Geodata on the web workshop – February 11th 2016 - Amsterdam

**Wrap up session minutes**

Pasquale Di Donato : Swisstopo

Interest coming from customers (federal offices asking if Swisstopo could publish its data as linked data so that they connect their data to them) but no clue, struggling with identifying the right implementation to experiment, the right vocabularies and CRS.

Opportunity : Getting more visibility on the Web.

Flemming Nissen, Danish Agency ‘Data supply’

We feel unsecure, one possibility is to explore but we are not certain about the need. If we put services based on linked data technology we are supposed to keep it in the air; We also need to keep ISO/OGC for existing customers and sense we need to start addressing new group of customers that is more web-enabled, they seem not to know what infrastructure is beyond (how much it cost actually).

Dimitris Kotzinos, U Cergy Pontoise

Linked data does not necessarily mean to put up triple stores. Important concerns are more: URI and how you publish resources but the technical implementation comes after. It is beneficial to differentiate when you try to solve the problem. His advice would be to start thinking about what data you want to connect to.

Lars Hagge, Lantmateriat, Sweden

3 weeks ago we had a workshop about how to put up a URI. But we don’t know the cost to maintain it. We say we need to combine data from different authorities but we can already do it. Do we need to facilitate others to do it themselves? Each time people come to talk linked data, they talk about triples and when we ask them why they just answer ‘because we need it’ but they don’t relate it to higher level needs. Where is the use case? Where is the money?

Arjen Santema, Kadaaster, Nederlands

One pb for Kadaster was to publish the data in a way users could know where they come from directly from the source.

We want to reach much more potential users who come from a completely different community and they need to understand the content.

We also want to have a larger community, like OSM, make data more open and linked. We want to be more open.

Peter Lentjen, Kadaaster, Nederlands

A first step was to distribute the data without charging the user. It is much more used. But there is still a lack of knowledge to use it more. We want to create an opportunity so that it is more easy for new customers.

B Bucher

Something specifically difficult with geo APIs is that the standard change all the time (e.g. : gml versions).

Erwin Folmer

There is an API learning curve; GI is a niche and web developers don’t want to learn something specific for this.

Dimitru Roman, Sintef, Norway

Have test to work with existing geodata from NMCA and can testify that it is difficult to work with these. It is so complicate that in many case the tradeoff between putting efforts to get NMA quality or using more simple data like OSM with maybe less quality is easy to do.

Use case1 : evaluation of real estate properties,

Use case2 : reporting of public properties. Every 4 years.

For both use case, OSM data do the job and are easier to use.

Raphaelle Lapotre, BNF, France

We need a mediation, a place on the web where we can explain to developers the nature of the data we distribute. There are places like github for libraries and we need the same for the data where we explain how to use the data. Because open data is not only free data.

Frans Knibbe, Geodan

Also think that focusing on metadata and semantics is a good idea. This need will stay.

There is the pb of how to deal for big changes. The way to do is to deal with it step by step. And the first step is semantics and metadata. See geodcat-AP.

Qn : for machines or for human?

Opportunity : logging what happen to the data (logs of HTTP servers), consolidating feed backs.

Joan Olive, Karteverket

OGC standards are complicate. What data should be considered? All data or just metadata?

P Lentjen : can use both? When you want to use data together you have to copy everything and store. Can new technology (like linked data) lead to more performant distributed architecture?

Dimitru Roman

You need to put simple REST API on data. We have developed datagraph platform which is a kind of github for data precisely, for sharing the transformations (happening to data) and hiding the complexity.

R Lapotre

Interesting to see how problems with geodata are common with data of librarians even if the data are different. We need to reconcile datasets who talk about the same object, we need to cluster data, aggregate and so on.

Another interesting thing about geodata is that you always include visualization.

N Abadie

When we tried to publish geodata on the web we spent much time publishing the metadata and the conclusion was that the most important part was the metadata publication because these metadata were stored in pdf files less accessible; besides, this publication as linked data allows to provide metainformation at a more fine grained level than the dataset.

