

# The eENVplus Harmonization and Validation toolkits

G. Martirano, F. Vinci, S. Morrone (EPSIT)



#### Outline

- Data requirements
- Harmonisation and validation toolkit
- Toolkit demonstration
- The Data Harmonization section in the collaborative platform forum



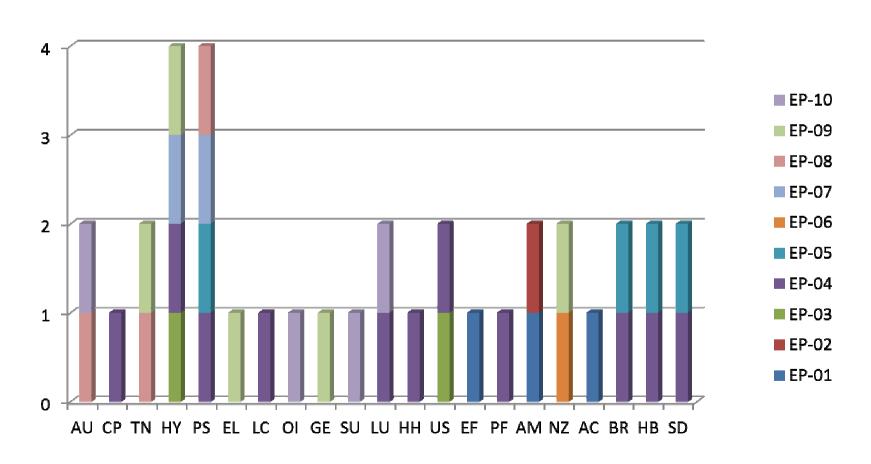
#### Data modelling requirements

- Identified the target schemas to be used in the harmonisation process by each pilot:
  - □ the relevant gml application schema (xsd) of the relevant INSPIRE Data Specification (21 DS)
  - □3 additional target schemas:
    - ■AQD schema of EEA for Air Quality Reporting
    - ■AGIV IMKL2.1 schema for cables and pipes in Flanders region (extending INSPIRE US DS v3.0rc3)
    - ■GeoSciML 3.2 schema for Geology in specific use case (extending INSPIRE GE v.3.0 as described in the Technical Guideline)



## Data modelling requirements

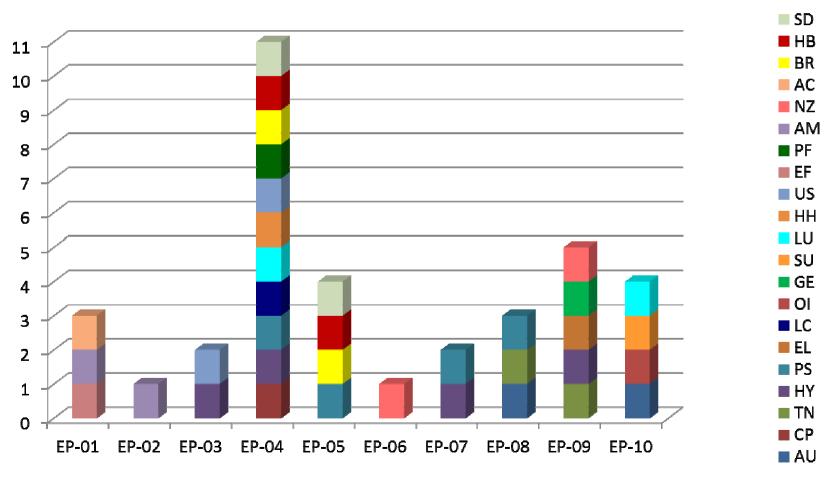
## Distribution of the 21 INSPIRE data-themes per Pilot (1/2)





## **ENUPLUS** Data modelling requirements

#### Distribution of the 21 INSPIRE data-themes per Pilot (2/2)





### Metadata modelling requirements

Identified the INSPIRE profile as target schema for the metadata, with the possibility to extend the profile using additional ISO19115 metadata elements (if required by eventual additional national requirements and/or specific pilot requirements).



#### **Toolkits**





The National State of State of

#### D 3.1 – Datasets and metadata harmonization toolkit

Revision: 3.0

Author(A)/Organiamion(A):

- Glacomo Martirano, Fablo Vinci and Safania Morrone (EPSIT)

Work package? Task:

WP9 - Harmonization and Validation

T9.1 Harmonization toolkis

	Feint selvent is in Businer Committee with its CT Policy Light Feigure	
	Supriger (and	data a
	Part of the Control o	- 00
E :	Confession, only to receive of the senseture and the Commission Senses	

for some \$10 forms or name to which have

April Inc. St





Priest brown all Carls

Programme of the State of the S

#### D 3.2 – Datasets and metadata validation toolkit

Revision: 1.0

#### Author(s)/Organisation(s):

Glacomo Martirano, Fabio Vinci and Stefania Morrore (EPSIT)

Work package / Tack:

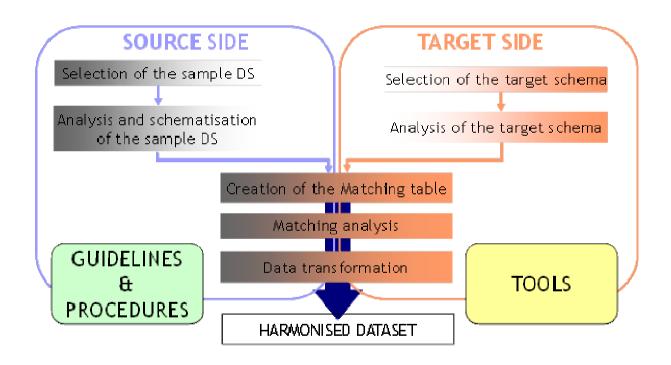
WP3 - Harmonization and validation

T3.2 Validation toolkit

	Project co-funded by the Burogean Commission within the ICT Policy Suggest Progra	anne:
	Discenination Level	
P	Public	×
c .	Confidental, only for members of the consortium and the Commission Services	1 10-



#### Steps of the data harmonization process



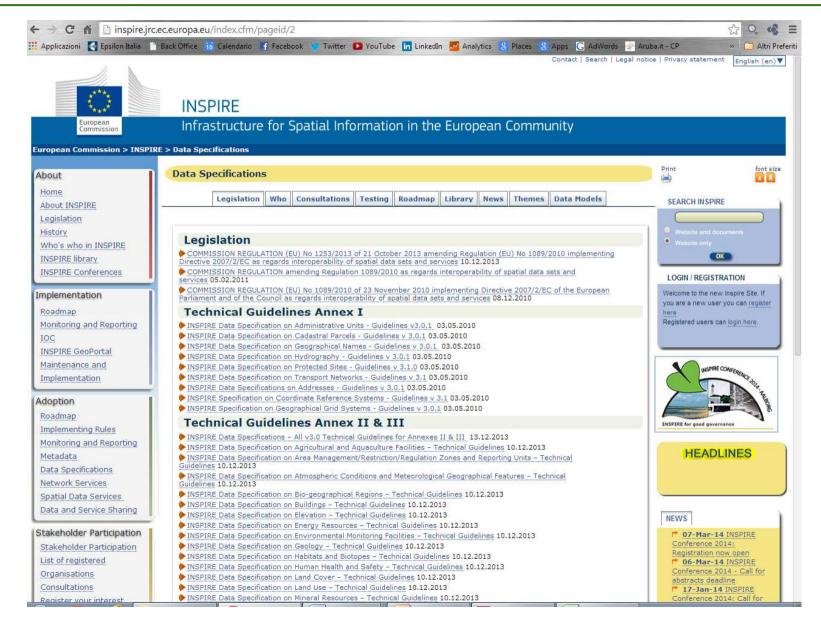


- 1. Analysis of the source dataset and its associated data model
- 2. Selection of the target schema best fitting for purpose with the source dataset and with the objective of the transformation
- 3. Analysis of corresponding Data Specification:
- by means of the relevant INSPIRE Data Specification and of its UML representation, both available in the INSPIRE website;
- by means of the available documentation when the target models does not correspond to a specific INSPIRE data theme.



