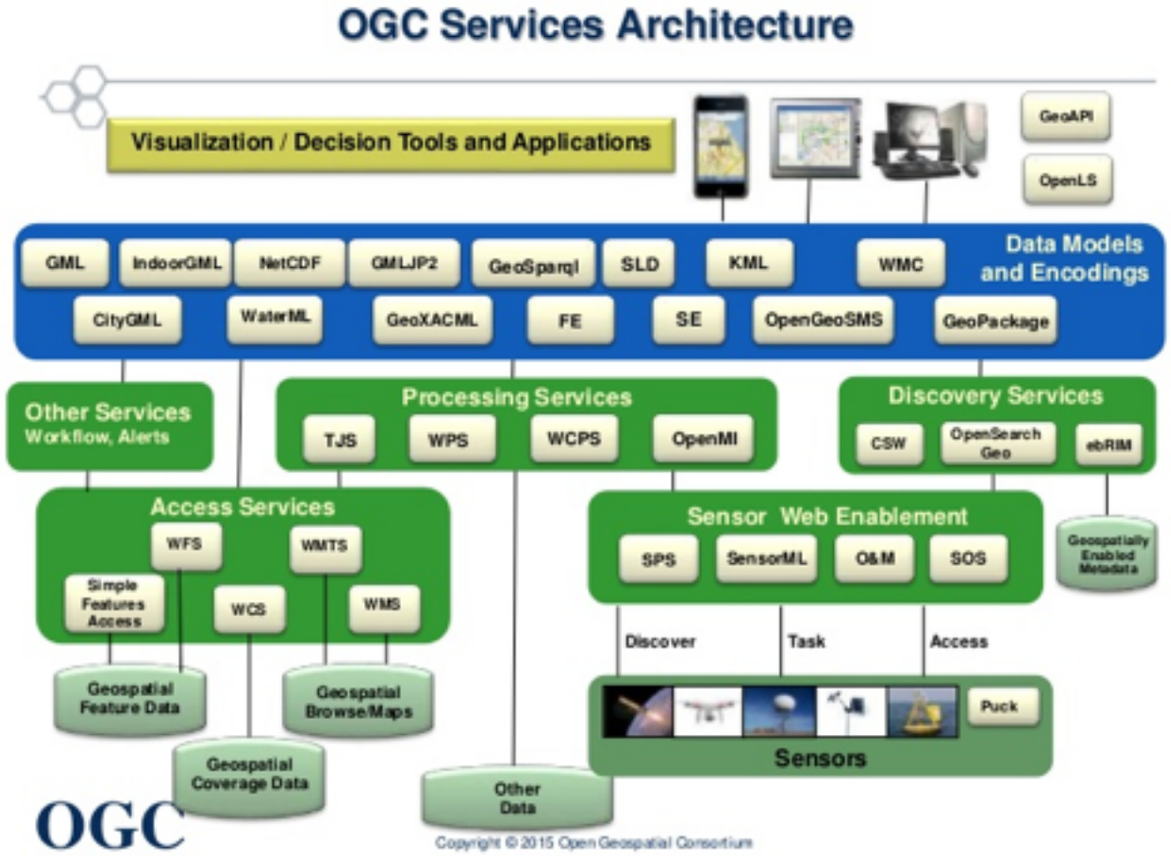

European Commission
INSPIRE GEOPORTAL
Enhancing access to European spatial data
What's new
Find a place in: Europe
Powered by GeoNames
Data Sources and Credits
Search: Enter your query in any language of the European Union
Advanced Search
geonovum
sorted by relevance most relevant
displaying 1 to 1 out of 1 results
< 1 >
[dataset] Agrarische bedrijfsinformatie Nederland
Ligging en dieraantallen van alle bij de Rijksdienst voor On ...(show more)
Active Layers: 0
Alkmaar
Hoorn
Purmerend
Lelystad
Zwolle
Amsterdam
Almere-Stad
Almere-Buiten
Haarlem
Amstelveen
Amsterdam Zuidoost
Hilversum
Apeldoorn
Utrecht
Amersfoort
Zeist
Nieuwegein
Veenendaal
Ede
Arnhem
Nijmegen
Zoetermeer
Capelle aan de IJssel
Rotterdam
Dordrecht
's-Hertogenbosch
Oss
Alphen aan de Rijn

