

Linked Open Data hackathon at Naturalis (2016)

Participants:

Jörg Holetschek (BGBM)

Ayco Holleman (Naturalis)

Dominik Röpert (BGBM)

Ben Scott (NHM)

Rutger Vos (Naturalis)

It wasn't really about Linked Open Data.

Primary goal was to get to know each other's data APIs.

- Which web-accessible APIs do the other parties have?
- How do I specify a query?
- How do I parse the result?

So why did we still call it an “LOD pilot”?

No LOD without a mature, reliable, flexible, queryable and web-accessible lower-level API.

We should get in the habit of **actually** querying and linking to each other's data.

We hope this will create a sense of urgency and a dynamic towards ever more formalism in getting the correct data: **LOD**

A bottom-up approach to achieving the formalism of Linked Open Data.

Being able to retrieve each other's specimens through a GeoName identifier is great.

Being at least able to do so via a verbatim locality is not bad.

Seeing “bad data” come back as a consequence of ambiguous, pre-LOD identifiers is a great incentive for implementing LOD techniques and for shared disambiguation techniques.

Conclusions from the hackathon

Both NHM's and BGBM's and Naturalis's data APIs were still very nascent at the time.

Except for specimen PURLs, accessing each other's data through other types of identifiers (e.g. GeoName or VIAF identifiers) still seemed a rather remote option.

(For one thing: we would first have to disambiguate our own data and enrich it with GeoName c.q. VIAF identifiers)

Conclusions from the hackathon

Ambiguity is only eliminated if all systems use the same stable identifier from the same concept authority.

Thus it became clear that we need to mutually agree which concept authorities to use and enrich our own data sets with stable identifiers provided by them.

Conclusions from the hackathon

For optimal linkability we need to develop shared disambiguation tools so that the same source data resolves to the same identifiers for all institutes.