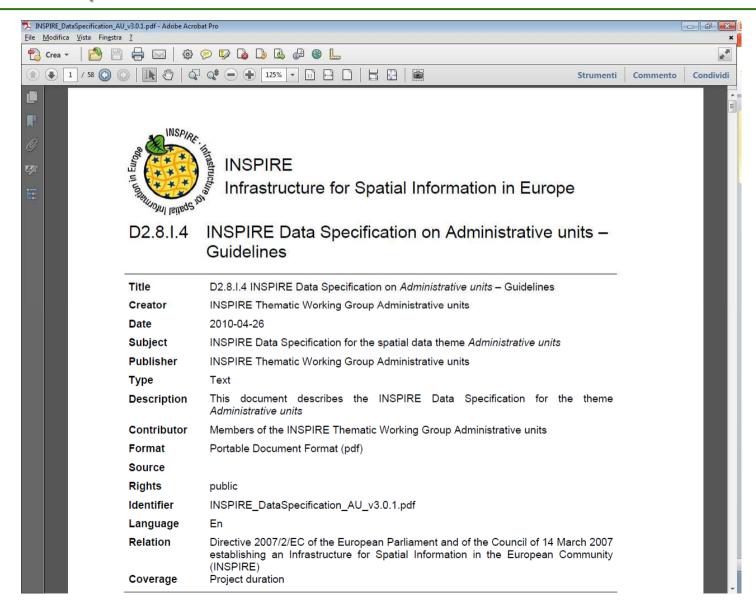
- 4. Filling-in of the mapping (matching) table.
- It's the most crucial harmonization step!
- Performing very carefully this exercise, analysing and solving the eventual mapping problems, strongly facilitates the transformation.
- 5. Transformation of the source dataset by means of software transformation tools.
- 6. Validation of the transformed dataset





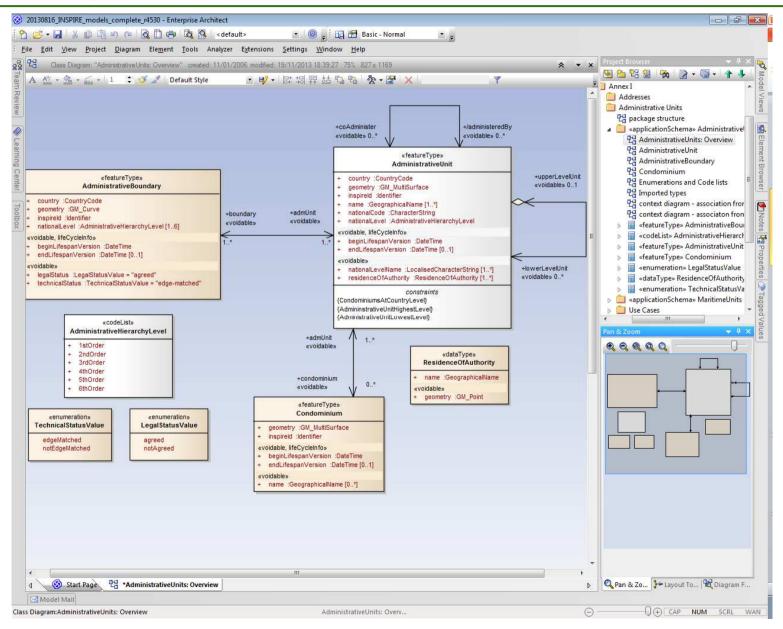


#### Target model





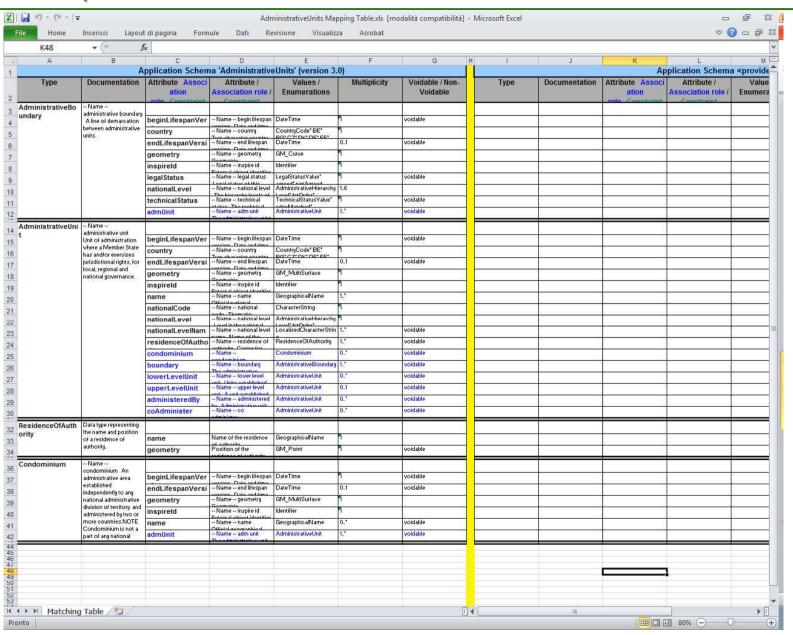
#### Target model





- In each mapping table there is a single row for each attribute of the feature types.
- Because some attributes have a complex data type with a tree structure, it may be useful to extended the JRC mapping tables in order to take into consideration the complex data types.
- each attribute of the source dataset has to be mapped to relevant attribute in the target schema.



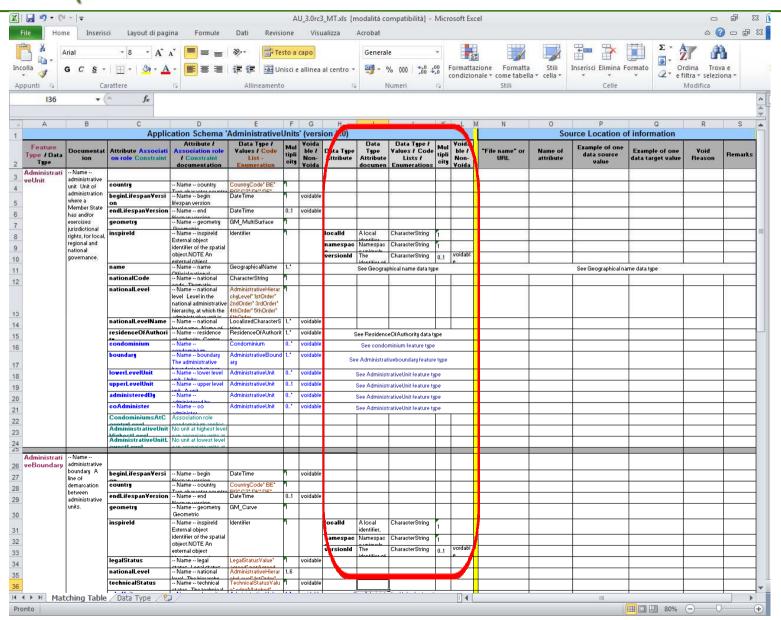




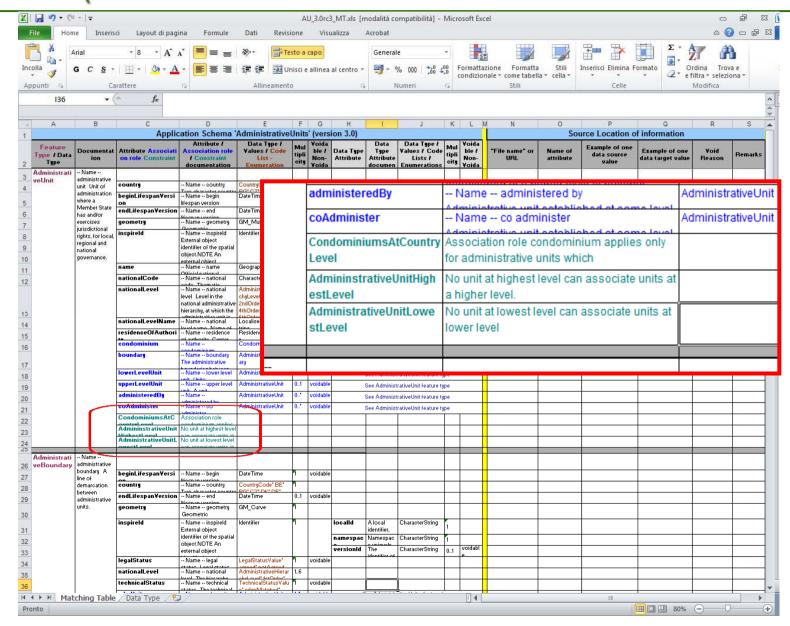
				The administration only			
14	AdministrativeUni	Name administrative unit					
15	L		beginLifespanVer	Name begin lifespan	DateTime	1	voidable
16		where a Member State has and/or exercises	country	Name country	CountryCode" BE"	1	
17		jurisdictional rights, for	endLifespanVersi	Name end lifespan	DateTime	01	voidable
18		local, regional and national governance.	geometry	Name geometry	GM_MultiSurface	1	
19			inspireld	Name inspire id	ldentifier	1	
20		_	name	Name name Official pational	GeographicalName	1*	

