Using controlled vocabularies in the framework of Inspire

Inspire Conference 2014

Dr. Benoît DAVID

Geographic Information Division
Ministry of Ecology, Sustainable development and Energy - France

June 2014

Outline

I want to share with you :

1. A tale from implementation :
   the development of an operational controlled vocabulary for Inspire metadata
2. An idea for a simple Inspire
   controlled vocabulary for Inspire metadata
3. A feedback from building this Inspire controlled vocabulary out of
   - The Inspire registers
   - The Inspire feature catalog

Development of a controlled vocabulary for metadata (1/6)

- In France we recommend to affect each dataset to a predominant Inspire data theme
- And to capture it as a keyword in the metadata
- But public authorities have difficulties to find the correct theme corresponding to their datasets
- Specially in the field of environment regulations

Development of a controlled vocabulary for metadata (2/6)

- To help them we built 2 controlled vocabularies of well-known terms :

1. The name of “official” feature concepts defined by
   - Legal texts (with the reference to the legal text)
   - Statistical authorities (with the reference to the definition given by the statistical authority)
2. The name of public policies as structured by French codes :
   - Agriculture, local authorities, environment, spatial planning, transport, national defence, forest, industry, …

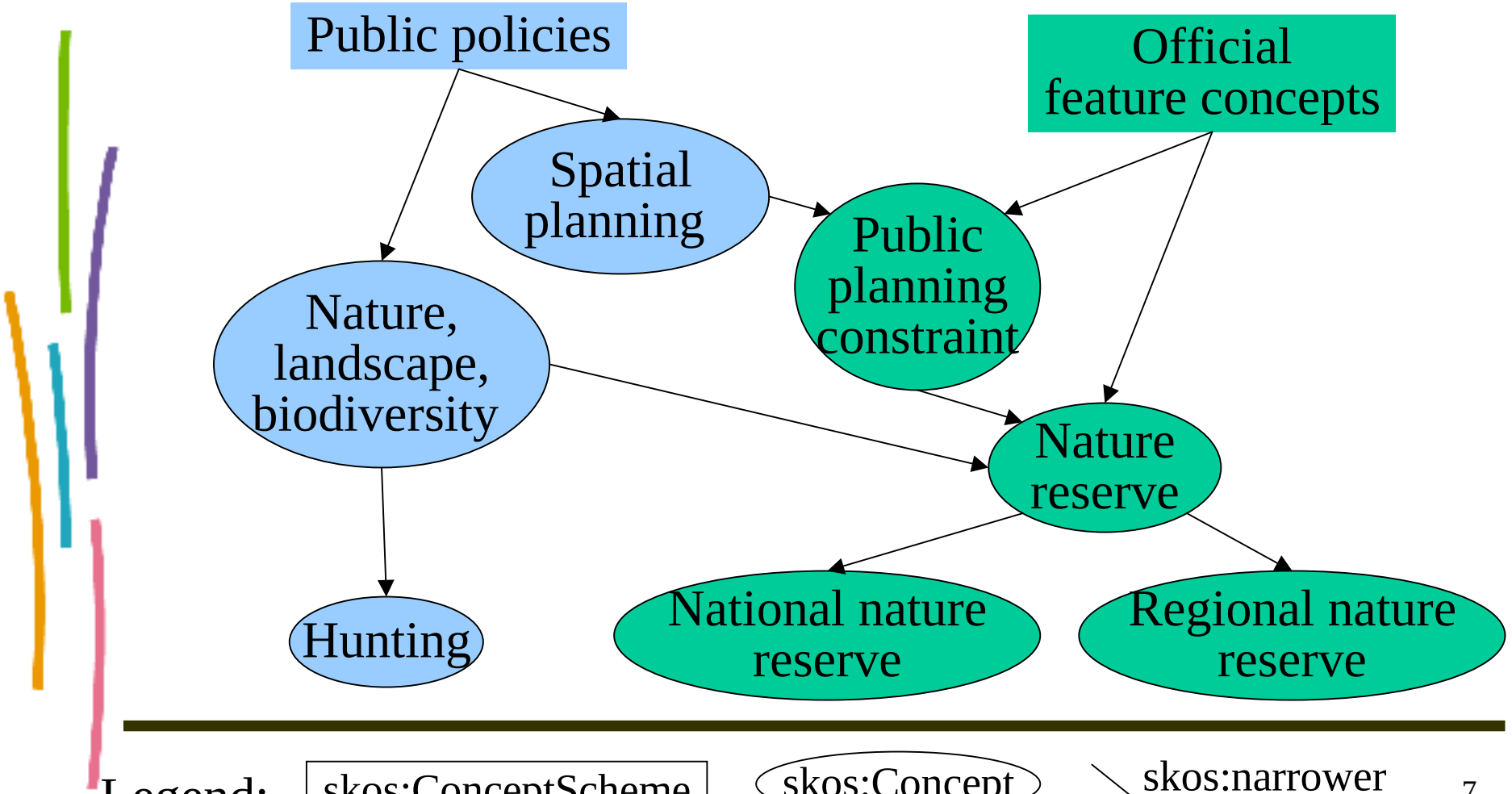
Development of a controlled vocabulary for metadata (3/6)

