



Infraestruturas de Dados Espaciais Harmonização de Dados Geográficos

9 e 23 novembro de 2017

Mestrado de Sistemas de Informação Geográfica – Tecnologias e Aplicações

Mestrado de Engenharia Geoespacial

FCUL, Lisboa

AULA(S) TEÓRICA(S)

André Serronha (DGT)

aserronha@dgterritorio.pt

Harmonização de Dados no âmbito da Diretiva INSPIRE

... e agora?

Conjuntos e Serviços
de Dados Geográficos

Harmonizar CSDG segundo os modelos de dados e as especificações INSPIRE...

De que necessita a Comissão Europeia?

(dados para «suprimir as necessidades» de 55 diretivas ambientais, etc...)

... e principalmente: de que necessita o nosso País?
Como nos temos que organizar?

... e o Cidadão?

- No âmbito da Diretiva INSPIRE, harmonização de dados é o processo de desenvolvimento de um conjunto comum de especificações de dados, que viabilize o acesso aos Conjuntos de Dados Geográficos (CDG), através de Serviços de Dados Geográficos (SDG), permitindo combinar os dados de forma coerente e interoperável.
- Neste módulo serão abordadas as diferentes fases do processo de harmonização de CDG, que envolve a análise dos modelos de dados (origem e destino), o mapeamento (*matching table* ou quadro de correspondências) entre os modelos, a transformação, a validação e a publicação dos CDGs (formato GML 3.2.1 e/ou em SDG).
 - Software a utilizar: QGIS (open source); hale STUDIO (open source); Gaia (gratuito)
 - É necessário acesso à Internet.

Glossário <http://inspire.ec.europa.eu/glossary> :

- - **dataset** -> <http://inspire.ec.europa.eu/glossary/DataSet> ;
- - **dataset series** -> <http://inspire.ec.europa.eu/glossary/DataSetSeries> ;
- - **dados geográficos** -> <http://inspire.ec.europa.eu/glossary/SpatialData> ;
- - **conjunto de dados geográficos** -> <http://inspire.ec.europa.eu/glossary/SpatialDataSet>

**“Uma coleção identificável de dados geográficos
com referência direta ou indireta a uma
localização ou zona geográfica específica”**

Example INSPIRE data specification: Hydrography (HY) – The World

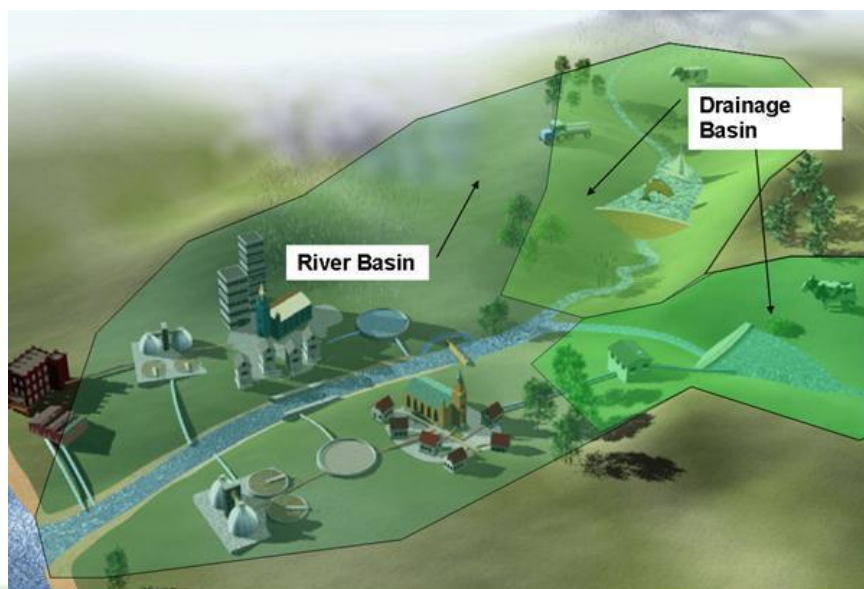
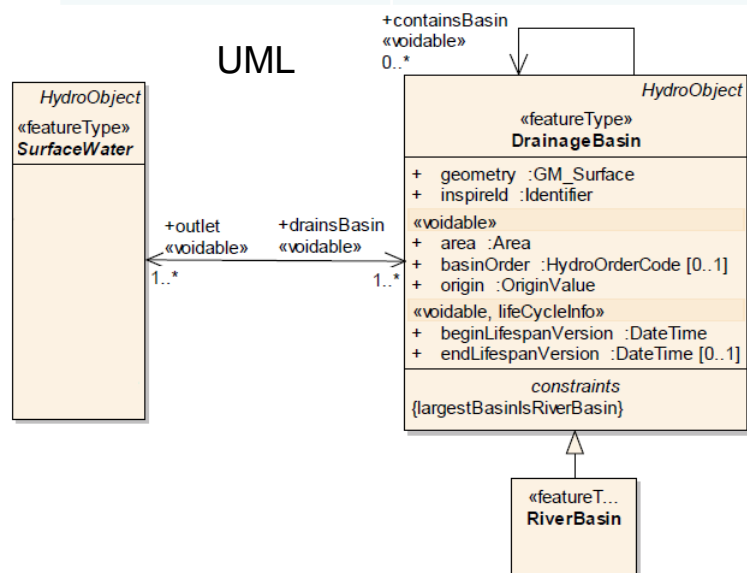


Example: HY – Mapping



Modelo de dados: Exemplo das Bacias (hy-p)

EN_label	PT_label	PT_definition
Drainage Basin	Bacia de drenagem	Área com uma saída comum para as suas escorrências superficiais.
River Basin	Bacia hidrográfica	A área terrestre a partir da qual todas as águas fluem, através de uma sequência de ribeiros, rios e eventualmente lagos, para o mar, desembocando numa única foz, estuário ou delta.



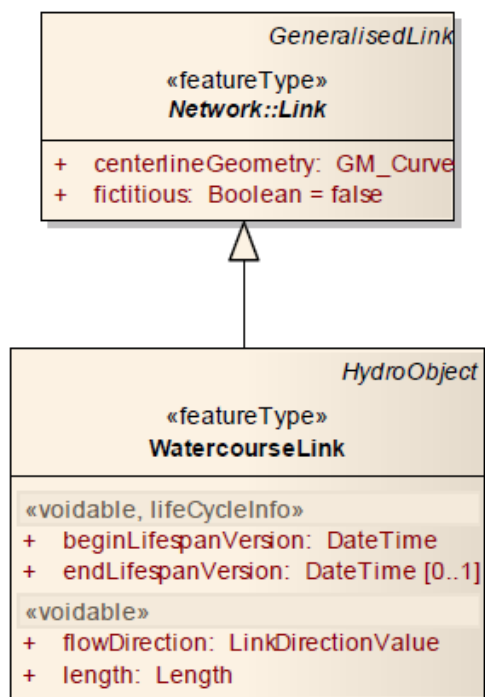
Example: HY – Network



Modelo de dados: Exemplo da Rede hidrográfica (hy-n)

EN_label	PT_label	PT_definition
Watercourse Link	Segmento de curso de água	Um segmento de um curso de água numa rede hidrográfica.

