

# DMworkshop – Notes from break-out sessions

---

Each of the four sessions of the workshop consisted of a number of presentations and a breakout session, each with three parallel breakouts.

Each breakout discussed amongst the participants some of the topics brought up during the presentations.

This document is a collection of notes from all breakouts.

The notes summarize discussions and views expressed by the individual participants in each breakout and do as such point out important topics within the areas discussed. However, the notes cannot be seen as official statements of the workshop.

## Linked data (A1)

*Documented by: Esa Tiainen*

What's the linked data approach equivalent to linear referencing? Is there a similarity to the approach used for time-series?

- see W3C <http://www.w3.org/TR/vocab-data-cube/> & best practices at <http://www.w3.org/2011/gld/> (old) & <http://www.w3.org/2015/spatial/charter> (new)
- Similar question in BIM – levels of detail / stages of development

Linked data as a gradual introduction

Identifier management (same problems in INSPIRE):

- How do you manage/coordinate authoritative publication? (same problem in INSPIRE)
- Whose URI should you reference? And/or who owns the id?

Who's responsible for the validity of data (fitness for purpose). It used to be the data producer, but (e.g. with open data) the consumer decides. What's the linked data answer to provenance/traceability?

- Finland has implicit & is planning to implement explicit links from HTTP URIs to metadata records in their service

Vocabularies – big problem, linked data brings one extra tool

- Linked data ideology is opposite to central control
- Who would maintain any (SKOS) super-ontology, which shows how the others relate? E.g. FAO, EC (several), WMO,
- (Perhaps) work towards national ones, initially within domains & between neighbouring domains.

Tool support

- Client side
- Performant search service: 'Elastic search' (similar to SOLR?) seems quite quick
- RAGLD – for creating web applications

## Data modelling cookbooks for SDIs (A2)

*Documented by: Knut Jetlund*

What do we need in cookbooks?

What do we particularly have difficulties with?

Is a cookbook a book? In this setting: Also a wiki, and a series of blog posts.

Adapted for different levels of detail.

Paul (NL): Positive that there are cookbooks, but difficult to find. Good that ISO is doing a work. Good that it is in a wiki - evolving. Not locked in a PDF. Should involve other communities as well, not just the geospatial communities. BP Wiki a way of communicating "I just don't know if what I'm doing is correct!" Should be on best practices, not an alternative on UML modelling literature. How to present information to stakeholders: Other ways of presentation: Tables (feature catalogue) Starting with simple overview models, moving into more complex models.

Debbie (UK): What we need: "How to's" Tools are important for using the models. Developing Plugins to EA.

Jef Vanbockruick (FGIA): Training is essential - important with a reference from ISO Tutorials Some thumbrules - Use roles/stakeholders - analyst, domain expert, etc to teach different things.

Heidi (DK): We need to learn from the IT community. How do we use UML different from the IT community. IT: For sketching. What can we reuse, what is special?

... (DK): If you do it like this, it might have these consequences, if you do it like this....

Gnäggi (CH): Should start by describing why we create models. Use cases.

Johannes (GmbH): Working with profiles - narrowing in Many ways to extend a model - hard to say what is wrong and not. Present solutions and pro/cons. Understanding metamodels - complicates even more.

Tatjana Kutzner (DE): Break it down to very small sections, for instance for teaching

... (FR): Is it possible to have the same model for implementation and for human communication?

Markus (DE): How to work with code lists etc in practice?

## Data Modelling and Implementation and documentation Cookbook (A3)

*Documented by: Morten Borrebæk*

**User perspective:** Who are the users, there are users at many levels, not only end users.

- usability
- background for policy decisions
- How to present

**Technicalities:**

- Model aware update process (how to keep the data model updated.)
- What is the process?
- How to manage extensions. (INSPIRE extension)
- Some model components are complex (large aggregates, but use just part of it.)
- Encodings. What is actually the intention of the INSPIRE directive – download data, but needs tools (like FME) to transform this to a format that GIS tools can use.

- Tools to handle INSPIRE data was expected to pop up, some actions, but less than expected. Why? For S57, vendors needed assurance that the spec should not change. Challenge for INSPIRE.
- Software innovations - More and smaller models, easier to implement, or larger model covering more use cases. Needs both.
- Associations, particularly bi-direction, used different in in INSPIRE. Needs clarifications. Requires more guidelines on direction of association
- Ensure desilofication.

## Tools support for modelling and model driven implementation of data distribution services (B1)

*Documented by: Stephan Mäs*

The methods and tools for model driven implementation of data distribution are under rapid development and some of these tool have been discussed in the presentations. But are the available tools sufficiently supporting the complex workflows? What are your experiences using standard tools and / or implementing own solutions? What is missing or causes problems?