Α	В	С		D	Е	F	G
		Applicati	ion Sch	ema 'Base Types'	(version 3.3rc3)		
Туре	Documentation	Attribute	Associa	Attribute /	Values /	Multiplicity	Voidable / Non-
		tion	1	Association role /	Enumerations		Voidable
		role Con	etraint	Constraint			
SpatialDataSet	spatial data.NOTE The type SpatialDataSet is						
		identifier		Identifier of the spatial	Identifier	1	
		metadata		Metadata of the spatial	MD_Metadata	1	voidable
ldentifier	identifier published by the responsible body, which	localld	$\neg$	A local identifier, assigned	CharacterString	1	
		namespace			CharacterString	1	
	may be used by external applications to reference	versionId			CharacterString	01	voidable



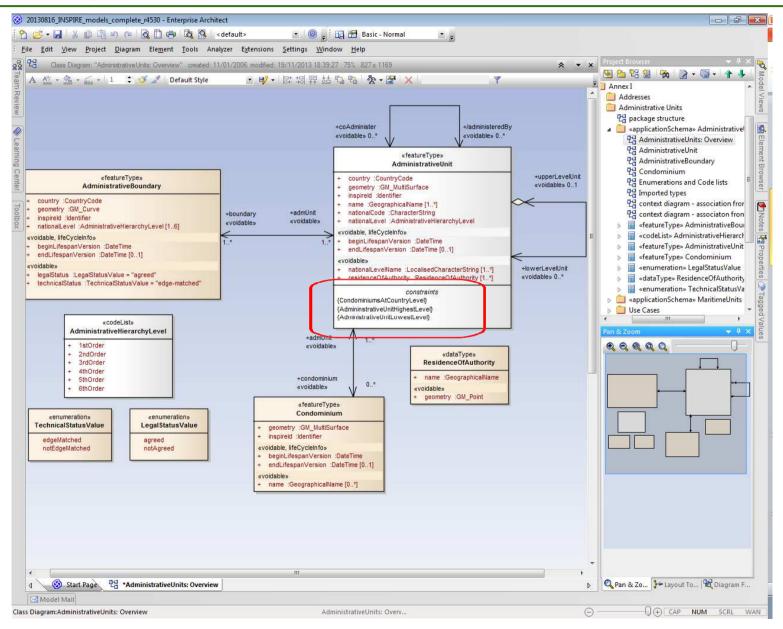




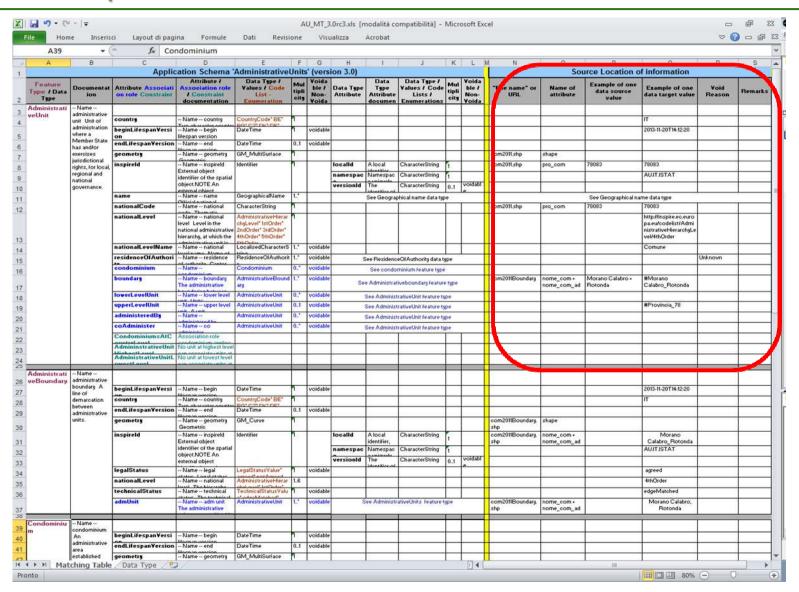




#### Target model









#### Data Transformation tools

■ The mapping between source and target properties defined in the matching tables can be used to set the encoding rules needed to obtain an harmonized dataset by means of a software transformation tool.



#### Data Transformation tools

- Among the many software transformation tools available, focus has been given on:
  - open source sw: HUMBOLDT Alignment Editor (HALE) open source tool to define and evaluate conceptual schema mapping and to transform geodata based on these mapping.

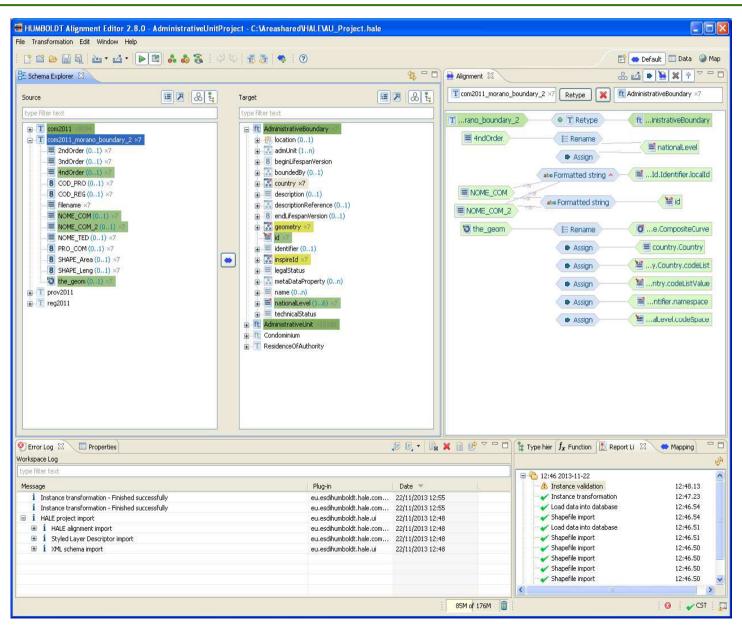
http://hale.igd.fraunhofer.de/2.8.0/help/inde
x.jsp

proprietary sw: GO Publisher (distributed by Snowflake Software -

http://www.snowflakesoftware.com/products
/gopublisher/)

