



Harmonização da Carta de Ocupação do Solo utilizando o HALE

Teresa Zuna, Alexandra Fonseca,
Danilo Furtado, Ana Luísa Gomes,
André Serronha, Paulo Patrício



Introdução

A DGT é a entidade responsável pela coordenação operacional do Sistema Nacional de Informação Geográfica (SNIG) e o Ponto de Contacto Nacional (PCN) para a Diretiva INSPIRE.

A DGT tem participado em projetos europeus no domínio das infraestruturas de informação geográfica especialmente focados na harmonização de dados geográficos requerida pela Diretiva INSPIRE (e.g. HUMBOLDT, GIS4EU, NatureSDIplus, HELM, eENVplus, EAGLE 6)

Introdução

A DGT é também responsável pela produção de alguma cartografia temática, como é o caso da Carta de Ocupação do Solo (COS).

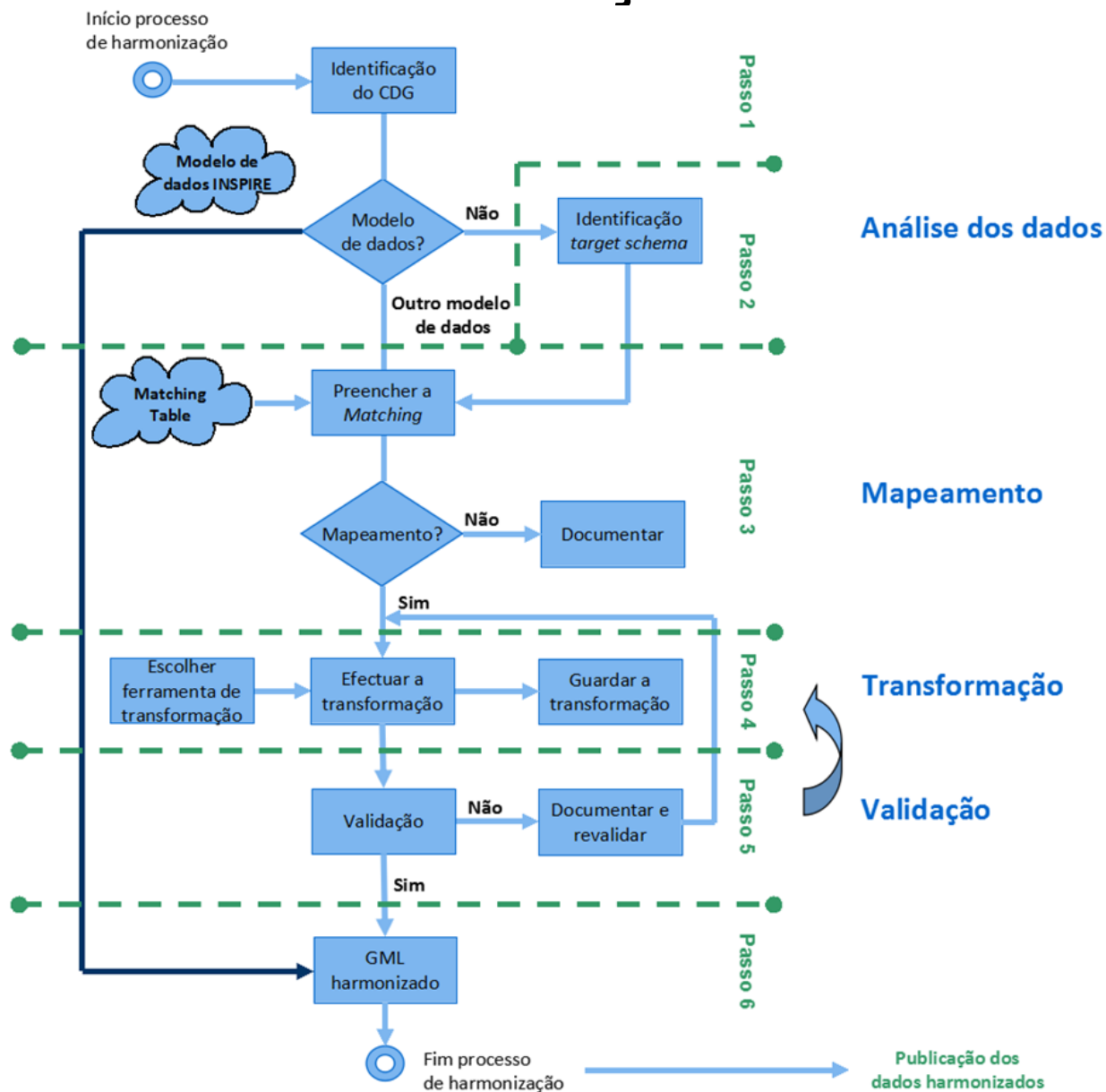
Sendo a Carta de Ocupação do Solo de Portugal Continental (COS) um dos CDG produzidos pela DGT com maior relevância para o desenvolvimento de estudos em gestão e planeamento ambiental em Portugal e na sequência da participação da DGT no projeto EAGLE 6, foi decidido avançar para a harmonização da COS de acordo com as especificações da Diretiva INSPIRE.