Two step search; unfamiliar protocols

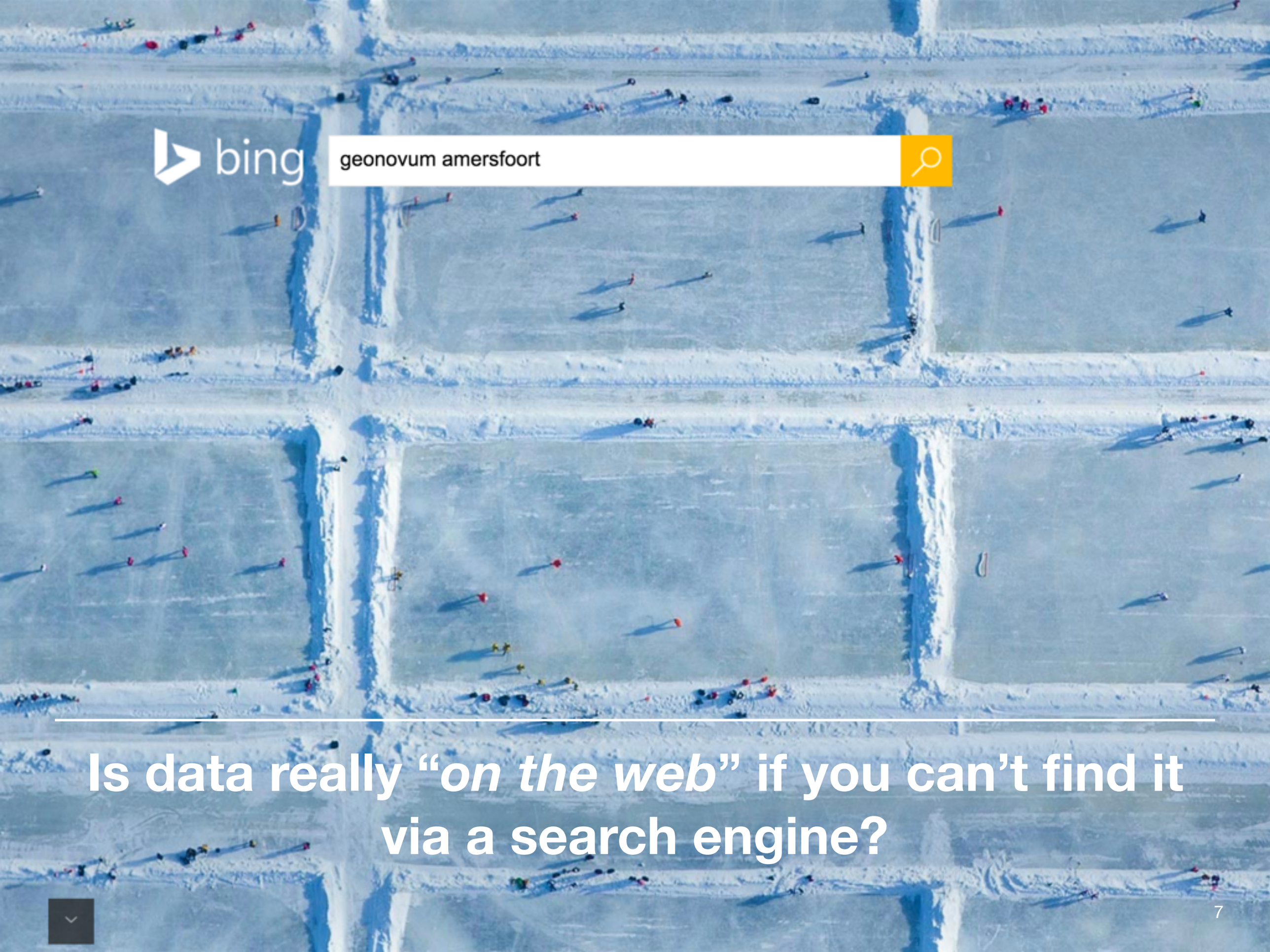




The geospatial ~~solution~~ niche



- Geospatial industry has developed it's own web services to publish Location information
- Dominated by large Government data publishers and Enterprise customers
- "Fit for purpose" for a niche industry?



Is data really "*on the web*" if you can't find it via a search engine?

So, we need to be more webby?

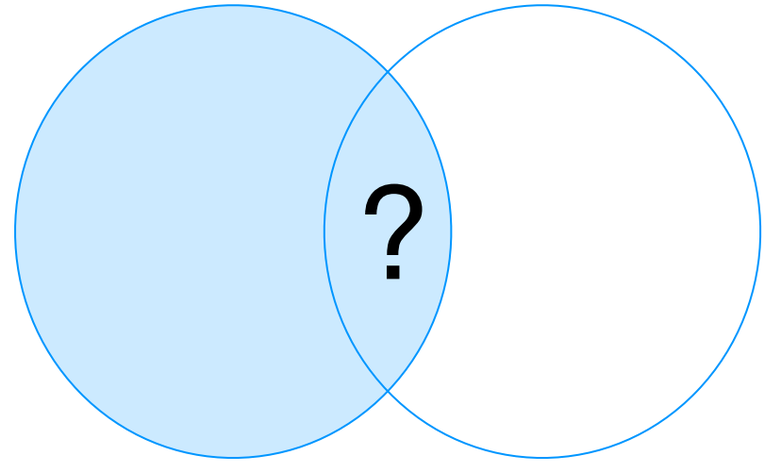
The reality? implicit and unstructured

- Most web content about places and location is unstructured
- Harvesting requires sophisticated NLP and inference
- Does not scale

```
<h3>Visiting Address</h3>
<p>Netherland Institute for Sound and Vision<br />Media
Parkboulevard 1<br />1217 WE Hilversum</p>
<p>If your navigation system does not recognize our
(new) adress, you can find us on Sumatralaan 45, 1217
GP Hilversum.</p>
<h2>Postal Address</h2>
<p>Nederlands Instituut voor Beeld en Geluid Media
Park, <br />Postbus 1060<br />1200 BB Hilversum</p>
<h2>Opening hours</h2>
<ul><li>The experience is open Tuesday - Sunday, 10.00
- 17.30h</li>
<li>Closed on monday</li>
</ul><h2>Entrance prices</h2>
<ul><li>Adults: € 16,00</li>
<li>Children between the ages of 4 and 12:  € 9,00</li>
<li>Children up to age 3: Free</li>
```

Did someone say Linked [Open] Data?

A problem shared …



W3C OGC **Linking Geospatial Data**, Mar 2014, London

Miss Globe and Mister Cube …

W3C OGC **Linking Geospatial Data**, Mar 2014, London

60 second summary: key questions

- how should we encode geometry?
- how and where should we implement topological functions?
- additional metadata is required for spatial datasets – how do we do that?
- where is the software support for spatial datatypes and functions?
- geometries expressed as WKT literals are large objects — the Linked data world is used to handling simple literals
- how do we help developers handle (or avoid) the steep learning curve to work with Linked Data?

https://www.w3.org/2014/03/lgd/report

W3C OGC **Linking Geospatial Data**, Mar 2014, London

Spatial Data on the Web working group

http://www.w3.org/2015/spatial

SDW working group charter

- to determine how spatial information can best be integrated with other data on the Web;
- to determine how machines and people can discover that different facts in different datasets relate to the same place, especially when 'place' is expressed in different ways and at different levels of granularity;
- to identify and assess existing methods and tools and then create a set of best practices for their use;
- to complete the standardization of informal technologies already in widespread use.