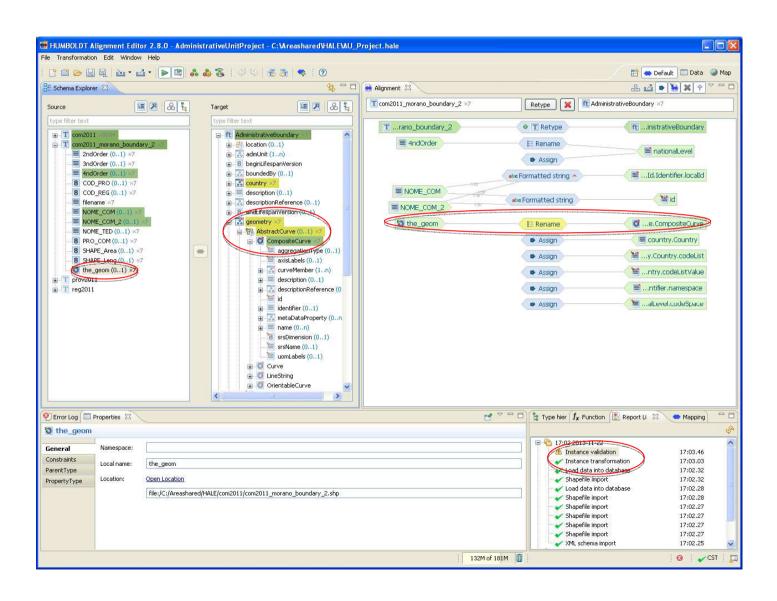






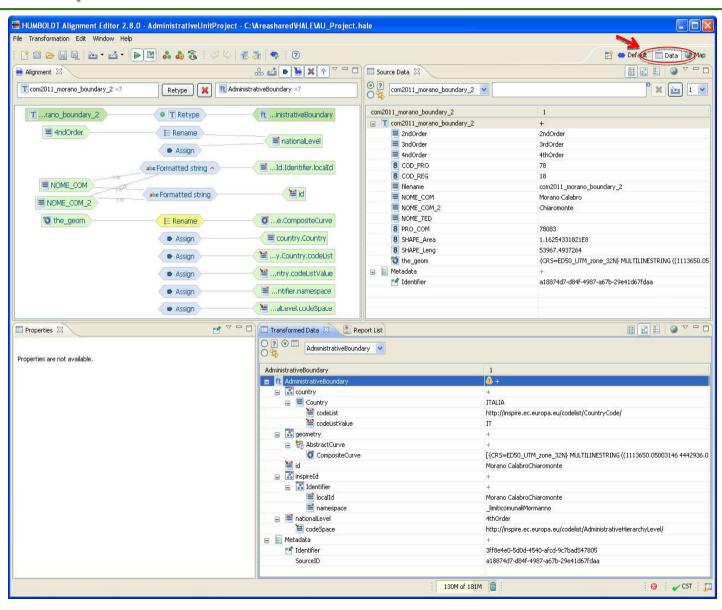






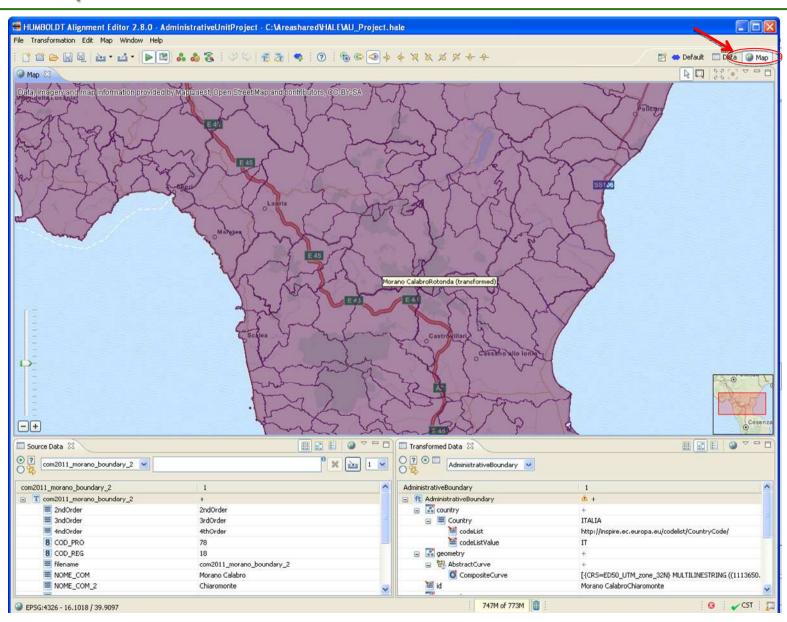


#### **HALE**





#### **HALE**

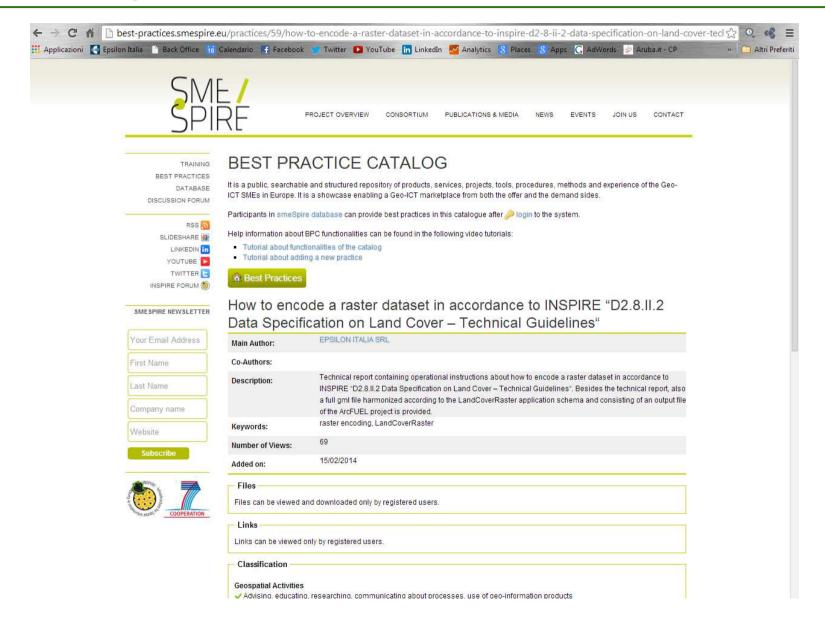




#### GoPublisher

GO Publisher Desktop				
File Edit Tools Help				
	R TO TO TO TO			
Project Settings SQL filters Desktop settings WFS settings Agent setting	s		198577	
Project name INSPIRE AU D: VAre	ashared\Limiti_Amministrativi\INSP	TRE All one	Form	nat: GML 3.2
Projectione Profession Dispersion		nar_uoright		
Target namespace Prefix				
Database to XML mapping				
Name Enabled	DB type or const value	XML path	Type in XML	Required
_ € metadata	Unknown	base:metadata/@gml:nilReason	gml:NiReasonType	Yes
□ COM2011	Table	base:member/au:AdministrativeUnit	au:AdministrativeUnitType	No
⊕ TPid	Column group	@gml:id	xs:ID	Yes
⊕ ¶ geometry ☑	Column group	au:geometry/gml:MultiSurface	gml:MultiSurfaceType	Yes
☐ PRO_COM	NUMBER	au:nationalCode	xs:string	Yes
⊕ 🎧 inspireId	Column group	au:inspireId/base:Identifier	base:IdentifierType	Yes
□ PRO_COM_2	NUMBER	base:localId	xs:string	Yes
C namespace	AU.IT.ISTAT	base:namespace	xs:string	Yes
C nationalCode	http://inspire.ec.europa.eu/co		gml:CodeType	Yes
□ C nationalLevelName □ □ □ □ country	Comune Column group	au:nationalLevelName/gmd:LocalisedCharacterString au:country/gmd:C	gmd:LocalisedCharacterString_Type gco:CodeListValue_Type	Yes Yes
□ 17 name	Column group	author antigeographicalName	gn:GeographicalNameType	Yes
C language	ita	nowerLevelUnit/@xlink:title	(anonymous)	Yes ≡
- C nativeness	endonym	au:lowerLevelUnit/@xlink:type	gn:(a. vnymous)	Yes
C nameStatus	Const	au:name	gn:(anon mous)	Yes
□	Column group	au:name/gn:GeographicalName	gn:(anony) ous)	Yes
─ C void_reason	Unknown	au:nationalCode	gml:NiReas nType	No
SourceOfName	OTHER	au:nationalLevel	gn:(anony ous)	No
□ □ pronunciation	Column group	au:nationalLevel/@gml:codeSpace	gn:Prony_ciationOfNameType	Yes
NOME_COM_1	VARCHAR2	au:nationalLevelName	▼ gn:(a onymous)	No
☐ ∏ spelling ☐ NOME_COM □	Column group VARCHAR2	g. nelling/gn:SpellingOfName gn:text	xs:string	Yes Yes
© script	Latn	gniscript	gn:(anonymous)	Yes
⊕ 🎧 residenceOfAuthority	Column group	au:residenceOfAuthority	au:(anonymous)	Yes
□ C beginLifespanVersion	2013-11-20T14:12:20	au:beginLifespanVersion	au:(anonymous)	Yes
€ NUTS	Unknown	au:NUTS/@gml:nilReason	gml:NiReasonType	Yes
⊕ ¶ upperLevelUnit ☑	Column group	au:upperLevelUnit/@xlink:href	xs:anyURI	No
☐ ☐ COM2011BOUNDARY	Table	au:boundary	gml:ReferenceType	Yes
⊕ 🏗 boundary 🔽	Column group	@xlink:href	xs:anyURI	No
PROV COM	NUMBER		71	No T
Preview XML Preview Schema Execution View				
Table Control of the				
<au: name=""></au:>		^		
<pre><gn:geographicalname></gn:geographicalname></pre>				
<gn:language>ita</gn:language>				
<pre><qn:nativeness>endonym</qn:nativeness></pre>				
<pre><gn:namestatus>official</gn:namestatus></pre>				
<pre><gn:sourceofname nilreason="Unknown" pre="" x<=""></gn:sourceofname></pre>				
	si: nii- crue />	E		
<pre><gn:pronunciation></gn:pronunciation></pre>		Gin		
<pre><gn:pronunciationofname></gn:pronunciationofname></pre>				
<pre><gn:pronunciationipa>Morano Calabre</gn:pronunciationipa></pre>	o <td>A&gt;</td> <td></td> <td></td>	A>		
· · · · · · · · · · · · · · · · · · ·				
Preview Sample Size: 1 - Update Preview Validate Preview	Y			