Carta de Ocupação do Solo (COS)



Processo de harmonização

O processo de harmonização envolve a análise dos modelos de dados, o preenchimento da *matching table* ou tabela de correspondências, a transformação do CDG para o *target schema*, a validação e a publicação, através de serviços, do CDG



1. Análise do
CDG



2.
Mapeamento



3.
Transformação
do CDG



4. Validação

1. Análise do CDG

Análise do modelo de dados de origem (*source schema*), nomeadamente:

- Formato dos dados
- Representação espacial
- Atributos
- Sistema de Coordenadas
- Metadados

1. Análise do CDG

A **Carta de Ocupação do Solo 2010** é uma cartografia temática que pretende caracterizar com grande detalhe a ocupação e uso do solo no território de Portugal Continental.

A COS2010 apresenta uma estrutura hierárquica com cinco níveis de detalhe, o que corresponde a 225 classes ao nível mais detalhado (N5). Esta nomenclatura é compatível com a nomenclatura da CORINE Land Cover aos primeiros três níveis.

Modelo de dados	Vectorial
Estrutura de dados	Polígonos
Sistema de referência	ETRS89 (European Terrestrial Reference System 1989) PT-TM06
Unidade Mínima Cartográfica (UMC)	1 ha
Unidade Mínima Cartográfica Distância mínima entre linhas	20 metros
Nomenclatura	Nomenclatura hierárquica com 5 níveis de detalhe e 225 classes

1. Análise do CDG

A harmonização envolve a transformação dos dados de origem (*source schema*) num modelo de dados descrito pelas especificações da Diretiva (*target schema*) em formato aberto e orientado para os serviços, para tal é necessário:

- Identificação do tema na Directiva INSPIRE
- Interpretação dos documentos INSPIRE, nomeadamente:
 - *General Conceptual Model*
 - *Data Specifications*



Title	D2.8.11.2 INSPIRE Data Specification on Land cover – Draft Technical Guidelines
Creator	INSPIRE Thematic Working Group Land cover
Date	2013-02-04
Subject	INSPIRE Data Specification for the spatial data theme Land cover
Publisher	INSPIRE Thematic Working Group Land cover
Type	Text
Description	This document describes the INSPIRE Data Specification for the spatial data theme Land cover. This version (version 3, release candidate 3) reflects the content of the draft amendment to Commission Regulation (EU) No 1089/2010 for the Annex II-III spatial data themes as submitted to the INSPIRE Committee.
Contributor	Members of the INSPIRE Thematic Working Group Land cover
Format	Portable Document Format (pdf)
Source	
Rights	Public
Identifier	D2.8.11.2_v3 Doc3
Language	En
Relation	Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)
Coverage	Project duration

ANNEX 1

	Addresses		Geographical names
	Administrative units		Hydrography
	Cadastral parcels		Protected sites
	Coordinate reference systems		Transport networks
	Geographical grid systems		