http://www.w3.org/2014/05/geo-charter

> *Where relevant, it will promote Linked Data using the 5 Stars of Linked Data paradigm, but this will not be to the exclusion of other technologies*

Best practices deliverable

- 5-star data with a spatial rosette?
- [best] practice not theory

Editors: Jeremy Tandy, Payam Barnaghi, Linda van den Brink

Editors draft: `http://w3c.github.io/sdw/bp`

We’re not alone …

W3C *Data on the Web Best Practices*

http://www.w3.org/TR/dwbp/

What’s the current status?

First Public Working Draft (FPWD)

- 30 best practices organised into 5 themes & underpinned by dozens of use cases …

Assigning identifiers | Expressing spatial data | Linking spatial data

Enabling discovery | Exposing datasets through web services

Best Practice 1: Use globally unique HTTP identifiers for entity-level resources
Best Practice 2: Reuse existing (authoritative) identifiers when available
Best Practice 3: Convert or map dataset-scoped identifiers to URIs
Best Practice 4: Provide stable identifiers for Things (resources) that change over time
Best Practice 5: Provide identifiers for parts of larger information resources
Best Practice 6: Provide a minimum set of information for your intended application
Best Practice 7: How to describe geometry
Best Practice 8: Specify Coordinate Reference System for high-precision applications
Best Practice 9: How to describe relative positions
Best Practice 10: How to describe positional (in)accuracy
Best Practice 11: How to describe properties that change over time
Best Practice 12: Use spatial semantics for Spatial Things
Best Practice 13: Assert known relationships
Best Practice 14: Provide context required to interpret observation data values
Best Practice 15: Describe sensor data processing workflows
Best Practice 16: Relate observation data to the real world
Best Practice 17: How to work with crowd-sourced observations
Best Practice 18: How to publish (and consume) sensor data streams
Best Practice 19: Make your entity-level links visible on the web
Best Practice 20: Provide meaningful links
Best Practice 21: Link to spatial Things
Best Practice 22: Link to resources with well-known or authoritative identifiers
Best Practice 23: Link to related resources
Best Practice 24: Use links to find related data
Best Practice 25: Make your entity-level data indexable by search engines
Best Practice 26: Include spatial information in dataset metadata
Best Practice 27: Publish data at the granularity you can support
Best Practice 28: Expose entity-level data through 'convenience APIs'
Best Practice 29: APIs should be self-describing
Best Practice 30: Include search capability in your data access API

FPWD: http://www.w3.org/TR/sdw-bp/

“

”

