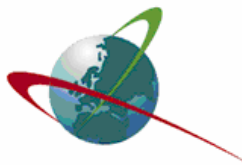


Metadata Model for the European Forest Information and Communication Platform



Dirk Tilsner
Carlos Figueiredo



Henrique Silva



Benjamin Chartier



Jesús San-Miguel Ayanz
Andrea Camia
Michel Millot

- Objectives and scope of EFICP
- Interoperability and system connectivity
- Overview of the EFICP metadata model
- Baseline metadata model
- Extended metadata model
- Outlook

Main goal

- Follow up of the European Council Regulation no 1615/89 concerning the creation of a European Forestry Information and Communication System (EFICS);
- Builds on previous initiatives and pilot projects such as EFIS and NEFIS;
- Main goal for EFICP is to become the European information and communication platform in the forest sector, providing facilities for the search, online access and analysis of all kinds of forest-related information;
- A major technical focus in interoperability, i.e. to enable EFICP to exchange data with a wide range of external systems and services, adopting INSPIRE guidelines and implementing rules.

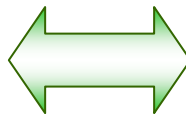
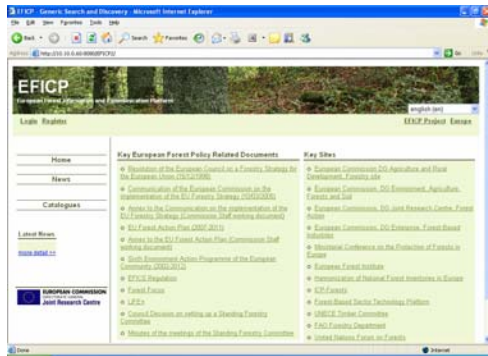
Users

- European / national agencies
- Land planners (local, regional, national level)
- Forest managers and forest-based industries
- Stakeholders
- Scientific community
- General public

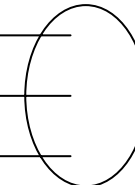
Key services

- 🌿 Resources search and discovery
 - by subject, location, time and resource type
- 🌿 Access and download
 - includes geospatial data download service
- 🌿 Data Visualisation and Analysis
 - geospatial data, forestry related statistical indicators and thematic mapping
- 🌿 Report Generation
 - personalisation
- 🌿 Platform and Data Management
 - includes metadata moderation and management

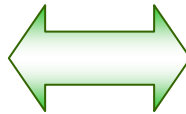
EFICP as front-door to forestry and related information in Europe



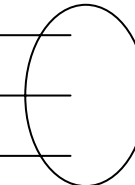
INSPIRE / Geo-Portal



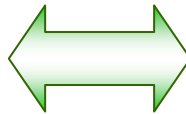
Data Centres
 GFIS



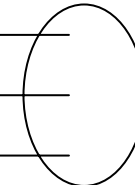
European Institutions



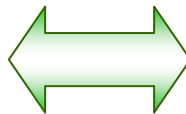
EFISCEN
 Forest Focus
 EFFIS
 Eurostat



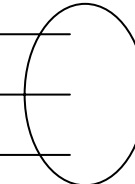
**National Institutions
 (e.g. NFIs)**



UN Comtrade
 CLC 2000



**Internat. Institutions
 & GFIS**



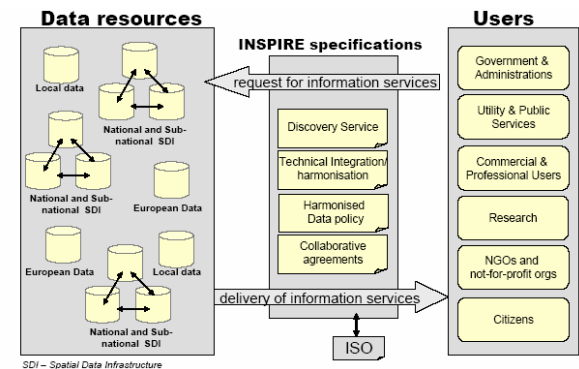
FAO
 CIFOR
 Inventories



...