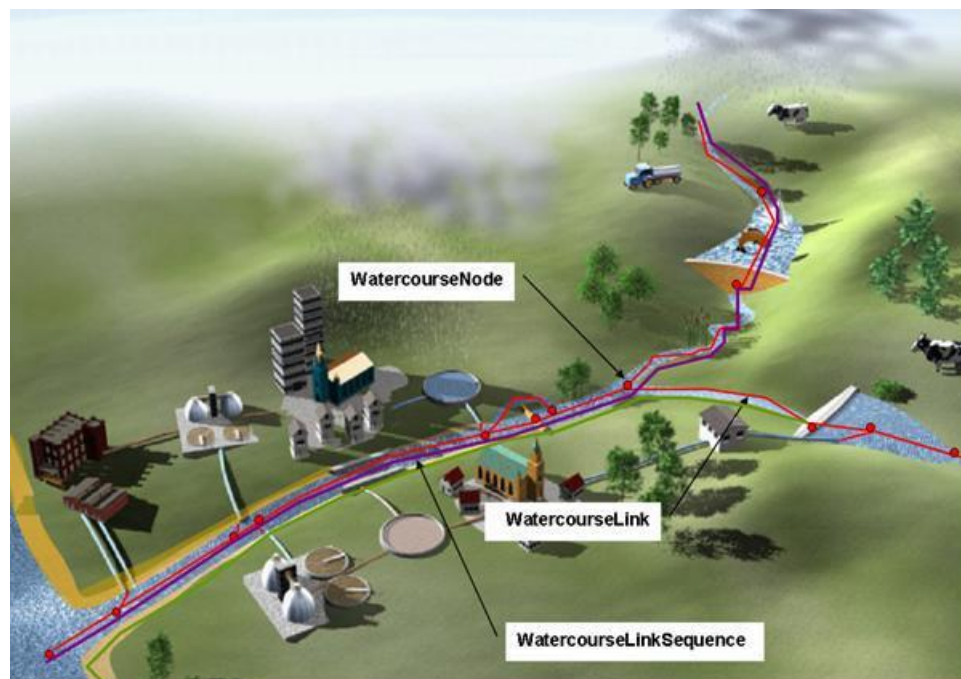
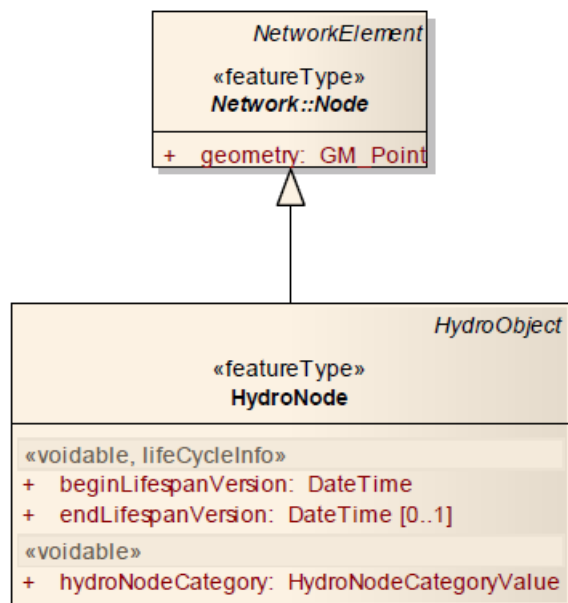
UML



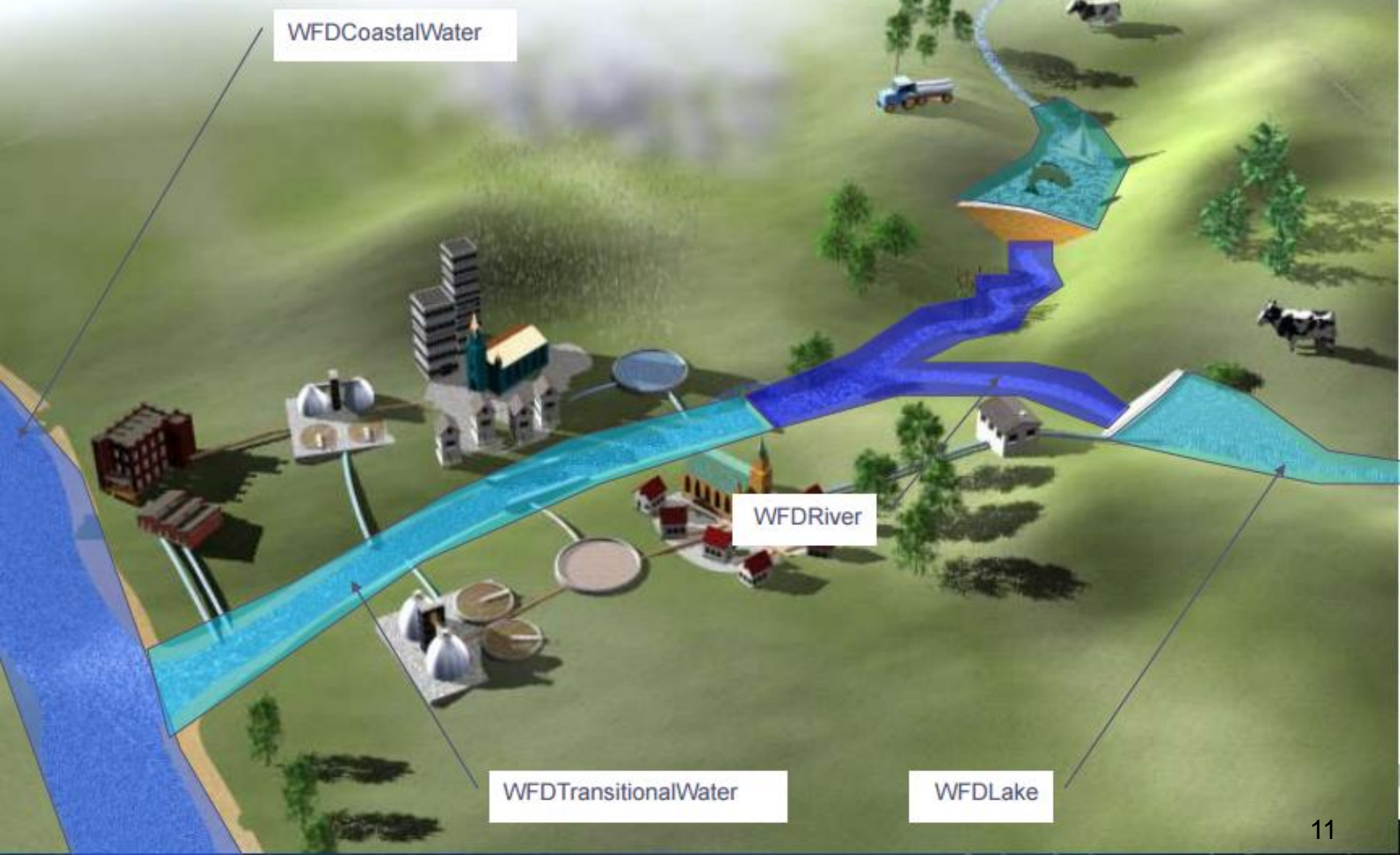
Modelo de dados: Exemplo de Nós hidrográficos (hy-n)

EN_label	PT_label	PT_definition
Hydro Node	Nó hidrográfico	Um nó na rede hidrográfica.