- **maintenance of data models**
  - consequences on the processes, schema
  - Versioning → tracking of versions, influences on transformations
  - differing version concepts: release vs. revision, logical vs. physical versions,
  - documentation of model changes,
  - update scripts / evolution scripts → SQL
- **INSPIRE Extensions:**
  - management of extensions, client support
  - consistency among the model components
  - replacements
- **consistency of models**
  - imperfect models, in terms of performance
  - scope of models
- **work with code lists and handling of updates**
  - bigger codelists
  - national extensions
  - interaction with registries
- **integrity constraints**
  - formalization → interpretation → validation
  - OCL extensions?
  - templates for common constraints

## INSPIRE compliance without GML; usability of INSPIRE data (B2)

*Documented by: Peter Parslow*

- If a data provider claims to satisfy INSPIRE regulation other than using the technical guidance, the onus is on them to prove it
  - Could perhaps be done by groups of providers

- Extend the question to cover service requirements
- INSPIRE & (something else in transport domain?) aren't working together – dynamic, online data
  - Dynamic data example: algae situation. There is a service, how to make it INSPIRE compliant?
- Different INSPIRE themes use different formats, relevant to the thematic needs
- Some themes have several use cases (others should have) – the different use cases would benefit from different encodings
- Tools for validating / proving whether you comply would vary from theme to theme, encoding to encoding
- MIG task 5 is creating 'abstract test suite' for each theme
- Much of INSPIRE Annex I (& some Annex II) **could** use Simple Feature GML
- INSPIRE is intended to support a range of use cases (including unknown – open data – ones) – hence the need to provide explicit semantics
- Complex (GML) structures are expensive to ingest
- Disappointed in lack of development on client side, to ingest complex data
- Fund open source development e.g. in QGIS?
- Complexity is sometimes necessary
- Complexity isn't an aspect of GML, but of the data model
- But GML/XML is verbose – large data volume (service & client support for maxFeatures) compression built in to HTTP & much software
- GeoJSON for some user communities

## User Interface for ShapeChange (B3)

*Documented by: Clemens Portele*

- Requirements
  - Non-Expert
    - Only control parameters like file locations, file types, application schema selection
  - (Irregular) Expert
    - "Full" control, but advanced aspects are out-of-scope for now, e.g. transformation workflows
    - Stereotype management
    - Namespace management
    - Map entry management
    - Encoding rule definition
  - Both
    - Include context-sensitive help and examples, links to shapechange.net
      - potentially adding additional explanations of options
    - Automatically adapting to ShapeChange capabilities (targets and their parameters)
    - Should be part of ShapeChange to simplify deployment
- Establish ShapeChange community to move activities like these forward

- interest by all participating in the breakout to contribute
- Clemens P to initiate and get in contact
- Need for finer grained building blocks of models as input?
  - currently 1..n (application) schemas
  - maybe also packages in schemas or a set of classifiers?
  - requires further discussion of use case
- SBVR has a lot of potential (OCL seen as very complex to learn)

## Modelling and use cases (C1)

*Documented by: Ziggy Vanlighthout*

### Modelling of business use cases

- How to link user requirements to UML data models?
  - ArchiMate for business use cases / user requirements?
  - UDL for data models?
  - Be aware of EIF (European Interoperability Framework) and EIA (European Interoperability Architecture) initiatives
- Who are the stakeholders
  - Differs between closed data and open data scenarios
  - Government in all cases - everybody in case of open access
  - Can/should government benefits carry all costs – and let everybody benefit
- What is the business of government/public administration?
- Up to what level of granularity should business use cases be modelled?
- You can't know what business use cases will emerge after you make your data available.
- Go for flexibility in the use of data!
- Don't confuse "flexibility" with "openness".

### Data models for data management vs. for data distribution

- Keep two different models
  - One for distribution – de-normalized – based on user requirements from business use cases
  - One for management – normalized – being in control of data
- You need metadata
  - Both on data management level and on product (data distribution) level
  - Managed data must be precisely described
- Using same modelling technology/language for both is an interesting approach

## Association classes, multiplicities, etc. (C2)

*Documented by: Jef Vanbockryck*

- Is the use of « association classes » for modelling core relations (extension of the association class via generalization) a way forward?
  - An area not widely explored
  - To be taken into the TC as a practice to be explored

- Feedback from Flanders into TC
- Should a data model for « management » of basic data have strong multiplicities? Should they be more loose and be managed via constraints/rules?
  - Have a model for each purpose with the exact multiplicity for that purpose
  - Model elements are re-used, only those with different multiplicities are copied and can be loosely referred (e.g. via dependency association) and repository structure (EA) can help