EFICP as Spatial Data Infrastructure

- Distributed data resources
 - EU Geoportal, external catalogues, OGC services, data providers
- Harmonisation of data specifications
 - namely ensuring comparability of forestry data from heterogeneous resources
- Metadata model for search & discovery and on-line access to data
 - Bibliographic, news, geospatial, statistical data resources
- Geospatial information services
 - OGC harvesting, view & overlay, download
- Technical interoperability
 - SOA & web services, open standards



Requirements for interoperability and system connectivity

Technical Interoperability

- SOA / web services architecture and related open standards
- ISO/OGC standards for geographical information exchange (web mapping)
- Metadata harvesting and publishing (CS-W, OAI-PMH, z39.50)
- Multi-lingual support and corresponding message encoding

Organisational Interoperability

- General compliance with INSPIRE guidelines and implementing rules
- Means for metadata upload (manual, automatic) and moderation based on defined data quality management procedures

Requirements for interoperability and system connectivity

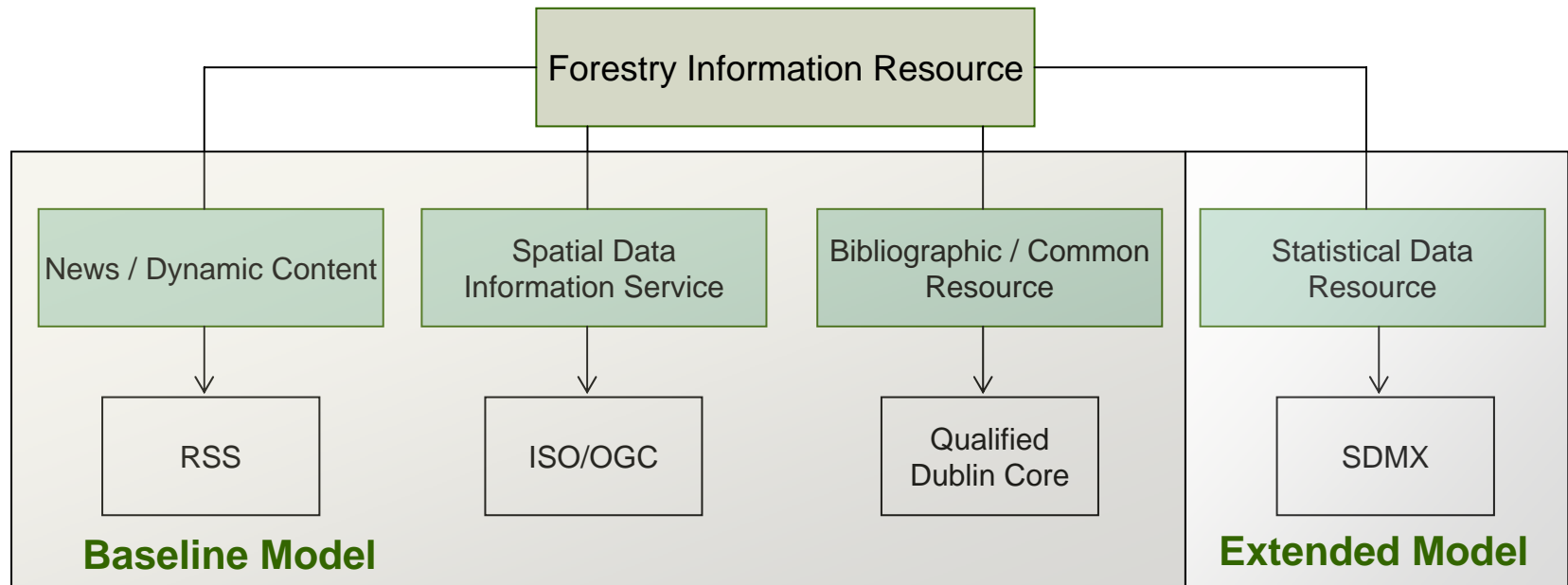
Semantic Interoperability

- Application of metadata guidelines provided by the previous NEFIS project (Dublin Core; proposal for controlled vocabulary for the forestry domain)
- Compliance with INSPIRE implementing rules for metadata (Level 2 for discovery)
- Compliance with ISO 19115 for the resource discovery and consideration of other (non-European) standards such as FGDC
- Extended metadata for the description of forestry-specific and related data resources, following a common domain model
- Support metadata interoperability (mapping of elements)