UML



Example: HY – Reporting



Example: HY – Complete Data Specification



Com a utilização da ferramenta *ShapeChange* (por exemplo) sistematiza-se a passagem entre o modelo UML e o meta-documento GML (application schema)



Linguagem UML e GML ?

- A linguagem UML (*Unified Modeling Language*) é uma linguagem de modelação aberta em que as entidades são modeladas por classes;
- A linguagem GML (*Geographic Markup Language*) foi desenvolvida para descrever especificamente informação geográfica (Lake, Burggraf, and Trninic, 2004).
 - As três características mais importantes:
 1. É uma linguagem de modelação para informação geográfica;
 2. É um formato XML para codificar informação geográfica;
 3. É um formato desenhado para a *web*, em particular para os *webservices*.
 - Este último ponto, fez com que fosse desenvolvida sobre um conjunto de normas do Consórcio W3C, como o XML, XML Schema, XLink e XPointer, com o objetivo de representar a informação geográfica num formato que maximizasse a sua divulgação através da Internet. O desenvolvimento do GML é ativamente promovido pelo consórcio OGC (*Open Geospatial Consortium* <http://www.opengeospatial.org>), criado com a finalidade de promover a interoperacionalidade dos sistemas que manipulam informação geográfica.

Fonte: <https://repositorium.sdum.uminho.pt/bitstream/1822/1243/1/tese.pdf>

Tool for data modelling

EnterpriseArchitect

Tool to derive FC, GML application schema, matching tables, etc. from EA UML model

ShapeChange

MODELO CONCEITUAL

É a próxima etapa do projeto de um sistema de aplicação em banco de dados.

Representa ou descreve a realidade do ambiente do problema, constituindo-se em uma visão global dos principais dados e relacionamentos, independente das restrições de implementação.

É uma descrição em alto nível (macro definição), mas que tem a preocupação de capturar e retratar toda a realidade de uma organização.

O resultado de um modelo conceitual é um esquema que representa a realidade das informações existentes, assim como as estruturas de dados que representam estas informações.

MODELO LÓGICO

Tem seu início a partir do modelo conceitual, levando em consideração três abordagens principais: Relacional (atualmente o mais utilizado), Hierárquica e Rede.

O modelo lógico descreve as estruturas que estarão contidas no banco de dados, mas sem considerar ainda nenhuma característica específica de SGBD, resultando em um esquema lógico de dados.

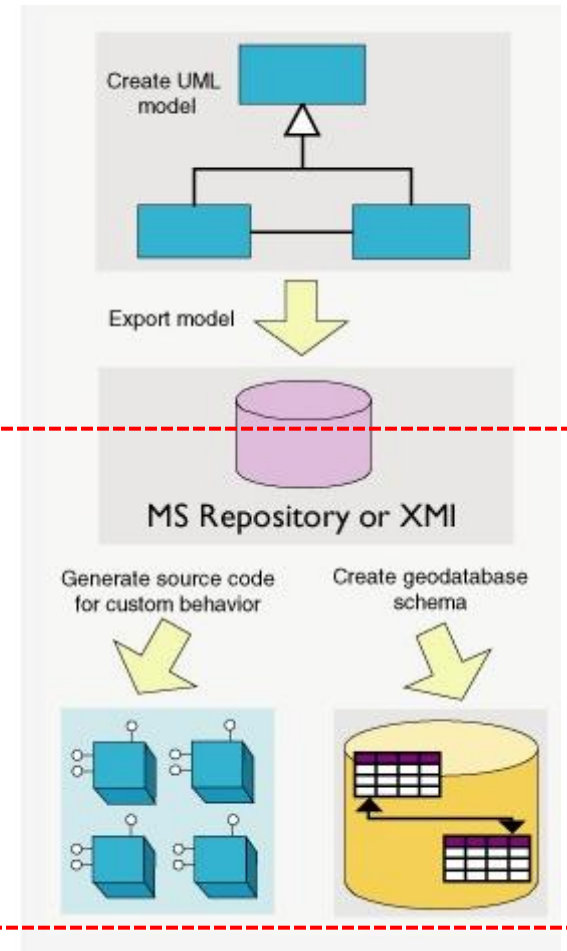
MODELO FÍSICO

Parte do modelo lógico e descreve as estruturas físicas de armazenamento de dados, tais como: tamanhos de campos, índices, tipos de dados, nomenclaturas, etc.

Este modelo detalha o estudo dos métodos de acesso do SGDB, para elaboração dos índices de cada informação colocada nos modelos conceitual e lógico.

É a etapa final do projeto de banco de dados, na qual será utilizada a linguagem de definição de dados (DDL), para a realização da montagem do mesmo no nível de dicionário de dados.

Na documentação do site INSPIRE apenas existem modelos UML para Enterprise Architect (EAP, XMI) e não existe nenhum modelo de dados físico feito a partir do modelo conceptual em UML, ou seja, não existe nenhum ficheiro XMI (XML Metadata Interchange) para nenhum SGBD de forma a gerar o schema da base de dados vazia (comercial ou open source) com todos os relacionamentos, estruturas físicas, etc...



- Os dados provêm de diferentes organizações, com diferentes formas de representar a informação geográfica, diferentes formatos e Sistemas de Referência Espacial (SRS), tornando a integração de dados uma tarefa complexa.
- As Infraestruturas de Informação Geográfica (IIG) ou Infraestruturas de Dados Espaciais (IDE) normalizadas contribuem para ultrapassar algumas complexidades do parágrafo anterior, fornecendo “Dados/Serviços” que possam ser úteis a uma comunidade de utilizadores que é cada vez mais dependente de dados geográficos.

- Uma IIG é um conjunto de tecnologias, políticas, acordos institucionais que facilitam a disponibilização e acesso a informação de natureza espacial.
- As IIG viabilizam a pesquisa, avaliação e exploração de informação geográfica por utilizadores diversos (administração pública, academia, centros I&D, empresas, ONGs e cidadãos)

SNIG2020: <http://snig.dgterritorio.pt/Inspire/documentos/ontheroad/SNIGontheroad-CCDR-Caetano.pdf>

❑ **Diretiva INSPIRE Estado de implementação e desenvolvimentos futuros (4ª Sessão Técnica RAA, 7 e 9 de junho 2016):**