- To each “official” feature concept, an Inspire theme is affected or a particular “Out of Inspire” term
- This methodology is rather simple and objective
- The terms are already known by a lot of peoples in administrative organizations
- The vocabulary covers many thematic fields except most of the datasets from mapping agencies (I will come back later on that point)

Development of a controlled vocabulary for metadata (4/6)

- Hierarchical organization of the vocabularies is important
  - Some official features may be specific to other more generic official features
  - Official feature concepts are included in a policy
- There is also polyhierarchy, eg. a feature in two different policies
- The vocabularies have been implemented using the W3C SKOS standard principles

Example



Legend: skos:ConceptScheme skos:Concept skos:narrower

Development of a controlled vocabulary for metadata (5/6)

- A first version of the vocabularies has been published as an HTML page
- http://www.geocatalogue.fr/www/affectthemes.html
  - 425 feature concept names
  - 49 policy and sub-policy terms
- This methodology can be applied in other countries

Development of a controlled vocabulary for metadata (6/6)

- To be easily used, these vocabularies should be imported in the tools used to manage metadata
- In France, a lot of public authorities use Geonetwork, or its French version Géosource
- Unfortunately Geonetwork/Géosource can’t cope with such hierarchically structured vocabularies
- We are going to implement this functionality in Geonetwork/Géosource in the next months
  - > For more information : E. Taffoureau (BRGM)

A simple Inspire controlled vocabulary for metadata (1/3)