## ID management and linking (C3)

*Documented by: Clemens Portele*

- **Use Cases for persistent IDs**
  - Distinguish objects
  - Traceability
  - (Explicit) linking
  - Lifecycle management of objects, incremental updates
- **ID Management Issues**
  - HTTP URIs vs Namespace + Local ID vs UUID
  - Dereferenceable
  - Register(s) needed?
  - Use of Well-known codes (NUTS etc.) for linking
  - Persistency
    - IDs for derived data, eg in INSPIRE transformations
    - Depends on life-cycle rules
  - Uniqueness
  - Object versions
  - Governance
  - Multiple "external" IDs
  - Multiple copies
  - Real world thing vs. feature identifiers
  - Support for linking in GIS systems
- **Current state**
  - In most cases using local identifier, typically with some namespace
  - Partial move to http URIs, but usually not yet dereferenceable
  - WFS can use GetFeatureById-URLs, but not persistent
  - Existing practice on http URIs exists in several countries, usually work-in-progress
  - Recommendations and explanatory material exist in INSPIRE
    - but may need more work
  - Activity in INSPIRE MIG, but low priority at the moment

## **Dynamic versus static Modelling. Information models shape the world or vica versa? (D1)**

Documented by: Jandirk Bulens/Paul Janssen

About:

- 1) Dynamic versus static Modelling. Information models shape the world or vica versa?
  - Flexibility, dynamics
  - Versioning
  - Runtime extending mechanism
  - New modelling paradigms - linked data?

Discussion:

Dynamics in modelling at different levels:

- 1) Modelling of dynamic processes: Time series - Simulations -> modelling patterns to design processes.
- 2) Static models adapting to changing requirements

Ad 1) Modelling of dynamic processes depends on use cases, is application related, i.e. Run time execution. It is less a concern for semantic information modelling as we deal with in this workshop.

Ad2) Adapting semantic model through time.

- Modelling approach results in versioned application schema's. These provide 'rigid' application environments. - update cycles are long
- usage adaptability is low
  
- Run time extending mechanisms sounds promising.
  
- Core models with specific application extension provide an flexible maintenance approach.
  
- ADE, application domain extension is a described and reusable extension pattern helping flexibility in modelling.
  
- UML Domain models are predominantly top down. Object orientation provides rigid views through feature model
  
- Linked data (OWL/RDF) provides open structure for adaptability by subject - predicate - object relation.
  
- Both can be complementary and applied in different areas of an SDI

## **Use and adaptability of the INSPIRE Network Model (D2)**

Documented by: Paul Scarponcini

- **Network topology**
- **Linear referencing**

Discussion of some relevant ongoing projects:

1) EULF Pilot Project - location framework based on INSPIRE Model

2) ROSATTE - Dynamic Location Services for dynamic notification, such as for accident occurrences; links to EULF

3) Denmark - need to integrate multiple Linear Referencing Method (LRM) / Linear Element (LE) combinations from various business units. Developing a hub base model consistent with INSPIRE from and to which each business unit can translate, instead of trying to develop individual translations between each pair of business units. Easier from a technical as well as governance standpoint.

Observation:

Linear Referencing (LR) may be new to GIS and perceived as being hard to do, because many GIS tools do not support it. However, in Transportation, it is required, is not hard to do, and software tools there support it.

Key findings:

1) INSPIRE only supports a single LRM (chainage) along a single type of LE (link geometry). This is inadequate as many people instead use relative or interpolative LRMs. Recommendation is to considering updating INSPIRE to better support ISO 19148.

2) The INSPIRE Road Transport Network Model apparently further restricts 19148 use by making all attributes events, even if they have a single value throughout the entire length of the LE. Several people stated that they define the LE boundaries such that a minimal set of attributes are homogeneous in value throughout this length. INSPIRE needs to allow for some attributes to be at the LE level, rather than being attribute events, obviating the need for LR for these attributes.

### **3D and information modelling serious gaming (D3)**

*Documented by: Jan Hjelmager*

- **Business use cases for 3D include both visualization and analysis**
  - Visualization and visual simulation for Defence
  - Noise monitoring
  - Pedestrian routing
  - Solar panels
  - Truck elevation change along the route
  - Underground cases
  - Weather service ...?
  - ...
- **Quality issues with respect to costs**

- Data representations - there are different ways to represent 3d (3d and 2.5d)
- What detail can you capture for a given costs?
- Different data representation needed?
- New quality assurance procedure needed?
- **Practical issues at the current software technology level to process 3d data / information.**
  - How to manage data 3d data using/together with the current solutions
- **National requirements for 3d model**
  - Buildings offered in Netherlands and UK
  - 3d cadaster?
  - Which LOD to use? Do we need some new ones?
  - Could NMCA provide generic CityGML model allowing others to add details using ADE mechanism?
- **Gaming**
  - Minecraft environment for DK, are there other similar activities?
  - Gaming companies use geo data, but convert them into a very specific, optimized form.