❑ <https://youtu.be/QaOUktVqBz8>

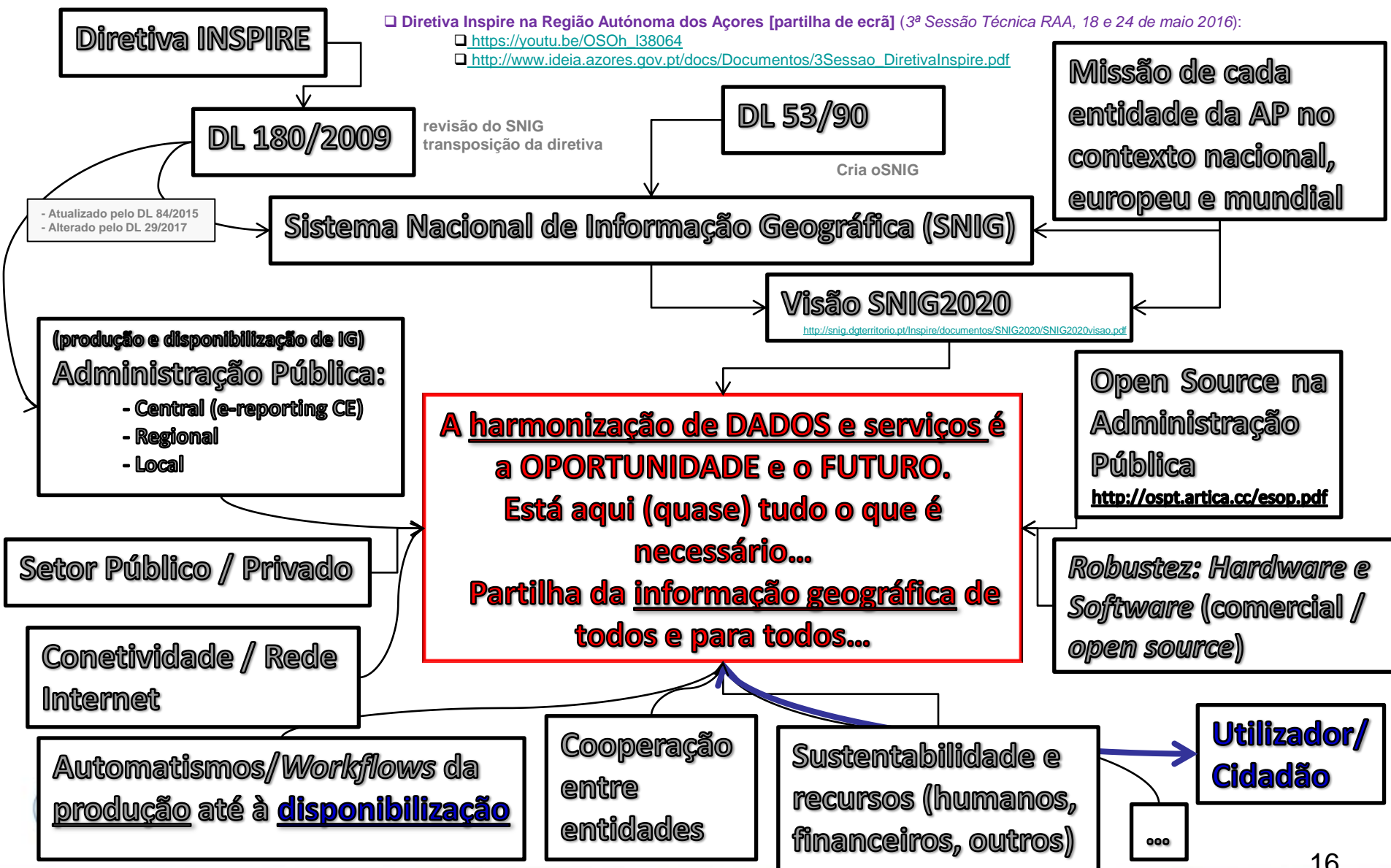
❑ http://www.ideia.azores.gov.pt/docs/Documentos/4SessaoTecnica_SNIGontheroad-azores-Caetano.pdf

❑ ENiiG2016

http://eniig.dgterritorio.pt/sites/default/files/ENiiG2016_MarioCaetano_DGT.pdf

Galeria (vídeos): http://www.dgterritorio.pt/fotos/1_encontro_nacional_de_infraestruturas_de_informacao_geografica_eniig_2016/

Harmonização de Dados no âmbito da Diretiva INSPIRE





Dificuldades?

Muitas e das mais variadas formas (harmonização de dados incluída “no bolo” deste processo):

- A necessidade da venda de informação geográfica é um entrave a políticas de dados abertas nos organismos. A sustentabilidade das instituições não pode ser posta em causa e tem de ser assegurada pela tutela.
- A informação geográfica continua a ser encarada como algo “acessório” e não fundamental. Ainda não existe a sensibilidade para aproveitar este veículo “chave” para o desenvolvimento (económico, etc.) do país.
- Diretiva que ainda não foi devidamente assimilada por todos os intervenientes (tutela, dirigentes das instituições e técnicos...) . A capacitação para a implementação da diretiva ainda está aquém das suas efetivas necessidades...
- Aspetos de índole política/organizacional levantam-se e ultrapassam muitas vezes as questões de capacidade técnica. Anos a fio de “desmantelamento” de recursos humanos, financeiros e *know-how* para estas temáticas na Administração Pública (AP).

Melhorou significativamente no ano de 2017

- A Comissão Europeia diz pouco (ainda estão em diversas fases de desenvolvimento) sobre como se deve fazer e validar. “Cada instituição/país” faz o que lhe parece melhor e com metodologias muito diferentes.

- Os dados existentes na AP (em geral) servem um universo pouco abrangente. Também por isso, muitas entidades produzem dados semelhantes sem cooperarem para finalidades comuns. Pouca “cultura de informação geográfica” de muitos dos intervenientes.

- A especificação de dados remete (muitas vezes) para informação geográfica que não existe na AP Central e, quando existe, é a um nível municipal, regional e empresarial em contextos muito diferentes de gestão e com muitas particularidades em que cada entidade serve os seus propósitos de forma desintegrada...

(...)

A “minha” instituição produz todos os dados que a diretiva INSPIRE pretende?

Sim. Porque...

- produz a geometria adequada e alguns atributos (produzidos na missão de âmbito nacional) conseguem encaixar nas especificações de dados INSPIRE através de um mapeamento (quase) direto dos valores dos atributos nacionais para os valores das *codelists* (listas de código) INSPIRE, ou através da extensão de uma *codelist*, quando é possível.

Não. Porquê? Porque “penso que”...

- a “minha” entidade apenas tem dados alfanuméricos;
- a “minha” entidade tem a geometria desadequada mas produz (de alguma forma) o(s) atributo(s) XPTO na sua missão;
- a “minha” entidade tem a geometria adequada mas o formato destes dados não permite saber onde começa e termina determinado atributo / propriedade / evento. (exemplo: dados ainda em formato CAD);
- a “minha” entidade tem a geometria adequada mas segmentada de forma desadequada daquela que se pretende na diretiva porque têm diferentes: definições, representações e objetivos.
- a “minha” entidade produz a geometria que se pretende mas não adquire nenhum dos atributos INSPIRE obrigatórios ou opcionais.



– etc...