ANNEX 2

	Elevation
	Geology
	Land cover
	Orthomagey

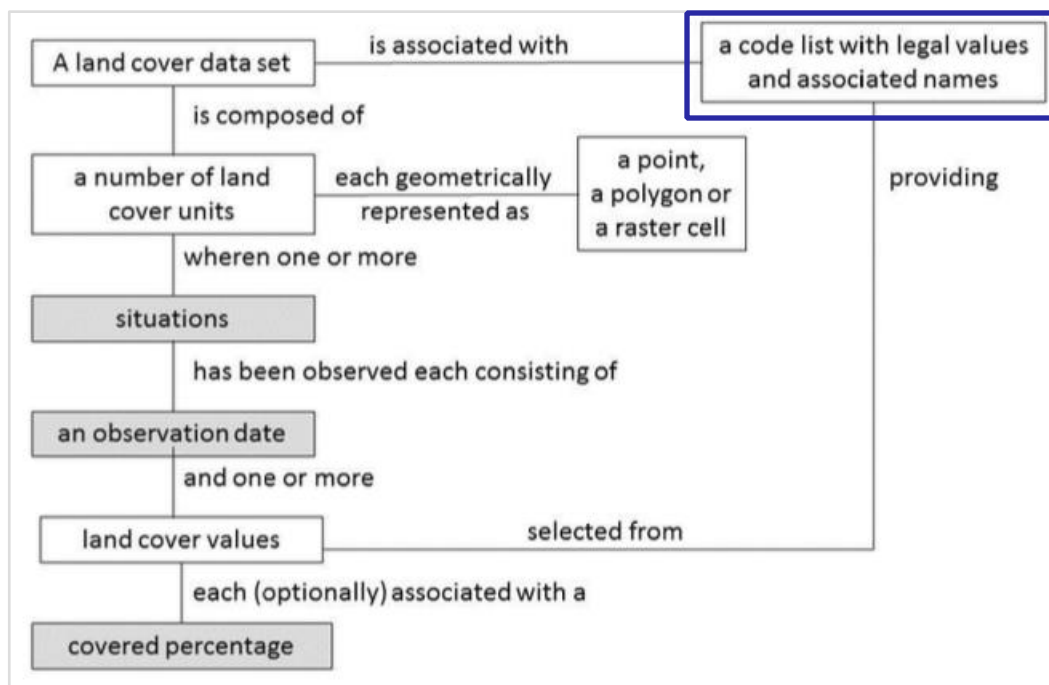
ANNEX 3

	Agricultural and aquaculture facilities		Habitats and biotopes		Population distribution and demography
	Area management / restriction / regulation zones & reporting units		Human health and safety		Production and industrial facilities
	Atmospheric conditions		Land use		Sea regions
	Bio-geographical regions		Meteorological geographical features		Soil
	Buildings		Mineral Resources		Species distribution
	Energy Resources		Natural risk zones		Statistical units
	Environmental monitoring Facilities		Oceanographic geographical features		Utility and governmental services

1. Análise do CDG

(Target data model)

- Tema INSPIRE – II.2 Ocupação do Solo
 - D2.5. INSPIRE Generic Conceptual Model versão 3.4rc3
 - D2.8.II.2 *Data Specification on Land Cover – Technical Guidelines*
 - *Application schema (LandCoverVector.XSD)*



Nomenclature Code List
URI <http://registo.igeo.pt/>

1. Análise dos dados

LandCoverNomenclature

Code list - é uma nomenclatura de classes de ocupação do solo, onde cada uma das classes é representada por um código e uma descrição

- É permitida apenas uma nomenclatura para a COS 2010
- Os valores são geridos fora da application schema *LandCoverVector.xsd*
- Lista hierárquica (parent value)
- INSPIRE register (http://inspire.ec.europa.eu/codelist_register/codelist)
- URI, legenda: <http://registo.igeo.pt/listadecodigo/CartaOcupacaoSoloValue>



The screenshot shows the INSPIRE Registry website. The header includes the European Commission logo and the text "INSPIRE REGISTRY Enhancing access to European spatial data". The main content area displays the "INSPIRE registry" details, including the ID, Label, Content Summary, Registry manager, and Other formats. A table of registers is shown at the bottom, listing various INSPIRE registers and their labels.

INSPIRE registry

ID: <http://inspire.ec.europa.eu/registry>

Label: **INSPIRE registry**

Content Summary: The INSPIRE infrastructure involves a number of items, which require clear descriptions and the possibility to be referenced through unique identifiers. Examples for such items include INSPIRE themes, code lists, application schemas or discovery services. Registers provide a means to assign identifiers to items and their labels, definitions and descriptions (in different languages). The INSPIRE registry provides a central access point to a number of centrally managed INSPIRE registers. The content of these registers are based on the INSPIRE Directive, Implementing Rules and Technical Guidelines.

Registry manager: **European Commission, Joint Research Centre**

Other formats: [XML \(ISO 19135\)](#) [XML beta \(ISO 19135\)](#) [RDF/XML beta \(ISO 19135\)](#) [JSON](#) [Atom](#) [CSV beta](#)

Registers

Label	Label
INSPIRE application schema register	
INSPIRE code list register	
INSPIRE feature concept dictionary	
INSPIRE glossary	
INSPIRE metadata code list register	
INSPIRE reference document register	
INSPIRE theme register	

Showing 1 to 7 of 7 entries

The INSPIRE Registry has been developed under Action 1.17 of the ISA Programme: A Reusable INSPIRE Reference Platform. Click here for more details.

1. Análise do
CDG



2.
Mapeamento



3.
Transformação
do CDG



4. Validação

2. Mapeamento

As tabelas de correspondência, designadas por *matching tables* são utilizadas para estabelecer correspondência entre os atributos contidos no modelo de dados de origem (*source schema*) e a estrutura do modelo dos dados de destino (*target schema*). A *matching table* identifica e descreve as classes, os atributos, as enumerações e listas de códigos e associações entre as classes de ambos os modelos