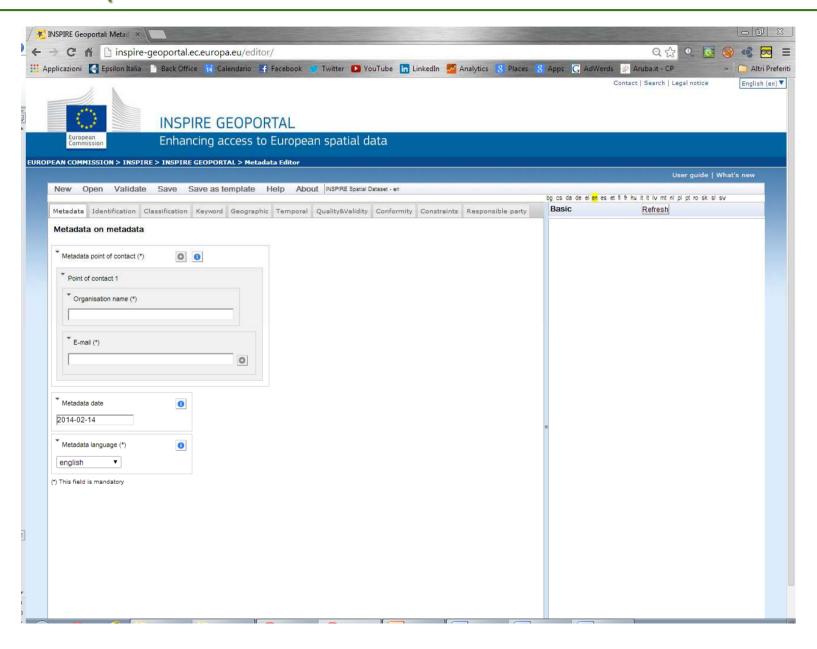


#### Metadata harmonization tools

- There are two possible options for metadata harmonization:
  - to transform existing metadata
  - to edit new metadata (from scratch or from an existing xml file)



#### Metadata editing tools



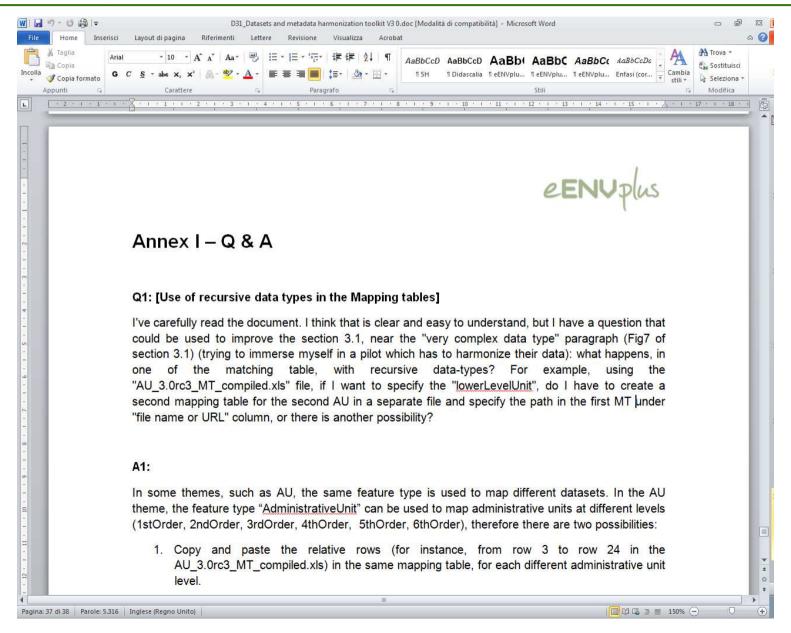


#### Metadata editing tools

P. TOURS AND		Reset Save	Save and close Ch	eck ① Other act	ions Cancel M	inor edit
INSPIRE view ISO		Nor				Parent/child metadata: Add or update parent metadata section
Minimum ISO Core			ailable			<ul> <li>Related service metadata:</li> <li>⊥ink service metadata</li> </ul>
ISO All  By Package  Metadata						Related feature catalogues:  Link feature catalogue
Identification  Maintenance  Constraints						
Spat. Info	IDENTIFICA	· · · · <u> -                                    </u>				
Ref. system	Title *	eu.Italy.Calabria.ArcFuel.Pri	oject_FuelClassif			
Distribution  Data quality	Date * Date type *	Creation 🔻				
App. schema	Unique resource identifie		CM			
Catalog	Codespace 🗵	ArcFuel_FCM	(Suggestion	s: 💌)		
Content Info Ext. Info	Abstract *	ArcFuel project aims develo	ping a generic 🔼			
XML view		methodology for creating fo which can be used for suppo	orting the			
		operational use of fire simul applications in context of for				
		management. ArcFuel uses recent effort of JRC Ispra, v	the results of a 💳			
		creating a standardized scho	eme of fuel			
		types representative of the forest regions and based on	this it defines			
		a methodology for producing	forest fuel			
	「▼Point of conta	act ⊕ ⊠				
	Organisation name	psilon Italia		Electronic mail address	info@epsilon-italia.it	
		oint of contact		address to M		
	I Graphic overview 🛨					
	Descriptive k	eywords 🛚 🖻				
	Keyword * ⊞ F	uel Load/Characteristic		= Thospumus n		
				▼ Thesaurus r	GEOSS - Earth Observ	vation Vocabulary, ver
				Date *	2000	<u> </u>
				Date type *	Publication 🕶	
	<sub>[</sub> ▼Descriptive k	eywords ⊠ 🗷 🗷 —				
	Keyword * ± Li	and Cover Fuel Classification Map				
	· ·	eywords # 🗷 🗷 🗷				
	Keyword * ⊞ Li	and cover		▼ Thesaurus r		
				Title *	GEMET - INSPIRE ther	nes, version 1.0
				Date *	Publication 💌	
In attesa di localhos	t			1		









#### Metadata validation

- Commission Regulation (EC) No. 1205/2008 implementing the INSPIRE Directive as regards metadata (<u>Discovery metadata</u>)
- Commission Regulation (EU) No 1089/2010 implementing the INSPIRE Directive as regards interoperability of spatial data sets and services (<u>Metadata for interoperability</u>)



## Discovery metadata

Metadata Regulation Section	Metadata element	Multiplicity	Condition
1.1	Resource title	1	
1.2	Resource abstract	1	
1.3	Resource type	1	
1.4	Resource locator	0*	Mandatory if a URL is available to obtain more information on the resource, and/or access related services.
1.5	Unique resource identifier	1*	
1.7	Resource language	0*	Mandatory if the resource includes textual information.
2.1	Topic category	1*	
3	Keyword	1*	
4.1	Geographic bounding box	1*	
5	Temporal reference	1*	
6.1	Lineage	1	
6.2	Spatial resolution	0*	Mandatory for data sets and data set series if an equivalent scale or a resolution distance can be specified.
7	Conformity	1*	
8.1	Conditions for access and use	1*	
8.2	Limitations on public access	1'	
9	Responsible organisation	1*	v.
10.1	Metadata point of contact	1"	
10.2	Metadata date	1	
10.3	Metadata language	1	ō.

#### Metadata for interoperability

#### 8.2 Metadata elements for interoperability

#### IR Requirement

Article 13

#### Metadata required for Interoperability

The metadata describing a spatial data set shall include the following metadata elements required for interoperability:

- Coordinate Reference System: Description of the coordinate reference system(s) used in the data set
- 2. Temporal Reference System: Description of the temporal reference system(s) used in the data set.

This element is mandatory only if the spatial data set contains temporal information that does not refer to the default temporal reference system.