Processo de Harmonização de CDG

Algumas das dificuldades

Diretiva INSPIRE

Problemas habituais na harmonização de dados:

- Formato dos dados;
- Sistemas de referência;
- Modelos conceptuais;
- Esquemas de classificação;
- Terminologia;
- Metadados;
- Representação espacial/ geometrias;
- Consistência topológica (fronteiras);
- Portrayal (estilos de representação);
- Multilinguismo.



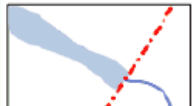
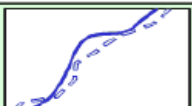
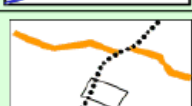
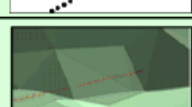
(A) INSPIRE Principles	(B) Terminology	(C) Reference model
(D) Rules for application Schemas and feature catalogues	(E) Spatial and temporal aspects	(F) Multi-lingual text and cultural adaptability
(G) Coordinate referencing and units model	(H) Object referencing modelling	(I) Identifier Management
(J) Data transformation	(K) Portrayal model	(L) Registers and registries
(M) Metadata	(N) Maintenance	(O) Quality
(P) Data Transfer	(Q) Consistency between data	(R) Multiple representations
(S) Data capturing	(T) Conformance	

... e outros:

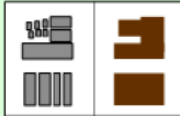

- Grande variedade de formatos de dados e duplicação dos mesmos, em que é quebrada a ligação com a principal fonte de dados;
- Muitas lacunas geográficas ainda permanecem;
- Problemas na captura, atualização, manutenção e qualidade dos dados;
- Modelos conceptuais quase inexistentes;
- Os sistemas de referência não estão harmonizados;
- Muitas fontes de dados não são consistentes;
- Escalas não são compatíveis;
- Os dados não são interoperáveis;
- Os custos e as restrições de acesso...

Processo de Harmonização de CDG

Algumas das dificuldades

Different Spatial representations		Limited capabilities - overlay of raster (orthoimage) and vector (roads) representations
Different representation geometries (3D vs.2D)		The same building represented in 3 and 2 dimensional geometries
Different planar representation geometries		The river is represented by a polygon on one side of a boundary, while on the other by the center line
Different boundaries		Possible causes: absence of agreement between authorities, measurement/transformation errors, different generalisation
Overlapping spatial objects and geometrical shift		Errors along a boundary presumably because of the different original projection systems
Inconsistency between data themes (Digital Elevation Model and Roads)		Violation of natural co-dependencies (the road crosses the land surface without a tunnel)

In: Toth K, Portele C, Illert A, Lutz M, Nunes de Lima V, 2012. A Conceptual Model for Developing Interoperability Specifications in Spatial Data Infrastructures. JRC Reference Report.

Examples of semantic differences		
Different aggregation level		The same real world entity is represented at different aggregation levels (houses vs. blocks)
Different classifications		The same entity differently classified at the two sides of a boundary (industrial zone vs. built-up area)

Dificuldades?

“Software CAD”:

- rapidez na aquisição dos dados;
- útil para cartografia topográfica de base/referência;
- ferramenta de desenho;
- ainda deve ser o mais utilizado (na administração pública em geral);
- estabilidade ao longo dos anos e grande conhecimento dos técnicos;
- realidade difícil de dissociar;
- ferramentas consolidadas;
- facilidade nos processos manuais e automatismos... rotinas CAD [AutoCAD, Microstation]:
 - macros, mdl (desenvolvido ao longo de muitos anos)...
 - ❑ estão implementadas muitas ferramentas relacionadas com:
 - ❖ Aquisição, manutenção e atualização de dados (controlo de qualidade);
 - ❖ Controlo de qualidade orientado para assegurar a consistência do produto (consistência do conteúdo dos ficheiros);
 - ❖ Simbologia... Regras de sobreposição... *Workflows* Vetor → Raster → Papel... etc...

MUITOS AUTOMATISMOS
CONSOLIDADOS AO LONGO
DOS ANOS NO PROCESSO
PRODUTIVO E DE CONTROLO
DE QUALIDADE DOS DADOS

TECNOLOGIA EM DESUSO NO
PARADIGMA ATUAL DA
INFORMAÇÃO GEOGRÁFICA E
DAS IDE

~~CAD~~ → “SIG” (BD ou outros
formatos)

IDE → CAD; “SIG”; outros

Dificuldades?

“Software SIG”:

- em constantes atualizações;
 - muita variedade;
 - velocidade de atualização de *software* (pode ser problemático para gerir processos);
 - implica muito acompanhamento;
- e constantes alterações aos automatismos bem como às instruções de trabalho das diversas tarefas a executar;

Algumas notas tendo em consideração a aquisição, manutenção, atualização de dados:

- Controlo de qualidade orientado para assegurar a consistência do produto (entre camadas);
- Gestão da vida dos elementos;
- Normas ISO;
- Simbologia para intervalos de escala:
 - Produto vetorial digital (WMS);
 - Produto vetorial digital no “SIG Desktop”;
- Simbologia para escala fixa adequada:
 - Mapas em papel, etc...

Recursos humanos/financeiros...

- Gestão/acompanhamento/evolução

É necessário “adequar /repensar” metodologias e processos automáticos para gerir estes “dados SIG” cada vez mais inseridos em IDE

Add-in ELF	24-06-2015 07:23	Pasta de ficheiros
ArcGISToolsWrappers.pyt	23-04-2015 14:36	ArcGIS Python Toolbox
ELFtools.pyt	29-05-2015 08:18	ArcGIS Python Toolbox
GeneraliseHydroHelper.pyt	22-05-2015 08:17	ArcGIS Python Toolbox
GeneraliseLandcoverHelper.pyt	15-01-2015 09:17	ArcGIS Python Toolbox
AdministrativeUnits.tbx	24-06-2015 07:22	ArcGIS Toolbox
CrossingsAndNetworkConnections.tbx	24-06-2015 07:22	ArcGIS Toolbox
ELF_Generalisation.tbx	24-06-2015 07:22	ArcGIS Toolbox
ELFBuilding.tbx	24-06-2015 07:22	ArcGIS Toolbox
GeneraliseHydro.tbx	22-06-2015 09:13	ArcGIS Toolbox
GeneraliseLandcover.tbx	24-06-2015 07:22	ArcGIS Toolbox
NetworkGeneralisation.tbx	24-06-2015 07:22	ArcGIS Toolbox
Partitioning.tbx	24-06-2015 07:22	ArcGIS Toolbox
Preprocessing.tbx	22-06-2015 09:13	ArcGIS Toolbox
Prototype.tbx	24-06-2015 07:22	ArcGIS Toolbox
Railways.tbx	22-06-2015 09:13	ArcGIS Toolbox
Utilities.tbx	22-06-2015 09:13	ArcGIS Toolbox
WaterTransportNetworks.tbx	22-06-2015 09:13	ArcGIS Toolbox
ArcGISToolsWrappers.AddLocationWrapper.pyt.xml	13-04-2015 08:29	Documento XML
ArcGISToolsWrappers.AppendWrapper.pyt.xml	21-05-2015 11:25	Documento XML
ArcGISToolsWrappers.BuildNetworkWrapper.pyt.xml	13-04-2015 08:29	Documento XML
ArcGISToolsWrappers.MakeRouteLayerWrapper.pyt.x...	13-04-2015 08:29	Documento XML
ArcGISToolsWrappers.pyt.xml	28-04-2015 13:27	Documento XML
ArcGISToolsWrappers.UnsplitLineWrapper.pyt.xml	28-04-2015 13:27	Documento XML
ELFtools.AnalyseJunction.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.CalculateMaxByIdentical.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.ConcatenateAttributes.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.DetectGraphicConflictWrapper.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.Exists.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.IfOnRowCount.pyt.xml	02-06-2015 13:16	Documento XML
ELFtools.KeepField.pyt.xml	29-05-2015 08:22	Documento XML
ELFtools.pyt.xml	29-05-2015 08:22	Documento XML
ELFtools.SelectDataIfExists.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.SnapWrapper.pyt.xml	02-06-2015 13:18	Documento XML
ELFtools.VoronoiPolygons.pyt.xml	02-06-2015 13:18	Documento XML
GeneraliseHydroHelper.AddLastSegmentToCenterlin...	18-06-2015 13:32	Documento XML