Application Schema 'LandCoverVector' (version 3.0)						Application Schema <COS2010_NS>									
Feature type		Feature type description		Feature type definition		Stereotype		Inspire theme		Dataset		Dataset definition			
LandCoverDataset		This representation allows Land Cover data being supported by a vector geometry						Land Cover Vector		COS2010_NS		COS 2007			
Application schema	Documentation	Attribute / Association role / Constraint	Attribute / Association role / Constraint documentation	Values / Enumerations	Multiplicity	Voidable / Non-Voidable	Attribute name	Documentation	Attribute / Association role / Constraint	Attribute / Association role / Constraint documentation	Values / Enumerations	Multiplicity	Voidable / Non-Voidable	Status	Feature type
gmBase	The attribute gpt for supports geometries in the database.	id		gpt:Id	1		gpt:Id				PT_COS2010NS_1_*			Not available	
LandCoverUnit	An individual element of the LC dataset represented by a point or polygon. Every unit supports Land Cover information.	id		gpt:Id	1						PT_COS2010NS_LCU_1_*			Not available	
		inspireId	External object identifier of the namespace uniquely identifying the data source of the object.	namespace	1		id	external feature nr id			1, 1005			Not available	
		version	The identifier of the particular version of the spatial object.	version	1						PT_GEO_LC_COS2010_PTCON_NS		unpopulated	Not available	
		beginLifespanVersion	Date and time at which this version of the spatial object	Date/Time	1	voidable								Not available	
		endLifespanVersion	Date and time at which this version of the spatial object	Date/Time	0..1	voidable								Not available	
		geometry	Spatial representation of the Land Cover unit.	GM_Object	1		the_geom		polygon					1:1	
		landCoverObservation	Land cover information at a specific time and place.	LandCoverObservation	1..*										
LandCoverDataset															
Modelo Inspire LandCoverVector.xsd															
		exists	Indicates the status of the data set.	EA_Exists	1									Not available	
		name	Name of the Land Cover data set.	Character/String	1		the_geom		polygon					Not available	
		nomenclatureDocumentation	Information about the nomenclature used in this data set.	LandCoverNomenclature	1		file_name							Not available	
		validFrom	The time when the phenomenon started to exist in the real world.	Date	1	voidable								Not available	
		validTo	The time from which the phenomenon no longer exists in the real world.	Date	1	voidable								Not available	
		member	A Land Cover unit being part of the data set.	LandCoverUnit	1..*									Easy	
LandCoverObservation															
	Land Cover information observed at a specific time and place.														
		class	The assignment of a land cover class to a land cover unit through a classification.	LandCoverClassValue	1									1:1	
		isActive	List of classification values resulting into details a land cover unit, identified with a percentage.	LandCoverValue	1..*	voidable								1:1	
		observationDate	The observation date associated of an observation.	Date/Time	1	voidable								Not available	
LandCoverValue															
	Generic class supporting Land Cover value and percentage.														
		class	Assignment of a land cover feature object to a land cover class.	LandCoverClassValue	1									Not available	
		coveredPercentage	Proportion of the area covered by the land cover object being concerned with the classification value.	Integer	1	voidable								Not available	

Modelo Inspire
LandCoverVector.xsd

Carta de Ocupação do
Solo 2010 N5

2. Mapeamento

As *matching tables* são utilizadas para documentar o processo de harmonização, através do preenchimento de diversos campos onde são registados aspetos relacionados com o processo de harmonização.

Application schema	Documentation	Attribute/ Association role/ Constraint	Attribute / Association role / Constraint documentation	Values / Enumerations	Multiplicity	Voidable / Non-Voidable
gmlBase	The attribute gml:id supports provision of a handle for the	id		gml:id	1	
LandCoverUnit	An individual element of the LC dataset represented by a point or polygon. Every unit support Land Cover information.	id		gml:id	1	
		inspireId	External object identifier of the Namespace uniquely identifying the data source of the particular version of the spatial object.	localId		
		beginLifespanVersion	Date and time at which this version of the spatial object	DateTime	1	voidable
		endLifespanVersion	Date and time at which this version of the spatial object	DateTime	0..1	voidable
		geometry	Spatial representation of the Land Cover unit.	GM_Object	1	
		landCoverObservation	Land cover information at a specific time and place.	LandCoverObservation	1..*	
LandCoverDataset	A vector representation for Land Cover data. This representation allows Land Cover data being supported by a vector geometry.	inspireId	External object identifier of the spatial object. NOTE An	localId		
		beginLifespanVersion	Date and time at which this version of the spatial object	DateTime	1	voidable
		endLifespanVersion	Date and time at which this version of the spatial object	DateTime	0..1	voidable
		extent	Contains the extent of the data set.	EX_Extent	1	
		name	Name of the Land Cover data set.	CharacterString	1	
		nomenclatureDocumentation	Information about the nomenclature used in this data set.	LandCoverNomenclature	1	
		validFrom	The time when the phenomenon started to exist in the real world.	Date	1	voidable
		validTo	The time from which the phenomenon no longer exists in the real world.	Date	1	voidable
		member	A Land Cover Unit being part of the data set.	LandCoverUnit	1..*	
LandCoverObservation	Land Cover information interpreted at a specific time and place.	class	The assignment of a land cover class to a land cover unit through a classification	LandCoverClassValue	1	
		mosaic	List of classification values describing into details a land cover unit, associated with percentages.	LandCoverValue	1..*	voidable
		observationDate	The observation date associated of an observation	DateTime	1	voidable
LandCoverValue	Generic class supporting Land Cover value and percentage.					