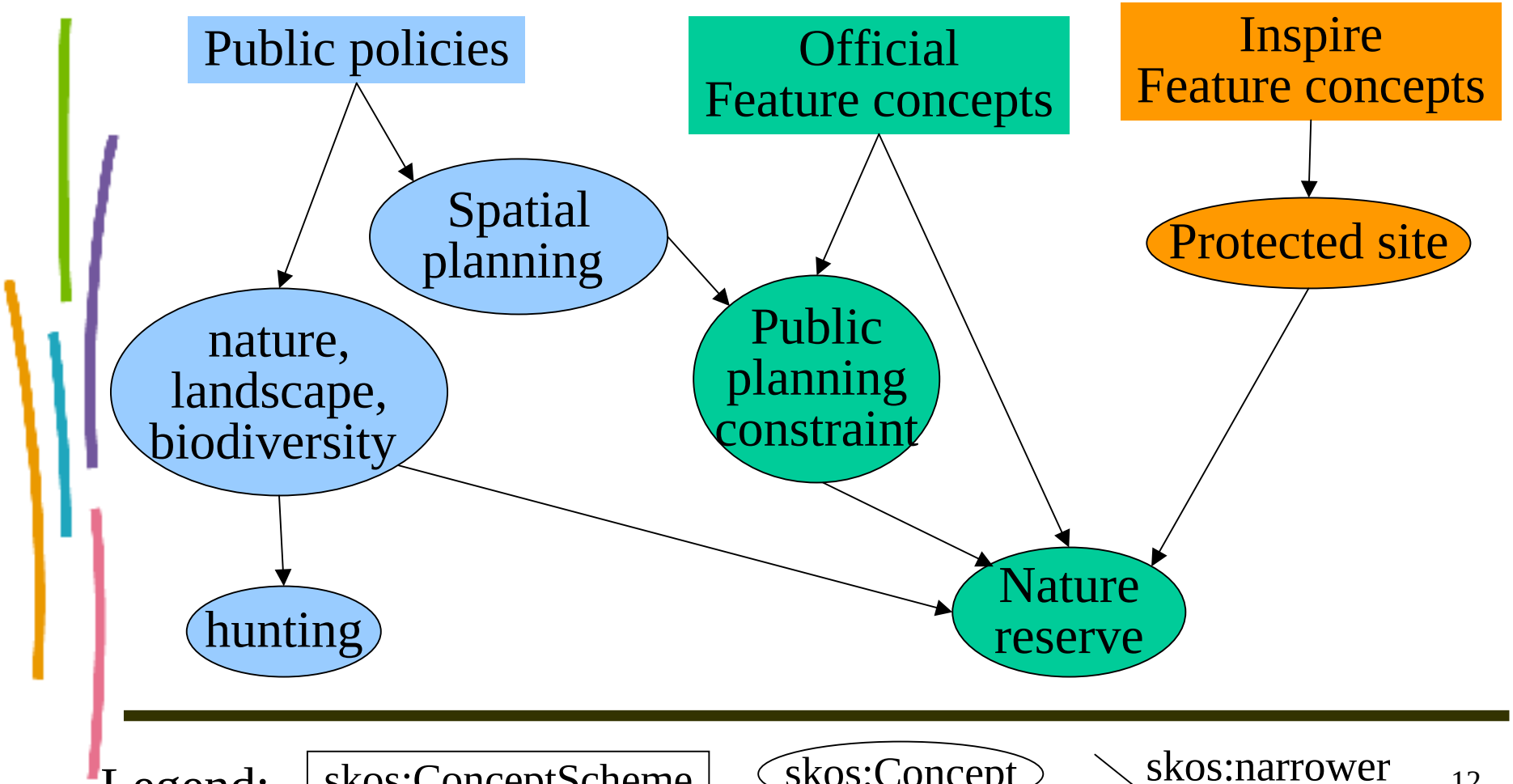
- A lot of datasets from mapping agencies do not correspond to official feature concepts, eg Elevation, Land cover, Orthoimagery, …
- For these datasets, it is easier to identify the correct Inspire theme
- A simple idea for these datasets is to use as keywords the names of the feature concepts from the Inspire interoperability regulation
- They could be completed with some values of the code lists which semantically refine some feature concepts
- Examples :
  - Breakline / top of slope , Spot Elevation / summit

A simple Inspire controlled vocabulary for metadata (2/3)

Advantages :

- To harmonize their spatial data sets, public authorities will have to know the corresponding feature concept
- Feature concepts are already defined in the different languages of the EU
- This vocabulary can be used with the two others by defining generic/specific relationships between terms of the vocabularies

Example



Legend: skos:ConceptScheme skos:Concept skos:narrower

A simple Inspire controlled vocabulary for metadata (3/3)

- This “Inspire controlled vocabulary” could also be “aligned” with some reference thesaurus such as
  - GEMET
  - EuroVoc
  - AGROVOC
- This alignment would only be done once and shared between all authorities

Feedback from building this controlled vocabulary (1/3)

- I tried to generate the simple Inspire controlled vocabulary from the Inspire registers
- It was not possible because there are no link in the registers between feature concepts and lists of codes
- I tried to use the Inspire feature catalog
- I had a lot of difficulties that I reported to the JRC