Types of resources vs. discovery and accessibility

Resource Type	Search and Discovery	Online access
News	by search term, time, language, location	direct access and visualisation
Bibliographic / common resource	by search term, time, language, location	direct access and file download
Geospatial service	by search term, time, language, location	WMS/WFS, view (with overlay) and download
Forestry (and related) statistical data resource	searched as common resource (indicators for specific resource type)	access at data set level (additional metadata) for visualisation & analysis

Types of resources vs. open standard for metadata



The Baseline Metadata Model is based in the **Dublin Core (DC)** standard and follows the same major principles:

- Simple model to facilitate the finding, sharing and management of information;
- Just 3 mandatory elements (Title, Description and Date). In Dublin Core each element is optional to achieve major interoperability;
- Each element may be repeated;
- Refinements of the DC Model following the DC rules.
- Use of encoding and vocabulary schemes;
- Simple XML encoding (no nested tags);
- xml:lang attribute for multilingual support.

Element Refinement: qualifiers that make the meaning of an element narrower or more specific.

Encoding Scheme: a term selected from a controlled vocabulary or a string formatted in accordance with a formal notation (e.g., "2000-01-01").

Examples of DC elements, refinements and encoding schemes

Element Type	Example	Comments
Simple DC Elements	<code><dc:title>Land Cover and Forest Classification of Finland</dc:title></code>	Namespace “dc”
Qualified DC Elements	<code><dcterms:modified>1997-07-27</dcterms:modified></code>	Date of the modifications in the resource. DC refinement of the “dc:date” element. Namespace “dcterms”
DC Encoding Schemes	<code>< dcterms:spatial xsi:type=“Box”> northlimit= 70.01;southlimit= 59.60; westlimit= 19.09; eastlimit= 31.58</ dcterms:spatial ></code>	Resource geographical bounding box. Attribute “Box” of “dcterms:spatial” element
EFICP Controlled Vocabulary	<code><dc:type xsi:type= “eficp:EFICPType”> Statistical data</dc:type></code>	EFICP Resource Type. EFICP Refinement of the “dc:type” element. Namespace “eficp”



Conformance with the INSPIRE directive

Although the Baseline Metadata Model is design to deal with all kind of forestry resources and INSPIRE is a directive for geographical resources, refinements where made to achieve conformance with INSPIRE Discovery Level 1 interface and Article 8.

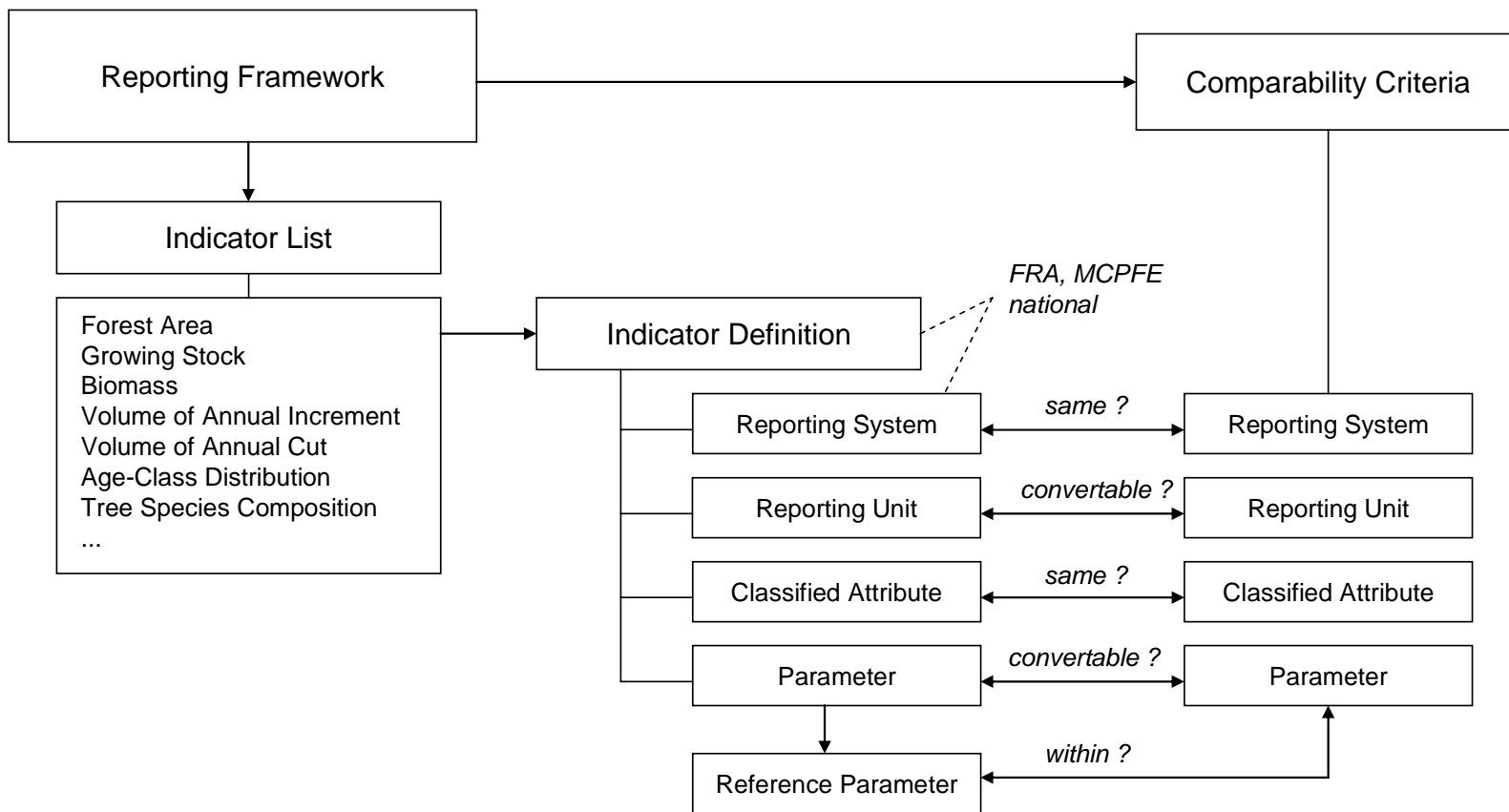
The following elements were included in the Baseline Metadata Model as refinements and encoding schemes of the DC model:

Quality, Topic Category, Spatial Resolution, Service Type and Rights Restriction Code.

Constraints for online access to statistical data resources

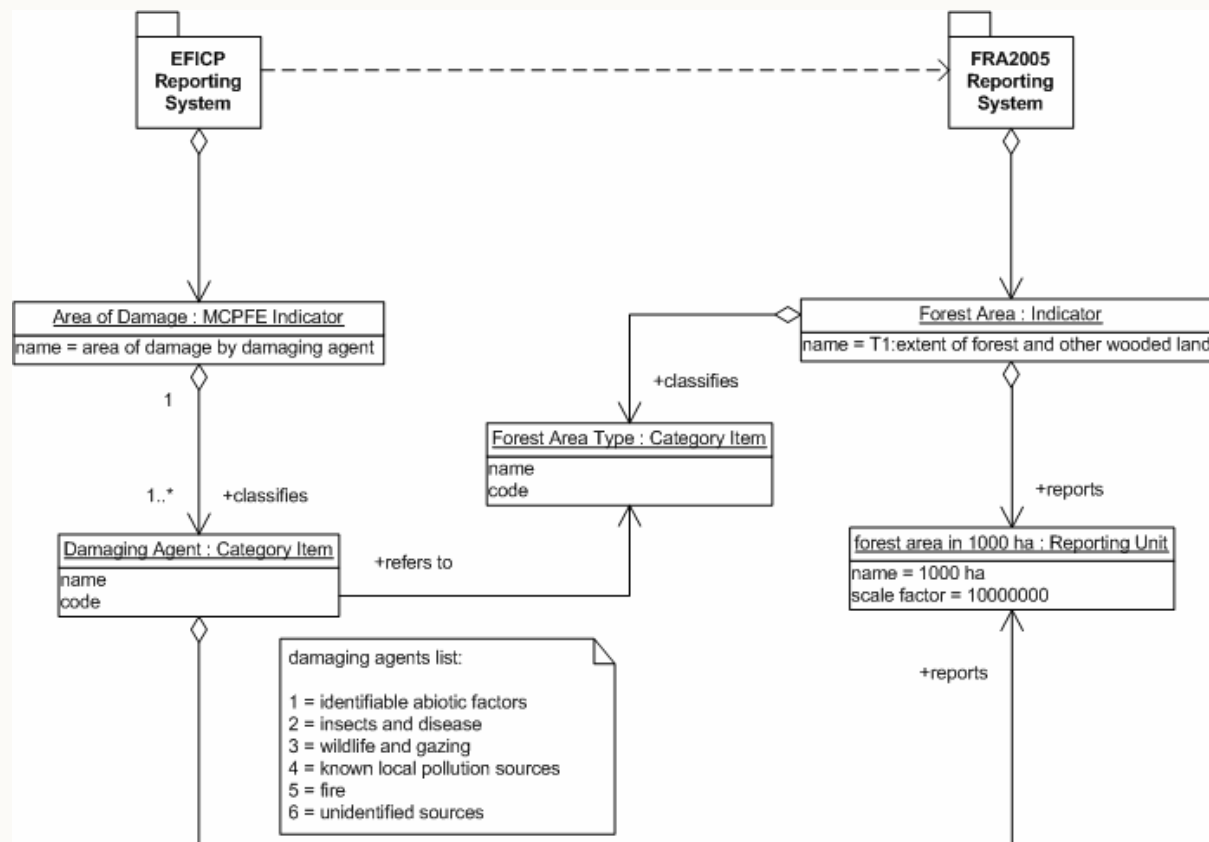
- Harmonisation of forestry data specifications is a long-term goal
 - Previous and ongoing initiatives such as EFICS study, EFIS/NEFIS, ENFIN, ...
 - Reporting at international level (FAO/EUROSTAT - FRA, MCPFE, ...)
 - European Forestry Strategy and EU Forest Action Plan (includes monitoring and corresponding harmonised information dissemination on forests in Europe)
- Scope of a general domain model for these resources
 - Adoption of rules and guidelines (FRA, MCPFE) for reporting of indicators
 - Definition of structural information of data sets and encoding
 - Criteria for comparability of data from heterogeneous sources
 - Conversion factors and rules
- Requirements of a Metadata model
 - General: simplicity, expressiveness, extensibility, semantic interoperability
 - Other: multi-lingual support, data quality, online query to services