gml: identificador único

Multiplicity

Voidable

Atributos

Application schemas

Descrição application schema

Member - LandCoverUnit

Descrição dos atributos

1. Análise do
CDG



2.
Mapeamento



3.
Transformação
do CDG



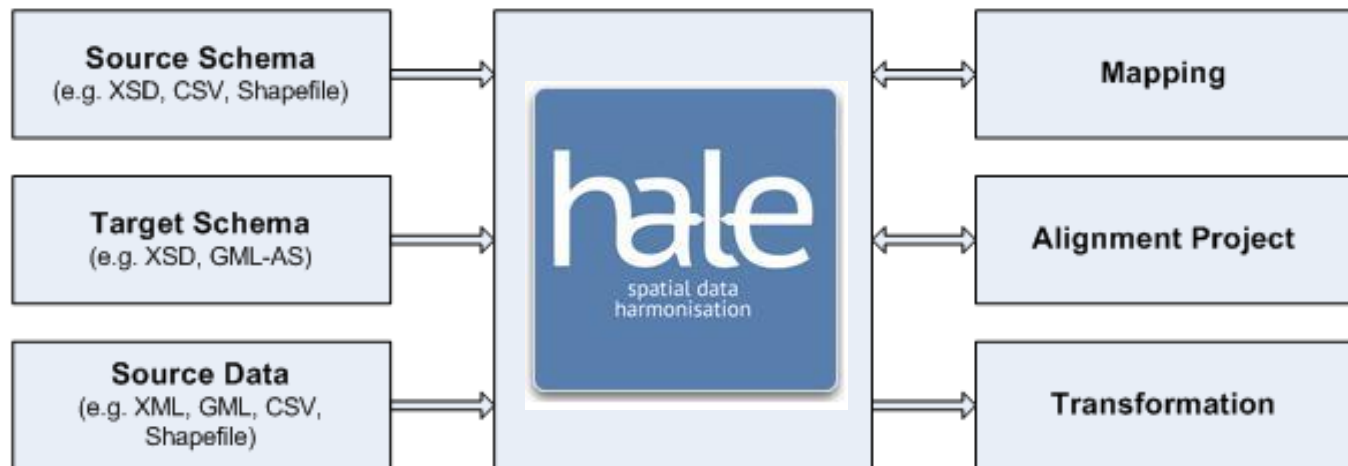
4. Validação

3. Transformação do CDG

HALE

O HUMBOLDT Alignment Editor (HALE) foi desenvolvido através do projeto europeu HUMBOLDT (www.esdi-humboldt.org) e pretende contribuir para a implementação da Diretiva INSPIRE.

O HALE é uma ferramenta Open Source, desenvolvida com o objetivo de dar suporte e de facilitar processos de harmonização e transformação de CDG. Permite ao utilizador, estabelecer relações entre schemas (*source* e *target*) e transformar CDG, de forma automática, com base nas especificações definidas nos application schemas.



3. Transformação do CDG

- **HALE:**
 - Exportação para GML 3.2.1
 - Adaptado à Directiva INSPIRE (*Code Lists, application schema ...*)
 - Transformação em tempo real com feedback
 - Validação online com o *application schema*
 - Permite a criação de scripts