“Software SIG”- Regras de qualidade para os dados: captura / atualização / generalização, etc...

ELF WP 4 – generalisation implementation

Prototype toolbox

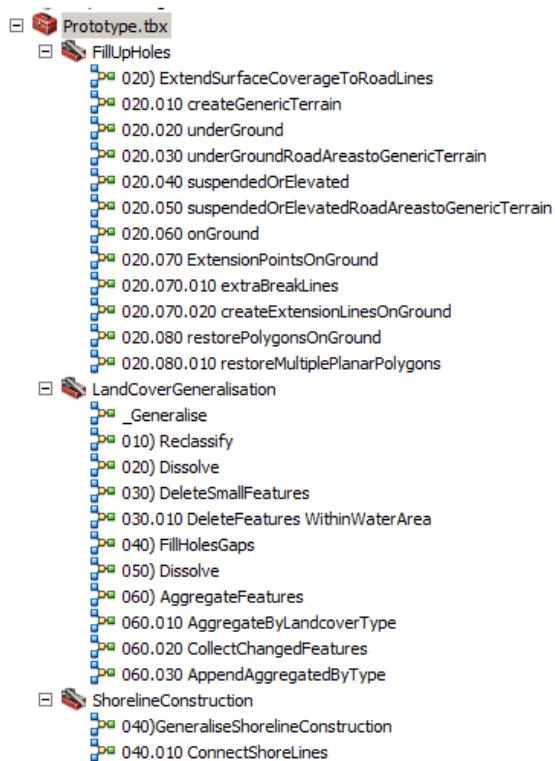


Figure 1. Prototype toolbox

	rule is unclear.
GEN.LC.4.6.	Combining two adjoining features. Actions seems equal to the previous step: <u>Combine</u> ., Snapping is not performed. If snapping is performed the feature between the LC feature has to be adjusted too.

2.4 Examples of situations

2.4.1 BuiltUpArea and LandCover

The figures below demonstrate the replacement of LandCover features with BuiltUpAreas. The first figure displays the initial situation, followed by one displaying an overlay with computed BuiltUpArea polygons and the last figure shows the integration of BuiltUpAreas in Landcover as final results.



Figure 1: Landcover and RoadLinks



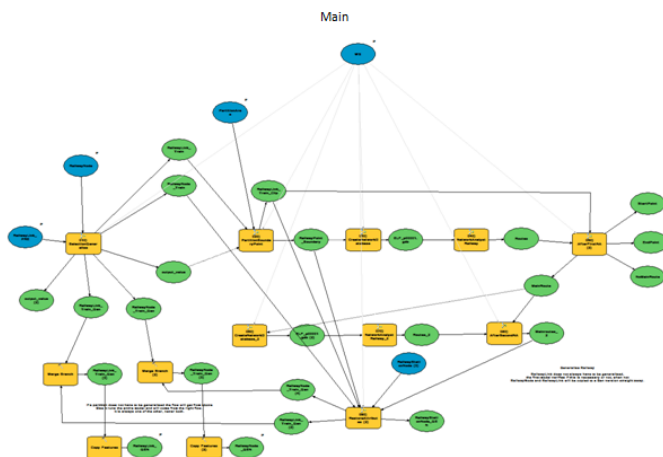
Figure 2: Landcover and builtUpArea



Figure 3: Integrate builtUpArea into landcover

“Software SIG”- Regras de qualidade para os dados: captura / atualização / generalização, etc...

ELF WP 4 – generalisation implementation of Railways



Main model Railway generalisation

Author: Kadaster-NL
Date: 30.06.2015
Version: 1.0

I Railway Generalisation

I.1 Used parameters

Minimallength: 500 meters

RailwayLinkCollapseDistance: 50 meters

1.4.1.1 Start & Endpoints

In this case begin and endpoints are the same so routes will be created from every point to the closest point.

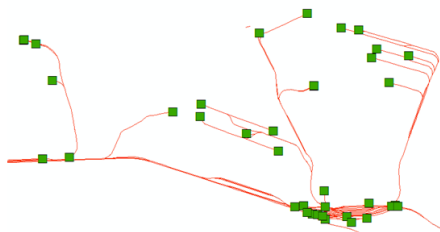


Figure 1, Start/End points where routes will be calculated from/to

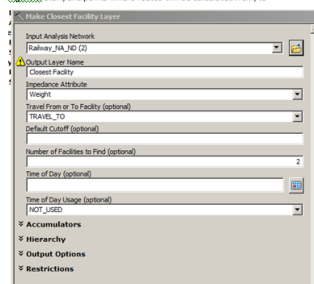


Figure 2, Tool 'Make closest facility layer' number of routes calculated is set at 2

I.3 Software requirements

The generalisation tools are built with:

Software:	ArcGIS for Desktop 10.3	
Version:	10.3.0.4322	
License type:	Advanced	
Extensions:	Spatial Analyst 10.3	Provides spatial analysis tools for use with raster and feature data.
	Production Mapping 10.3 (build 4163)	Provides user-defined integrity rules and behaviours for efficient data creation as well as data-driven cartographic production tools for creating repeatable standardized products.
	Network Analyst 10.3	Provides a complete set of tools to create, maintain, and perform analysis on network datasets.
Scripting	Python 2.7	

Start & Endpoints second Network Analyst => Mainroute + remaining RailwayLink (NotMainRoute)