EFICP supporting reporting of forestry indicators



General domain model

Modelling for Specific Indicators



Metadata model as 'SDMX – implementation' of the domain model

- **SDMX: Statistical Data and Metadata eXchange**
 - has emerged (since 2001) as most advanced, international standard for statistical data and metadata access and exchange
 - is a general / conceptual information model from which syntax (domain) specific implementations can be developed
 - Currently version 2.0, version 1.0 already approved as ISO standard
 - Powerful and flexible solution for the definition of complex structures and data set specific metadata structure, satisfying all mentioned requirements

- **EFICP metadata model**
 - Structural definitions (key families) per indicator + generic format
 - Reporting definitions implemented on the basis of classified attributes (code lists), parameters, reporting units + simple metadata structure (data set)
 - Use of particular SDMX-ML messages for encoding, querying and transport

SDMX Example

- <Concepts>

```

- <structure:Concept id="CO_FOREST_AREA_TYPE" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">FOREST_AREA_TYPE</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_FREQ" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">FREQ</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_INDICATOR_TYPE" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">INDICATOR_TYPE</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_OBS_STATUS" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">OBS_STATUS</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_REGION" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">REGION</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_REP_UNIT_TYPE" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">REP_UNIT_TYPE</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_SCALE_FACTOR" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">SCALE_FACTOR</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_UNIT_TYPE" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">UNIT_TYPE</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_YEAR" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">YEAR</structure:Name>
</structure:Concept>
- <structure:Concept id="CO_OBS_VALUE" agency="EDI" version="1.0" uri="http://www.edisoft.pt">
  <structure:Name xml:lang="en">OBS_VALUE</structure:Name>
</structure:Concept>
</Concepts>

```

Dimensions Key		
Type	Concept	Representation / Code List
Dimension (role is Frequency)	FREQ	Code List: CL_FREQ
	OR_TYPE	Code List: CL_INDICATOR_TYPES
		Code List: CL_NUT_REGIONS
	_AREA_TYPE	Code List: CL_FOREST_AREA_TYPES
		DateRange

reporting unit type is "area".
 recommended unit of measurement (scale factor *) is 10⁷ m².

- Contract started in January 2006
- Currently a pre-operational version is under evaluation
- Workshop to invited guests from Member States in July 2007
- Further development and integration of external providers by end of 2007
- Final system fully available in early 2008