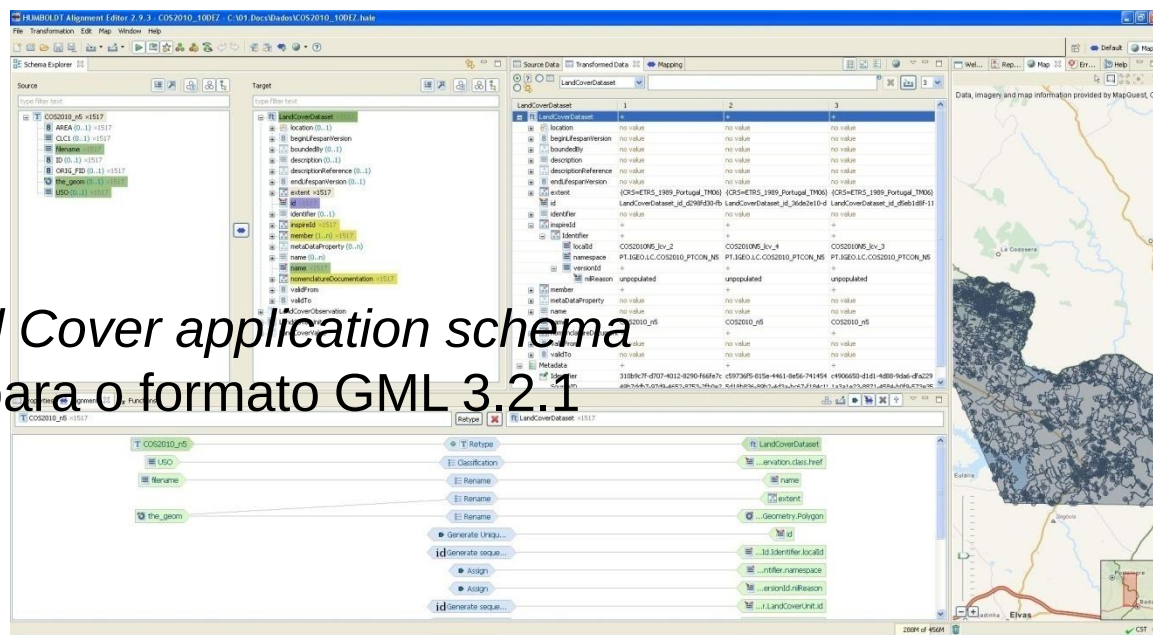
3. Transformação do CDG

HALE:

1. importar a *source* e *target* schemas
2. importar a *source data* (ter em conta a ISO 8859-1 para o alfabeto latim)
3. importar as *code list* nos formatos .XML ou .CSV
4. mapeamento entre as entidades do *source* e *target* schemas usando a *matching table*
5. estabelecer as relações de correspondência entre os *schemas*, utilizando as seguintes funções do HALE:

- *Re-type*
- *Assign*
- *Formatted string*
- *Rename*
- *Classification*

6. validação com o *Land Cover application* schema
7. exportar o resultado para o formato GML 3.2.1



1. Análise do
CDG



2.
Mapeamento



3.
Transformação
do CDG



4. Validação

4. Validação

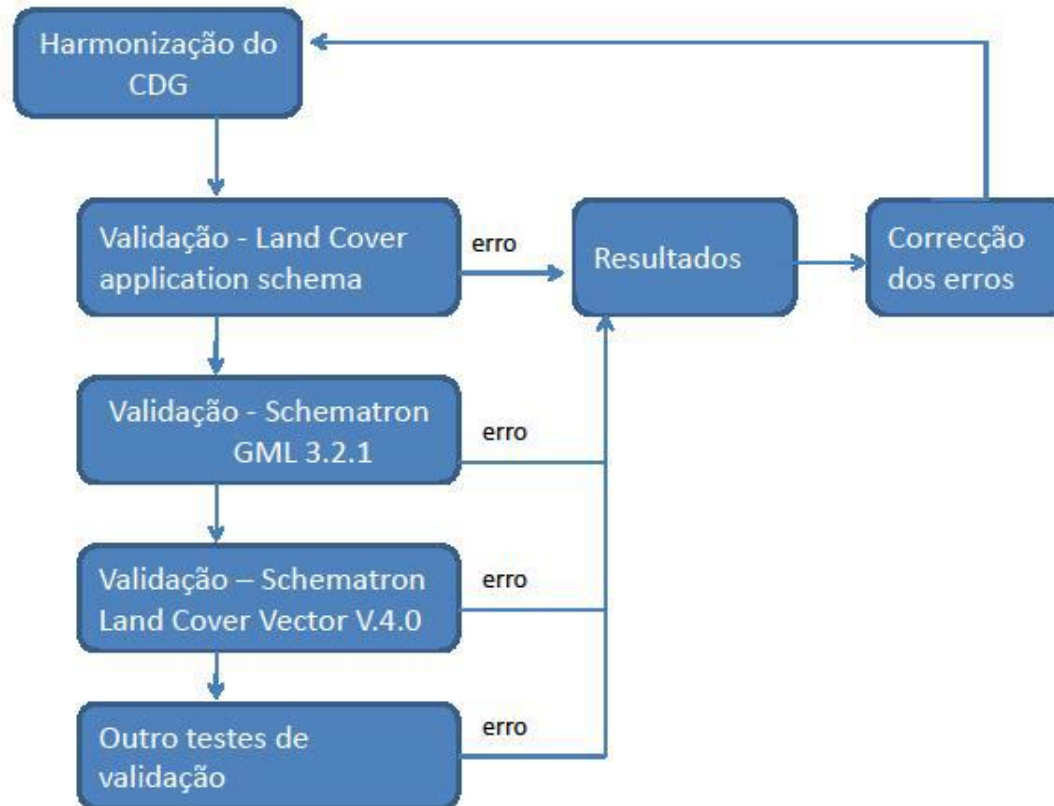
Abstract Test Suite (ATS): No Anexo A das especificações de dados

- Abstract Test Suite (ATS)
 - Grupo 1 – normativo
 - Grupo 2 – informativo
- Validação automática do GML
 - LandCoverVector.xsd
 - GML Schematron 3.2.1
 - LandCover Schematron 4.0
- Validação manual do GML
 - Verificação manual da existência no gml, das características especificadas pelos ATS

