Metadata on Biodiversity: Definition and Implementation

Etienne Taffoureau, BRGM
Overview

> Context
  • SINP & ECOSCOPE
  • INSPIRE
  • Biodiversity concepts

> Metadata on biodiversity
  • Profiles definition
  • Standards implementation
  • Relation between the profiles
  • Publication & Access

> Conclusions
SINP & ECOSCOPE

> Two scientific infrastructures on biodiversity relying on different data sources and producers
> Main objective: document and share information on biodiversity in France
> Provide tools for metadata management
> **SINP**: Information system on nature and landscape
> **ECOSCOPE**: Observation for research on biodiversity data hub

DC-2016, Copenhagen, 2016-10-13
INSPIRE directive

> **Objective**: to establish a spatial infrastructure in Europe

> **Different obligations, including**: 
  - to do metadata related to datasets & services
  - to provide metadata through a discovery (web)service

> **Definition of ~20 metadata elements**
  (mandatory & conditional)

> **Implementation**: 
  - ISO 19115 (datasets) and 19119 (services)
  - XML 19139 (exchange format)
  - CSW 2.0.2 ISO AP (discovery service)
Biodiversity concepts

OBSERVATION SYSTEMS
A data acquisition system as part of a project (monitoring of marine invertebrate, …)

DATASETS
- Nature (Observation, protected sites, …)
- Landscape
- Geology (geological sites)

Occurrence
Collection
Reference
Geographic

SERVICES

DATABASES

ACTORS
Metadata profiles - Definition

**ECOSCOPE**
- General description
- Taxonomic coverage
- Temporal coverage
- Geographic coverage
- Content
- Material and methods
- Maintenance & change history
- Distribution, access and use of the data
- Associated collection
- Additional information
- Contact data
- Contact metadata

**SINP**
- General description
- Point of contact
- Thematic
- Format
- Reference data
- Geographical features
- Constraints on access & use
- Distribution
- Quality
- Metadata
Choice of the metadata standards

> It depends on:
  - objective of the project
  - utilisation of the metadata
  - data type
  - level of details

> We needs to have:
  - an architecture arborescent & modular = flexible & adaptable
  - to ensure interoperability between the systems
    - metadata transfer
    - description of data organisation in the information system & their acquisition

> Use of « eXtensibleMarkup Language (XML) » schema
  - Content & structure definition
  - Transfert protocol (import/export)
  - Allows format conversion (interoperability)
Metadata standards for SINP & ECOSCOPE

**SINP**
- Extension of ISO 19115 & 19139 (XML implementation)
- XSD schema extended
- Description of spatial information
- INSPIRE compliant

**ECOSCOPE**
- EML (Ecological Metadata Language)
- Description of information acquired from ecological researches
- For observation & experimentation
ISO 19115:2003 - UML diagram

DC-2016, Copenhagen, 2016-10-13
EML (Ecological Metadata Language)

Dataset
- Dataset description

Citation
- Publication & citation

Software
- Software & data organisation

Protocol
- Methods & protocols

Resource
- General information

(needs adaptations)

Other modules:
- eml-party – information related to persons, organisations, …
- eml-entity – information related to files
- eml-attribute – information related to attributes
- eml-access – information related to access conditions
- eml-distribution – information related to data distribution

DC-2016, Copenhagen, 2016-10-13
Relations between metadata profiles

Mapping between the different metadata profiles

- Definition of the metadata elements
- Cardinality
- Lexicons

DC-2016, Copenhagen, 2016-10-13
Publication & access of biodiversity information

Producer -> EML

XML ISO 19139
Harvesting CSW

Producer -> User

XML ISO 19139
Harvesting CSW

Producer

DC-2016, Copenhagen, 2016-10-13
Conclusions

> SINP & ECOSCOPE have the same objectives but data sources & producers are different

> This has led:
  • to define different metadata profiles
  • to use different standards

> In order to describe data on biodiversity, ISO 19115 standard:
  • is adapted for general & spatial information
  • but not for specific information (taxonomic information, actors, …)

> Solutions:
  • SINP: creation of an extension of ISO 19115
  • ECOSCOPE: utilisation of additional metadata elements of EML

> In order to ensure interoperability it is necessary:
  • to do a mapping between profiles
  • to provide conversion tools
Thank you for your attention!

http://www.naturefrance.fr/sinp/
http://www.fondationbiodiversite.fr/fr/recherche/programmes-frb/ecoscope.html