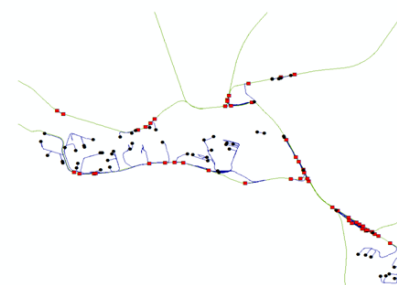


Figure 3, Black dots are startpoints, red points are endpoints

RailwayLink_Gen



Figure 4, Result after second Network analyst = RailwayLink_Gen

Dificuldades na legislação

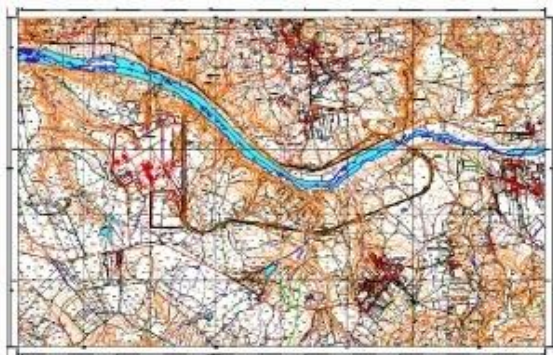
- A aquisição de dados para produção de cartografia topográfica vetorial rege-se pelo DL 141/2014. Este é a republicação do DL 193/95.
- As principais alterações introduzidas pelo DL 141/2014 foram a introdução da obrigatoriedade, para fins públicos, da cartografia topográfica ser produzida de acordo com as Especificações Técnicas (ET) da DGT, com relevância para a adoção normativa do sistema de referência ETRS89 para o continente e ITRF93 para as ilhas.
- As ET que a DGT tem na sua página são as que vigoram e que são obrigatórias para cartografia topográfica de base para fins públicos (isto desde 2014 com o DL 141/2014).
- Nestas ET a aquisição de dados é feita por fotogrametria e levantamento direto topográfico, preferencialmente.
- Pode-se recolher informação já existente (as ET só prevêm isso claramente para a rede geodésica e para a CAOP (limites administrativos) mas outras hipóteses podem acontecer, desde que os dados respeitem as exatidões de posição e completude das ET da DGT). Compete ao produtor, a menos que especificado doutra forma no Caderno de Encargos, representar a informação recorrendo ou não a dados já existentes.
- Quanto aos catálogos base das ET da DGT (escalas 1:10000 e 1:1000 a 1:5000) têm vindo, ao longo dos anos, como seria inevitável, a sofrer ajustamentos. Há ainda a situação de projetos em que o promotor deseja cartografar objetos que não estão nos catálogos. Neste caso, a DGT, se considerar ajustado o acréscimo, cria um objeto novo, com código novo, que fica guardado num catálogo à parte, não passando para o catálogo base.

Dificuldades?

Realidade dos dados de base / referência até aos dias de hoje:

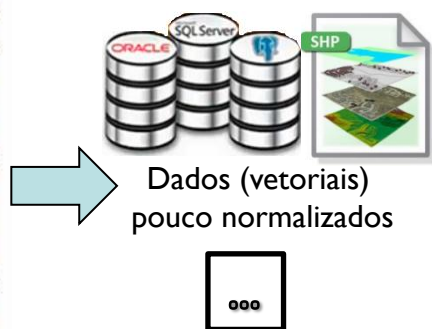
• A cartografia digital topográfica de base/referência da escala 1:10 000 [\[Link\]](#) ainda está “pensada para CAD”, objeto a objeto, para a elaboração de modelos numéricos altimétrico, topográfico e cartográfico (com regras e convenções cartográficas que derivem da uniformidade dos produtos: vetorial → raster → impressão em papel). As dificuldades prendem-se com:

- “abandono”, ao longo dos anos, do conceito de série cartográfica (existem vários “donos de obra”);
- não existe cobertura nacional nem homogeneidade no produto como um todo;
- pouca normalização em todo o *workflow* de produção/validação no contexto público/privado ao longo de mais de 20 anos em PT Continental;
- multicodificação (conceito bom para a rapidez na aquisição e garantia de consistência geográfica entre objetos mas, de difícil manuseamento em que a comunidade não entende bem o fluxo de extração a partir deste formato de dados; o controlo de qualidade por amostra não isenta os erros associados; tecnologia está em desuso há bastante tempo, etc.);
- dificuldades de interligação com o “processo SIG” nos produtos derivados/temáticos;
- nunca houve recursos/visão para materializar o produto de forma contínua e adaptado a um SGBD.

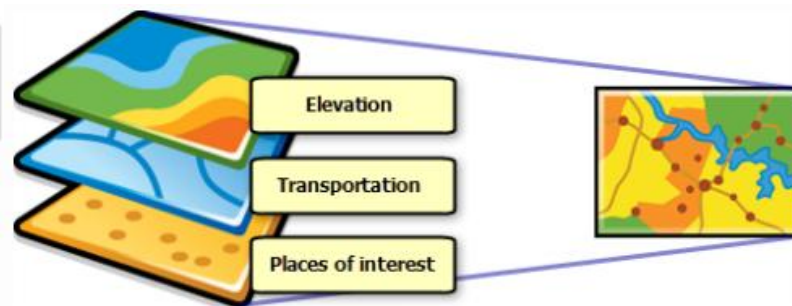


Cartografia digital tradicional

“Ambiente CAD” -> difícil associar aos dados a relação geometria/atributos, etc.
Exemplo: Onde é que começa e acaba Y?



<https://www.safe.com/integrate/>



Temas de CDG

“Ambiente SIG” -> mais difícil de assegurar qualidade aos dados na representação cartográfica e na impressão (offset), etc.

Futuro? Adaptação a novas realidades e novas tecnologias?
Estratégia nacional?
Política de dados abertos?

Entidades da AP (central, regional, local)

Produção

Informação geográfica de base / referência

Disponibilização / Cidadão

normalização

Informação geográfica temática

Produto "à medida"

- "Entendimentos" entre entidades
- Articulação entre fontes e formatos de dados

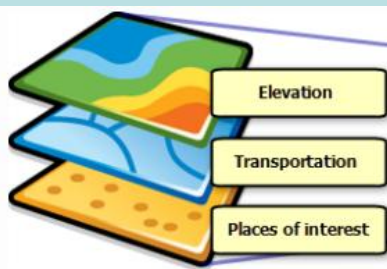


Informação geográfica



SGBD

[normalização e consistência geográfica (escalas de produção?!), alfanumérica (semântica), temporal, ...]



Representação geográfica

(regras aligeiradas para a visualização na sobreposição entre CDG de um determinado tema; produto de um ponto de vista mais isolado)



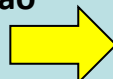
Output para WEB (mais simplista e temático)



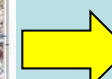
- SIG desktop/web (simbologia);
- Serviços de visualização WMS (intervalos de escala);
- Vários formatos vetoriais de descarregamento (WFS, GML, SHP, KML, ...)

Representação cartográfica

(regras bem definidas de visualização na sobreposição entre todos os dados/temas)



Cartografia digital (sub-produto) normalizada e topográfica de base/referência (foco no produto como um todo)



Output em papel (mais qualidade, mais recursos)

Como devemos encarar esta diretiva e a harmonização dos dados?