ATS	Conformance classes	Abstract Tests	Related ET
Part 1 (normative)	A.1 Application Schema Conformance Class	A.1.1 Schema element denomination test	E.1
		A.1.2 Value type test	E.1
		A.1.3 Value test *	E.1
		A.1.4 Attributes/Associations completeness test	E.1
		A.1.5 Abstract spatial object test	E.1
		A.1.6 Constraints test *	E.1
		A.1.7 Geometry representation test*	E.1
	A.2 Reference Systems Conformance Class	A.2.1 Datum test *	E.1
		A.2.2 Coordinate reference system test *	E.1
		A.2.3 Grid test	E.2
		A.2.4 View service CRS test	E.2
		A.2.5 Temporal reference system test	E.2
		A.2.6 Units of measurements test	E.2
	A.3 Data Consistency Conformance Class	A.3.1 Unique identifier persistency test	E.3
		A.3.2 Version consistency test	E.3
		A.3.3 Life cycle time sequence test*	E.1
		A.3.4 Validity time sequence test *	E.1
		A.3.5 Update frequency test	E.3
	A.4 Metadata IR Conformance Class	A.4.1 Metadata for interoperability test	E.4
	A.5 Information Accessibility Conformance Class	A.5.1 Code list publication test	E.5
		A.5.2 CRS publication test *	E.1
		A.5.3 CRS identification test *	E.1
		A.5.4 Grid identification test	E.5
	A.6 Data Delivery Conformance Class	A.6.1 Encoding compliance test	E.1
	A.7 Portrayal Conformance Class	A.7.1 Layer designation test	E.6
Part 2 (informative)	A.8 Technical Guideline Conformance Class	A.8.1 Multiplicity test	E.1
		A.8.2 CRS http URI test	E.7
		A.8.3 Metadata encoding schema validation test	E.8
		A.8.4 Metadata occurrence test	E.8
		A.8.5 Metadata consistency test	E.8
		A.8.6 Encoding schema validation test	E.1
		A.8.7 Coverage multipart representation test	E.9
		A.8.8 Coverage domain consistency test	E.9
		A.8.9 Style test	E.10

4. Validação

No processo de validação da COS 2010 relativamente às especificações de dados da Diretiva INSPIRE, foi seguida a seguinte metodologia:



4. Validação

A DGT esteve envolvida em 2015 num projeto para a Agência Europeia do Ambiente, **EAGLE 6**, que procedeu à harmonização de dados do CORINE Land Cover e Urban Atlas de acordo com a Diretiva INSPIRE.

Neste projecto, em parceria com a empresa **EPSILON Italia** colaborou na produção do *ficheiro **schematron Land Cover v.4***, disponível no validador do eENVplus



O eENVplus disponibiliza um serviço de validação online e gratuito, ([http://cloud.epsilon-italia.it/eenvplus new/ATS.htm?](http://cloud.epsilon-italia.it/eenvplus_new/ATS.htm?)), que permite a implementação dos ATS (*Abstract Test Suite*) incluídos no Anexo A das especificações de dados.

Este conjunto de testes executável (ETS) verifica a conformidade dos conjuntos de dados GML em relação aos *application schemas* e também em relação à ISO 19136:2007 (*schematron* GML 3.2.1).

Permite ainda a validação com ficheiros de tipo *schematron* para os temas já disponibilizados.

4. Validação

O validador **eENVplus** foi utilizado para a validação do GML da COS 2010 com o application schema Land Cover Vetor, o schematron GML 3.2.1 e o schematron Land Cover v.4:

TestNG Results

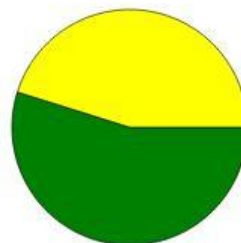
[Results overview](#)

[Reporter output](#)

gml32-3.2.1-r18	55%
0 Groups	
0 / 17 / 31	
All GML application schemas	
GML application schemas defining features and feature collections	
GML application schemas defining spatial geometries	
GML application schemas defining time	
GML application schemas defining spatial topologies	
GML Documents	

Test suites overview

■ Failed (%)
■ Passed (55%)
■ Skipped (45%)



eENVplus

gml32-3.2.1-r18	0	17	14	31	55%
All GML application schemas	0	7	0	7	100%
GML application schemas defining features and feature collections	0	2	0	2	100%
GML application schemas defining spatial geometries	0	2	0	2	100%
GML application schemas defining time	0	0	2	2	%
GML application schemas defining spatial topologies	0	0	2	2	%
GML Documents	0	6	10	16	38%

Vantagens do validador **eENVplus** :

- Guia metodológico sobre o processo de validação
- Recurso online
- Representação gráfica dos resultados

4. Validação

Como auxiliar de validação do GML transformado foi ainda utilizado o **oXygen**, um software de edição de XML proprietário, permitiu a visualização do ficheiro GML e também a validação com os seguintes esquemas:

- LandCoverVector.xsd
- GML Schematron 3.2.1
- Land Cover Schematron 4.0



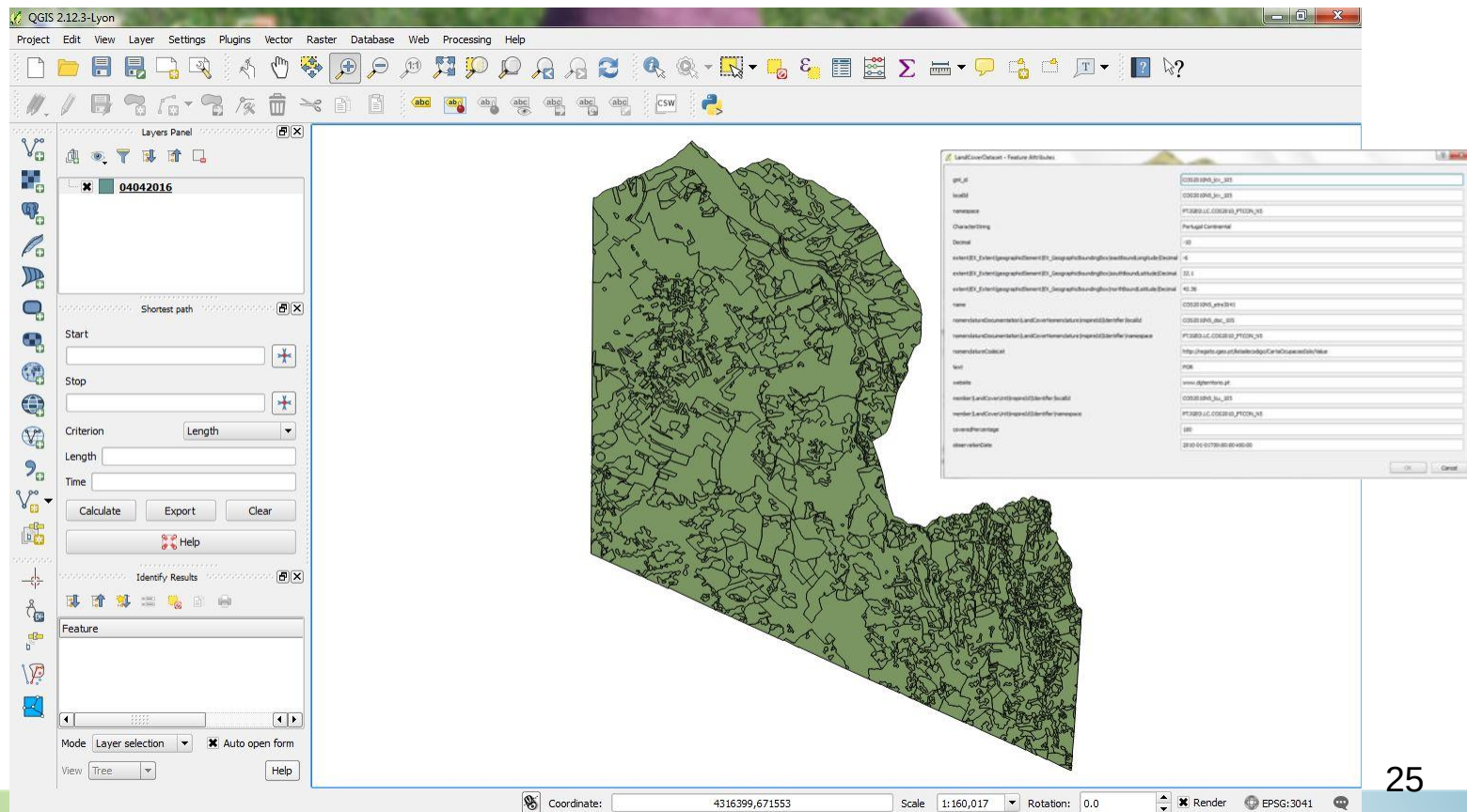
O **oXygen** XML editor ajudou a compreender os erros e a corrigi-los

The screenshot shows the oXygen XML editor with the following details:

- File Name:** 08MAR2016.gml
- Current File:** LandCoverVector.xsd
- XML Content:** The code starts with a GML FeatureCollection and includes a large block of coordinate data (lon, lat) for a specific area.
- Status Bar:** Shows 'Validation successful'.
- Interface Elements:** Includes a toolbar at the top, a sidebar on the left with 'Outline', 'Open Find Resource', and 'Archive Browser', and a bottom status bar with 'U+000A', '9: 83', and 'U+000A'.

Considerações finais

O principal objetivo deste trabalho foi a aplicação das DE relativas ao tema II.2 Land Cover da Diretiva INSPIRE à Carta de Ocupação do Solo de Portugal Continental de 2010 (COS), de forma a produzir um ficheiro harmonizado de acordo com as regras estabelecidas pela Diretiva INSPIRE. Tendo como resultado final do processo de harmonização um ficheiro GML válido.





OBRIGADO

Paulo Patrício

ppatricio@dgterritorio.pt