- 3. Encoding: Description of the computer language construct(s) specifying the representation of data objects in a record, file, message, storage device or transmission channel.
- Topological Consistency: Correctness of the explicitly encoded topological characteristics of the data set as described by the scope.

This element is mandatory only if the data set includes types from the Generic Network Model and does not assure centreline topology (connectivity of centrelines) for the network.

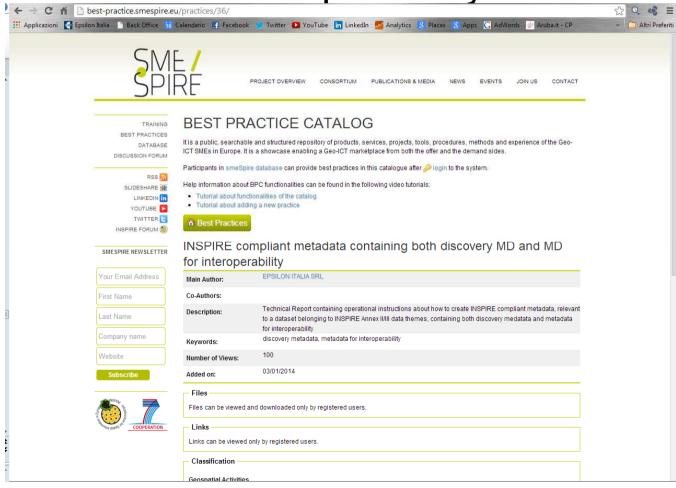
5. Character Encoding: The character encoding used in the data set.

This element is mandatory only if an encoding is used that is not based on UTF-8.

Spatial Representation Type: The method used to spatially represent geographic information.

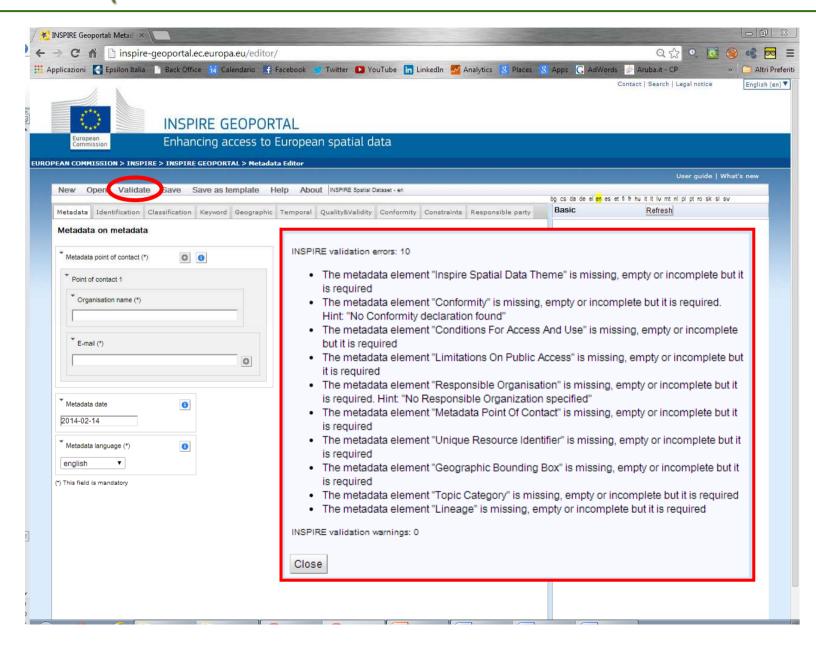


How to create INSPIRE compliant metadata containing both discovery metadata and metadata for interoperability?

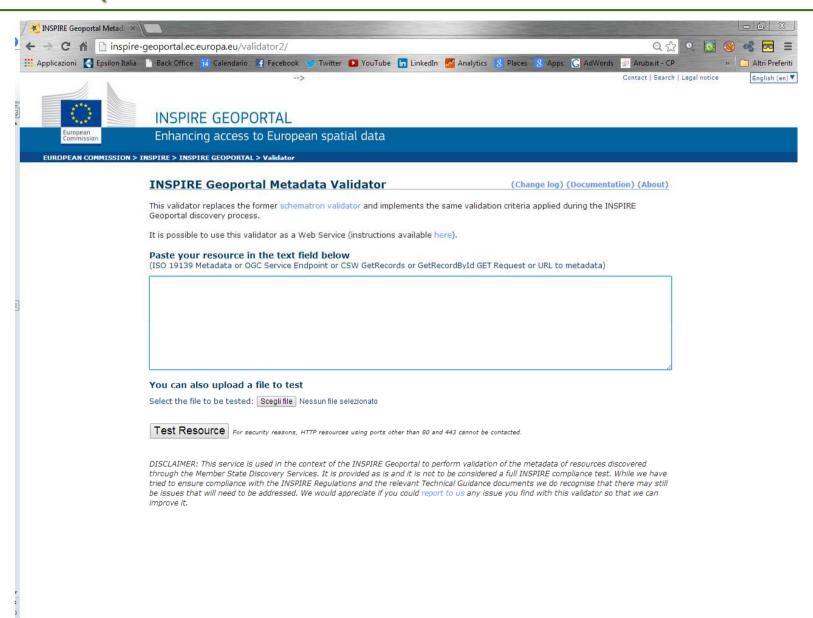




## **INSPIRE Metadata Editor**









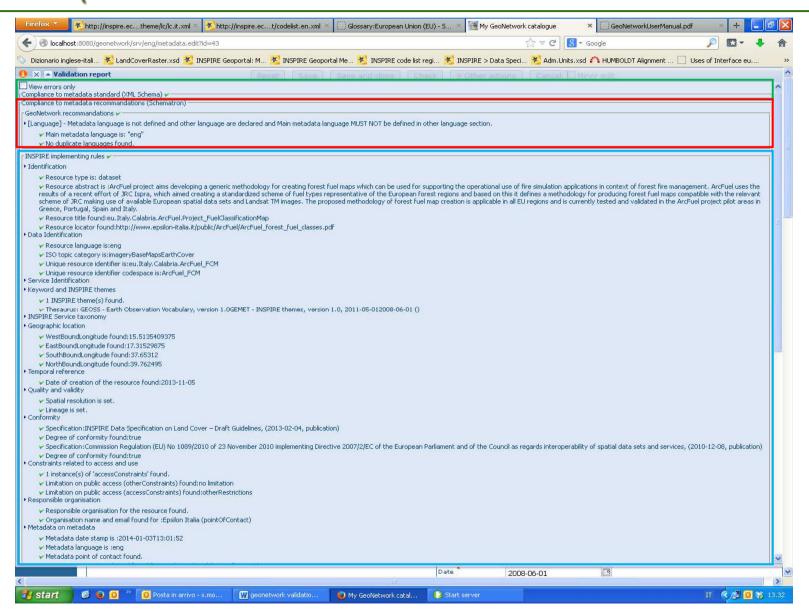
In order to insert the six additional metadata elements for interoperability defined by the regulation 1089/2010, namely:

- Coordinate reference system (mandatory)
- Temporal reference system
- Encoding (mandatory)
- Character encoding
- Spatial representation type (mandatory)
- Data Quality Logical consistency Topological consistency you may use Geonetwork