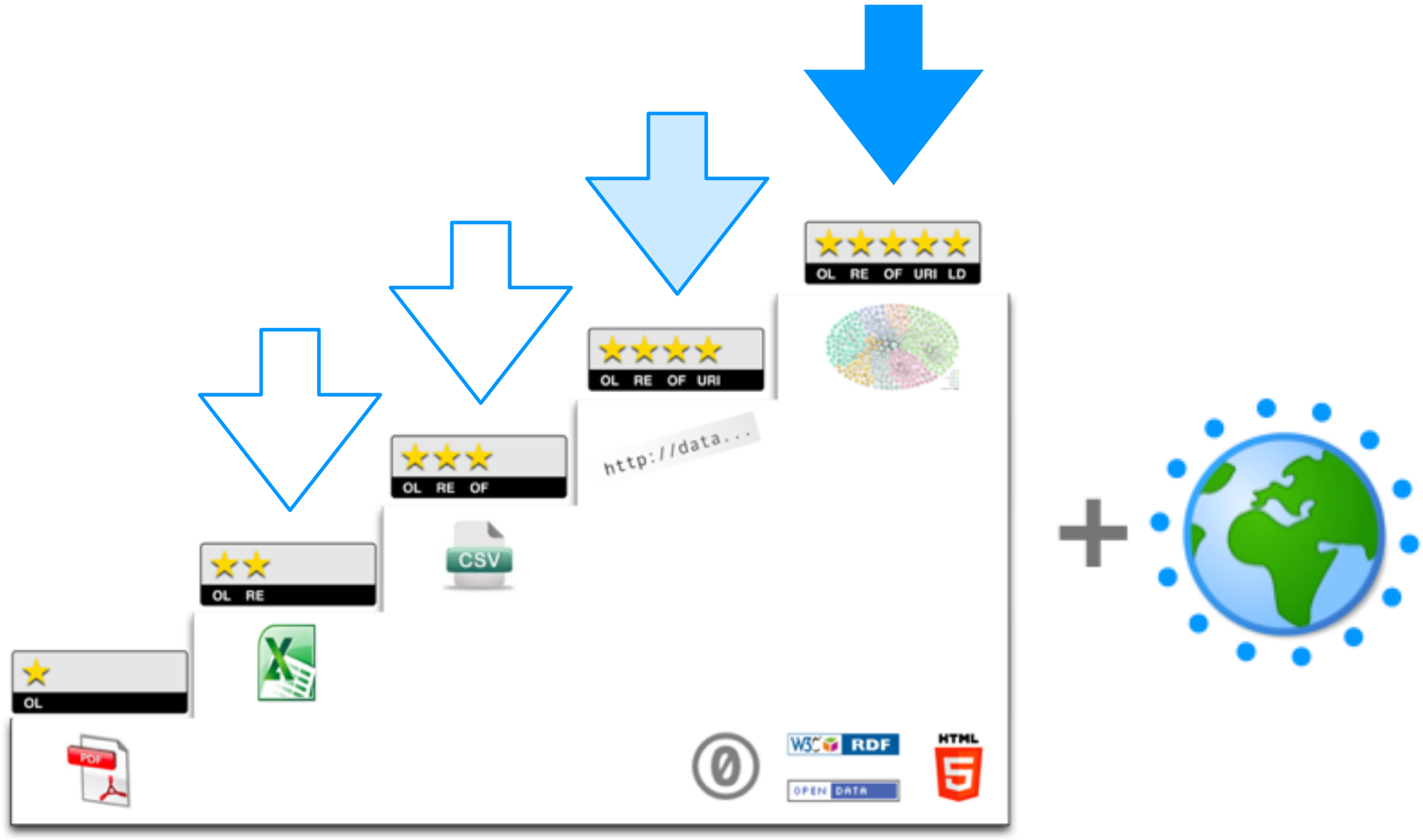
<u>Best Practice 1</u>: Use globally unique HTTP identifiers for entity-level resources

Who are we trying to help?

- Everyone - because spatial data is really useful! …

- But we really need to focus on the gatekeepers:
  - Web developers
  - Spatial data custodians & publishers

Is this for '*Linked-Dataverts*' only?



Too focused on the <http://dbpedia.org/resource/holygrail>*?

source: http://5stardata.info/

* of course, this doesn't exist in dbpedia :-)

And finally …

We need you!

- help us provide something that's useful for you …
- review the working drafts;
- participate in discussion;
- …
- maybe even join the working group.

- specific requests:
    1. what [software] architectures are you using in your spatial applications?
    2. which formats do you typically publish data in today?
    3. what kind of spatial data are you trying to publish? 1d, 2d, 3d etc.
    4. what is preventing you from using spatial data in your applications / decision making?
    5. help us structure the BP doc to suit a short attention span
    6. tell us about examples of implementation from the real world
    7. tell us what we are missing
    8. tell us what you want from this

Questions please …

mailto:public-sdw-comments@w3.org

Jeremy Tandy
<jeremy.tandy@metoffice.gov.uk>
@JeremyTandy

Ed Parsons
<eparsons@google.com>
@edparsons