F Knibbe

Again one piece at a time. Metadata about coordinates are relevant and it will improve the quality of the metadata to distribute them.

E Folmer

Opportunities : google search , community platform, relation with INSPIRE is also important. It cost us a lot of money. ? how to assess the benefice : obvious, google ranking, avoiding redundancy (why somebody wants to search in a national registry whereas you have google)

P Lentjes

The important thing to do now is : let’us explore and see how far we can go.

Peter Hallahan, OS Ireland

What we want is to make sure the customer get the best information. Metadata management should be at the object level (e.g. about the currency of the data).

We need an improved solution to encode and distribute multiple representation data. Also explaining fitness for use is an issue.

R Lapôtre :

to ease metadata production during production process so that it is easier to curate them after.

Lars Hägg

Are these data cross boarder?

Researchers are using OSM data.

Our models are complicate and the problem is rather on the semantics part : let us define unite cross boarder data. Eg : for buildings.

(Dimitri Kotzinos) : you don’t need to solve all conflicts before linking.

(N Abadie) : this point relates to multiple representation issue mentioned before, especially wrt geometry representation

Peter Melin, Landmateriat, Sweden

Like the idea to use a simple restful API

Using WFS et WMS is not so difficult its is more a question of incentive than of learning curve.

Heidi Vanparys

Agency for data supply and efficiency, Denmark

It is a question of teaching material

Pasquale Di Donato

And a question of visibility

D Kotzinos

It is important to start from the problems : what problems do we have? How to solve them.

Think of problems where SW techno can help: improving openness, machine readability, linking to non GI sources. Relate to public open data. Raise the number of applications that use our data.

Than separate the pbs : what kind of data do you want to link and at what level, semantic conflicts, web publication, discoverability,

Eero Hietanen

Finland has engaged in publishing topo data sets with URIs. This is a 2 years pilot. We look for best practices, linked data is a small part of it. We need to know how to share ontology and link them to existing ones. Which data model to use?

Pasquale Di Donata

I am looking for implementation guide lines, and tutorials.

Dimitru Roman

Technology now is mature. But Linked data is really complex and we need to look how to simplify stuff. Make these things more easy to use. We also need reliable infrastructure.

Flemming Nissen, Danish Agency ‘Data supply’

We did education gimmic, Danemark Minecraft for children to get a sense of what is coming up. It was a success and we put it down after one year because we did not think it was supposed to be maintained. Do we have to be a testbed also for linked data and put it down after some months. As a gvt agency we have some decision making problems because we need to have explainable use cases for decision makers to risk it.

Heidi Vanparys

Agency for data supply and efficiency, Denmark

Can we use these principles to make it easier for us to maintain our data?

Pb : how to convince people to document data

We also have an issue about cross boarder, maybe there is no obvious interest but somehow we can guess it is a stake. Lars : yes, the only user so far is EU.

Abdefettah Feliachi

Important to think of opportunity and added value of linked data –not to replace what already exist but to add value to it-. To open to new communities, to link to other (community) data. The possibility to exist in the big cloud will give more value to our data.

Maybe techno is mature but the infrastructure is not that mature. Lot of works must be done on that part.

B Bucher

SW is supposed to provide interesting results for discovery, recommendation (linking data with users and usages, storing data about users that will be used to make recommandations by clustering usages) - And GI specific field provides interesting results about visualisation, quality management (gap with reality and what can be assumed about reality based on the data)

GI people tend to think GI is everywhere and sw is a niche, whereas SW people think GI is a niche. Semantic web has evolved: TimBernersLee first wanted to describe on the web what exist, but now it is more to describe what exist on the web (probably it is a preliminary step).

Sometimes difficult to remember concrete impacts we want/can have.

Maybe SW people know more about successful economic models in digital world. When we think of impacts, the is a question of priority –must have- Avoid flood, getting crowd to contribute, etc