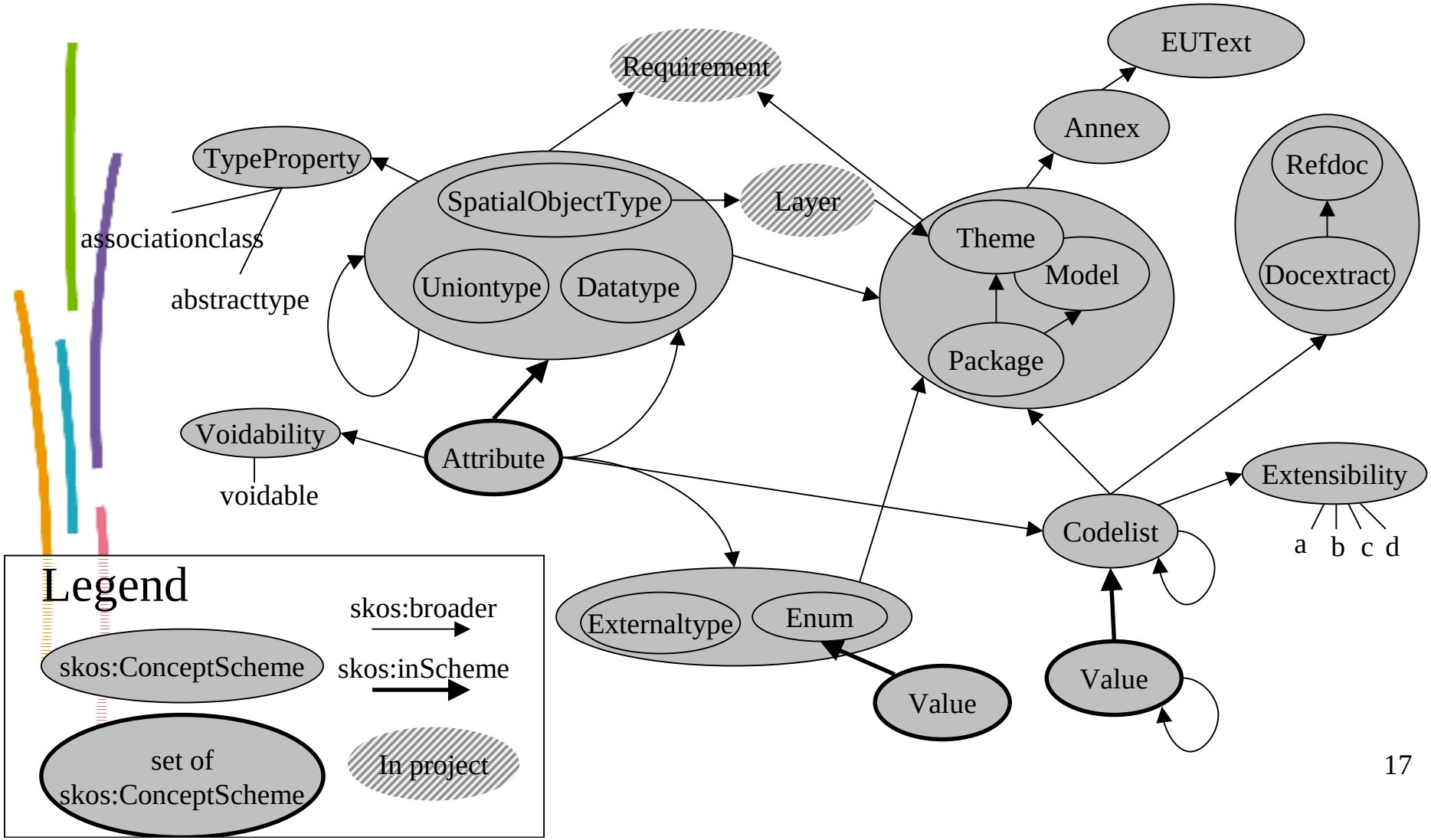
Feedback from building this controlled vocabulary (2/3)

- I also found a lot of discrepancies between
  1. the interoperability regulation
  2. the view given by the registers and the feature catalog
- Examples:
  - In the registers, no distinction between feature concepts from the regulation and those from technical guidelines
  - 2 fake themes in the theme register : Metadata and Generic Conceptual Model
  - FC which should only countain regulation elements countains other elements (OilGasChemicalsPipeExtended in approved/r4618-ir/fc/)
- Comparisons are difficult because the regulation is only text
- **Need for a set of registers corresponding to the perimeter of the interoperability regulation**

Feedback from building this controlled vocabulary (3/3)

- Finally, a new set of registers has been built
  - With the concepts defined by the interoperability regulation
- With the following objectives :
  1. Help people to browse though the regulation
  2. Compare the regulation to the Inspire registers and the feature catalogue managed by the EC
  3. Infer new information, such as, for example, the Inspire controlled vocabulary I defined before
- A first result of this work, still experimental, is available at http://docinspire.eu/
- Inconsistencies in the regulation were also reported

Concepts extracted from the interoperability regulation



Conclusions

- A first version of the first 2 controlled vocabularies (official feature concepts and policies) is available but difficult to use because of limitations of Geonetwork/Géosource which have to be improved
- We plan to extend these controlled vocabularies with Inspire feature concepts and some values of code lists
- Some improvements are needed on the Inspire registers & feature catalog
- There should a clear distinction between elements defined by regulations and those from TG