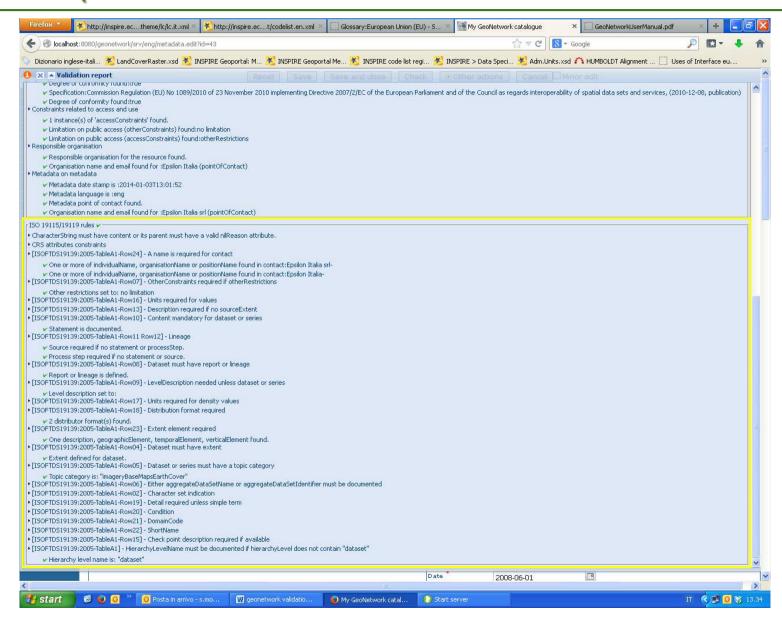


Default view		Reset Save	Save and close Check ⊕ Oth	er actions Cancel Mi	inor edit
INSPIRE view  By Group					Parent/child metadata:  Add or update parent metadata section
ISO Minimum ISO Core		No pre availa			Related service metadata:
ISO All By Package Metadata					Related feature catalogues:  Link feature catalogue
Identification Maintenance					
Constraints Spat. Info	▼IDENTIFICATION INFO				
Ref. system	Title *	eu.Italy.Calabria.ArcFuel.Proje	t_FuelClassif		
Distribution  Data quality	Date * Date type *	Creation 🔻			
App. schema	Unique resource identifier	eu.Italy.Calabria.ArcFuel FCM			
Catalog	Codespace 🗵	ArcFuel_FCM	(Suggestions:		
Content Info Ext. Info XML view	Abstract *	methodology for creating fores which can be used for supporti operational use of fire simulati applications in context of fores management. ArcFuel uses the recent effort of JRC Ispra, which creating a standardized schem types representative of the Eur forest regions and based on the	ArcFuel project aims developing a generic methodology for creating forest fuel maps which can be used for supporting the operational use of fire simulation applications in context of forest fire management. ArcFuel uses the results of a recent effort of JRC Ispra, which aimed creating a standardized scheme of fuel types representative of the European forest regions and based on this it defines a methodology for producing forest fuel		
	r ▼ Point of contac	74 1 1 5			
	Organisation name Enc		Electronic mai	il info@epsilon-italia.it	
	×	nt of contact	address ⊞⊠	ini og opsnon reanane	
		int of contact			
	Graphic overview ±				
	Descriptive ke				
	Keyword * ⊕ Fue	Load/Characteristic	and the second s	ırus name 🗵 ———————————————————————————————————	
			Title * Date *	GEOSS - Earth Observ	ation Vocabulary, ver
			Date type *	Publication 🕶	
	_ Descriptive ke	ywords ⊠ ■ 🗨 —			
		d Cover Fuel Classification Map			
	Descriptive ke	ywords 🛍 🗉 🗵 🗷 📉 💮 💮 💮			
	Keyword * ⊕ Lan	d cover	r ▼ Thesau	ırus name 🗵 —————	
			Title *	GEMET - INSPIRE then	nes, version 1.0
			Date *	Publication 🗸	
In attesa di localhost			Date type T	Publication 💌	











- Abstract Test Suite (ATS) included in Annex A to all the D2.8.II/III.x -Data Specifications on Annex II/III data themes Technical Guidelines, v3.0, published the 10th of December 2013
  - Part 1 (normative) IR Requirements
  - Part 2 (informative) TG Requirements



Conformance Class	Tests
	A.1.1 Schema element denomination test
	A.1.2 Value type test
A.1 Application Schema Conformance	A.1.3 Value test
	A.1.4 Attributes/associations completeness test
Class	A.1.5 Abstract spatial object test
	A.1.6 Constraints test
	A.1.7 Geometry representation test
	A.2.1 Datum test
A 2 Deference Cyclema Conforman	A.2.2 Coordinate reference system test
A.2 Reference Systems Conformance	A.2.3 Grid test
Class	A.2.4 View service coordinate reference system test
Class	A.2.5 Temporal reference system test
	A.2.6 Units of measurements test
	A.3.1 Unique identifier persistency test
	A.3.2 Version consistency test
A.3 Data Consistency Conformance Class	A.3.3 Life cycle time sequence test
	A.3.4 Validity time sequence test
	A.3.5 Update frequency test
A.4 Data Quality Conformance Class	
A.5 Metadata IR Conformance Class	A.5.1 Metadata for interoperability test
	A.6.1 Code list publication test
A.6 Information Accessibility	A.6.2 CRS publication test
Conformance Class	A.6.3 CRS identification test
Comormance etass	A.6.4 Grid identification test
A.7 Data Delivery Conformance Class	A.7.1 Encoding compliance test
A.8 Portrayal Conformance Class	A.8.1 Layer designation test



#### Part 1

(normative)

### Conformity with Commission Regulation No 1089/2010

### A.1 Application Schema Conformance Class

#### Conformance class:

http://inspire.ec.europa.eu/conformance-class/ir/ef/as/<application schema namespace prefix>

#### A.1.1 Schema element denomination test

- a) <u>Purpose</u>: Verification whether each element of the dataset under inspection carries a name specified in the target application schema(s).
- b) Reference: Art. 3 and Art.4 of Commission Regulation No 1089/2010
- c) <u>Test Method</u>: Examine whether the corresponding elements of the source schema (spatial object types, data types, attributes, association roles, code lists, and enumerations) are mapped to the target schema with the correct designation of mnemonic names.

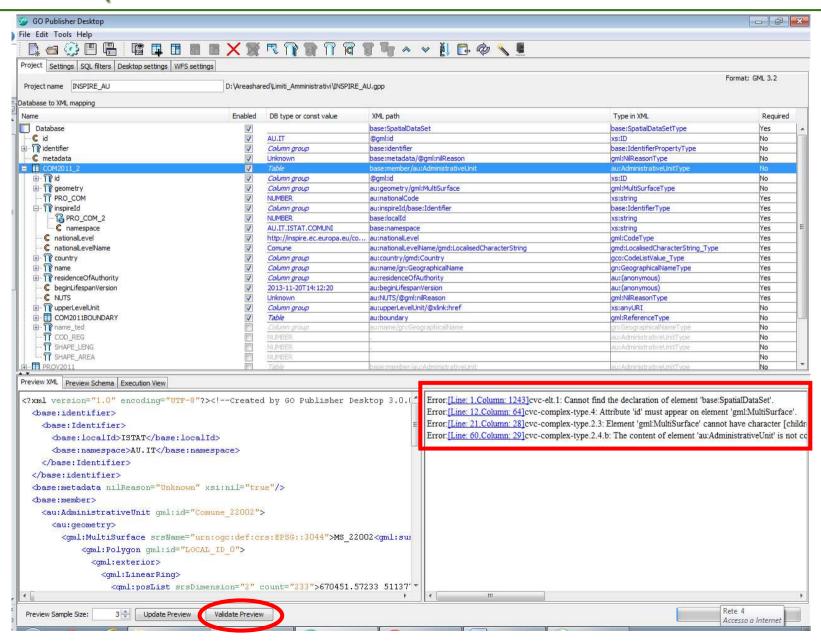
NOTE Further technical information is in the Feature catalogue and UML diagram of the application schema(s) in section 5.2.



Conformance Class	Tests
	A.1.1 Schema element denomination test
	A.1.2 Value type test
A.1 Application Schema Conformance	A.1.3 Value test
	A.1.4 Attributes/associations completeness test
Class	A.1.5 Abstract spatial object test
	A.1.6 Constraints test
	A.1.7 Geometry representation test
	A.2.1 Datum test
A 2 Deference Confermence	A.2.2 Coordinate reference system test
A.2 Reference Systems Conformance	A.2.3 Grid test
Class	A.2.4 View service coordinate reference system test
Class	A.2.5 Temporal reference system test
	A.2.6 Units of measurements test
	A.3.1 Unique identifier persistency test
	A.3.2 Version consistency test
A.3 Data Consistency Conformance Class	A.3.3 Life cycle time sequence test
	A.3.4 Validity time sequence test
	A.3.5 Update frequency test
A.4 Data Quality Conformance Class	
A.5 Metadata IR Conformance Class	A.5.1 Metadata for interoperability test
	A.6.1 Code list publication test
A.6 Information Accessibility	A.6.2 CRS publication test
Conformance Class	A.6.3 CRS identification test
Comormance class	A.6.4 Grid identification test
A.7 Data Delivery Conformance Class	A.7.1 Encoding compliance test
A.8 Portrayal Conformance Class	A.8.1 Layer designation test

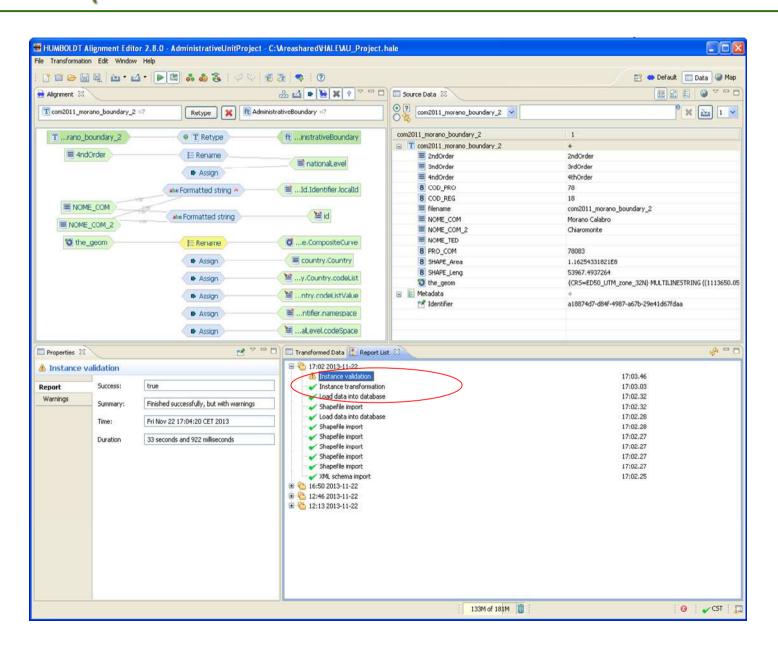


## Dataset validation tools

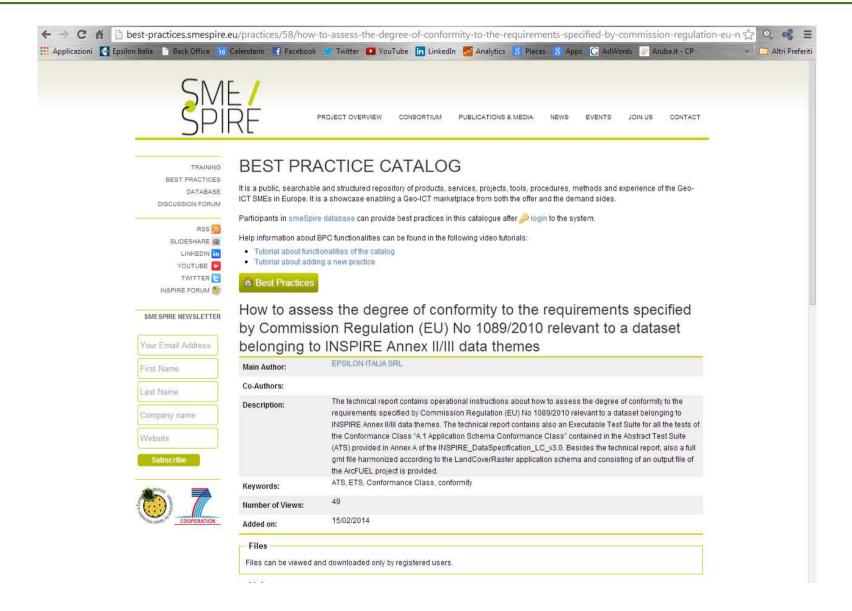




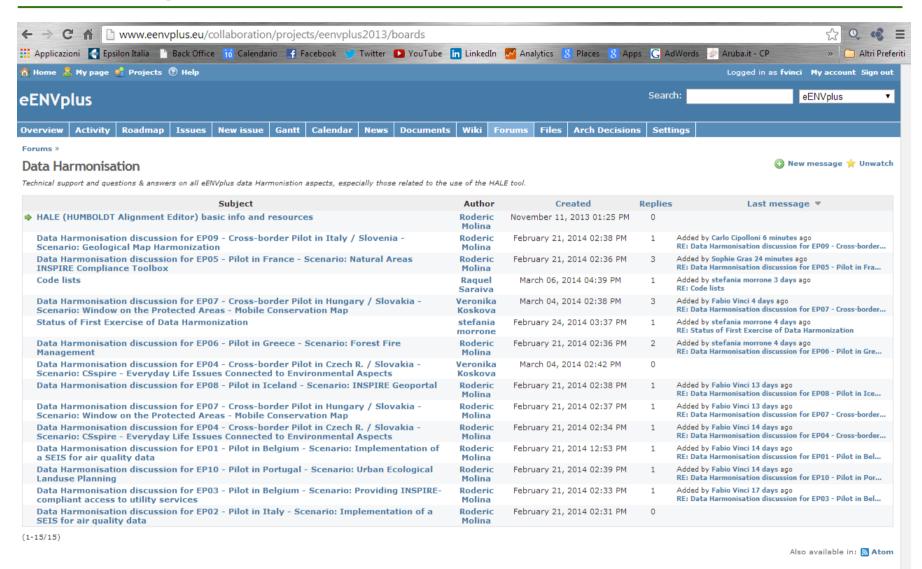
## Dataset validation tools





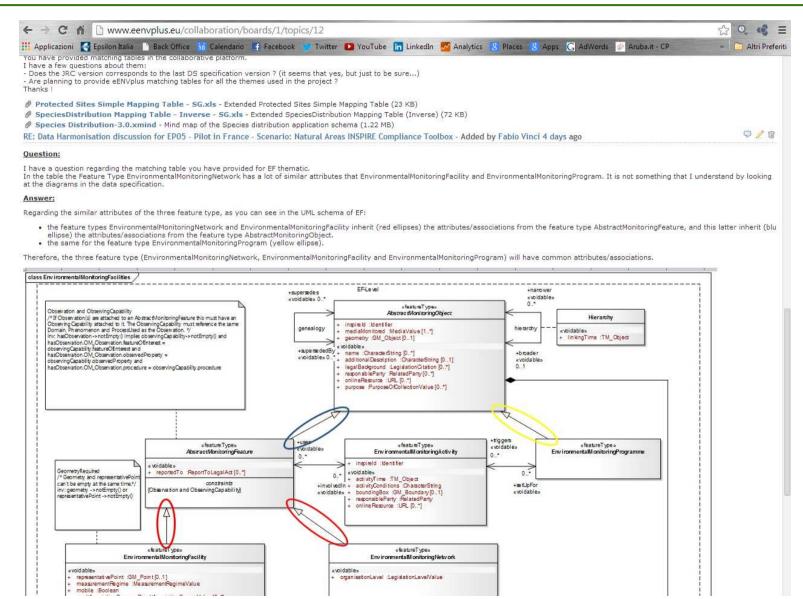






Powered by Redmine © 2006-2013 Jean-Philippe Lang

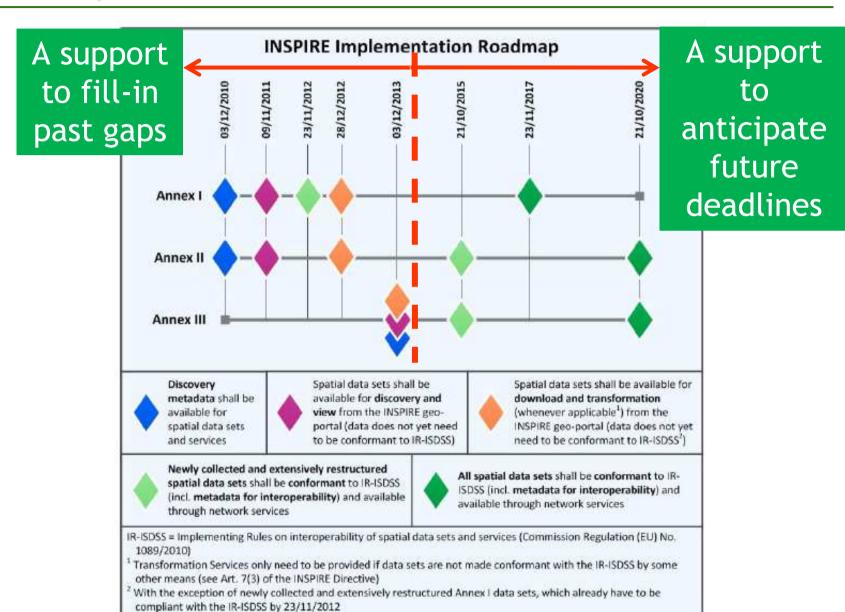






# Why?







# Thank you