Como uma oportunidade para:

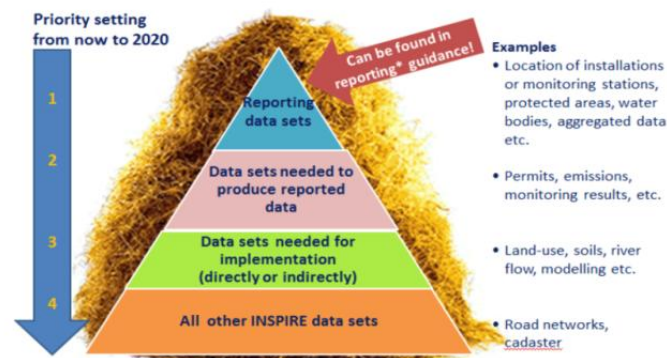
- (Re)organização dos organismos da Administração Pública (AP) com responsabilidades no domínio da informação geográfica;
- Reforçar Cooperação/Protocolos entre entidades:
 - Os dados deixarem de ser vistos isoladamente mas numa perspetiva integrada;
 - Quem utiliza os dados/serviços sabe que estão harmonizados num modelo comum;
 - Começar a perceber que os dados têm que ter muito mais consistência/organização e valor/riqueza na sua representação geográfica/cartográfica e alfanumérica/semântica, etc.

e-reporting e CDG Prioritários!?

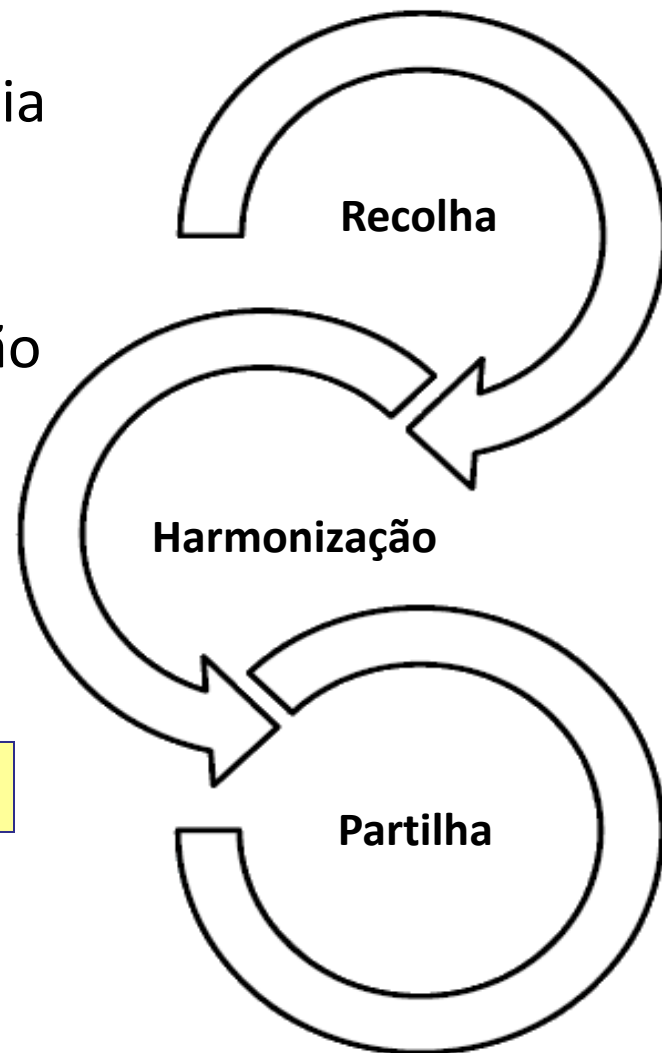
Sub-elements/spatial data sets prioritários para as diretivas ambientais...

Title with link to source	Short description	Theme	Detailed reporting requirement	Number of spatial data sets in reporting	Sub-elements/spatial
Directive 2002/49/EC relating to the assessment and management of environmental noise.	List of major roads, railways, airports and agglomerations (DF1_5)	Noise	location of agglomerations, major airports and major roads/railways	3	5) Roads, railways and air network 6) Agglomerations 7) Population
Directive 2002/49/EC relating to the assessment and management of environmental noise.	Noise reduction measures already in place (DF6_9)	Noise	location of agglomerations, major airports and major roads/railways concerned by noise reduction measures in place	3	5) Roads, railways and air network 6) Agglomerations 7) Population

... várias diretivas ambientais que “pretendem”
os dados de diferentes formas...
... foco na disponibilização...
... foco no cidadão?



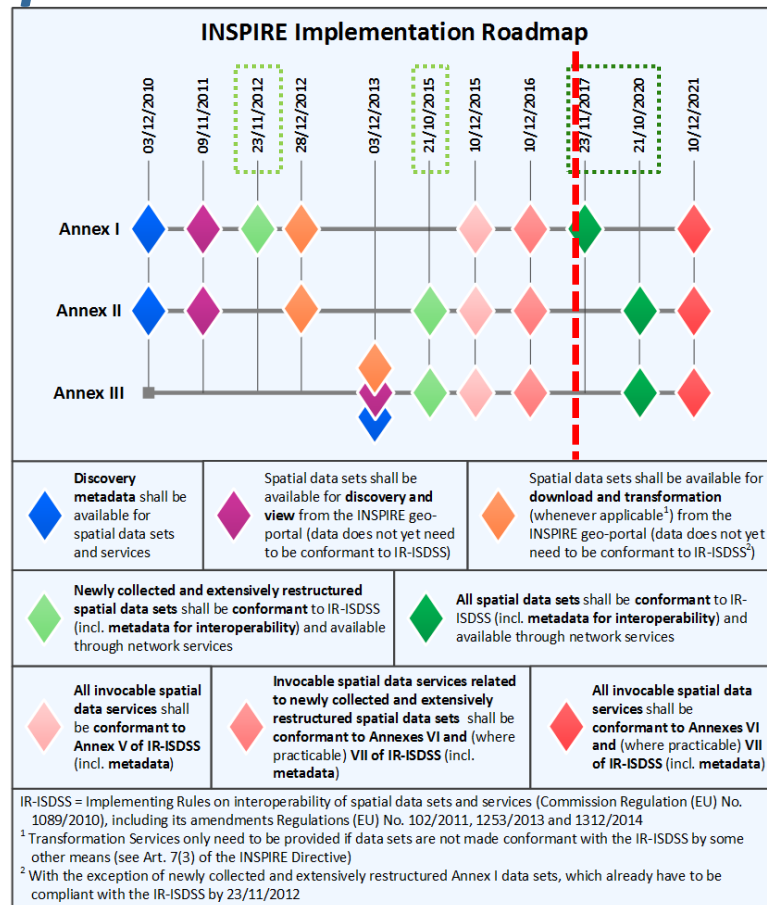
- Diretiva INSPIRE
 - Criação da Infraestrutura Europeia de Informação Geográfica
 - Disponibilizar aos utilizadores serviços integrados de informação geográfica
- Principais exigências
 - Metadados
 - **Dados e Serviços interoperáveis**
 - Serviços de Rede
 - Acesso e partilha de dados
 - Monitorização e *reporting*



Processo de Harmonização de CDG Roadmap

- Fases da implementação da Diretiva INSPIRE

- 1. Documentação** - criação e disponibilização de metadados
- 2. Acessibilidade** - por intermédio dos Serviços de Dados Geográficos
- 3. Harmonização** - por forma a garantir a interoperabilidade dos CDG

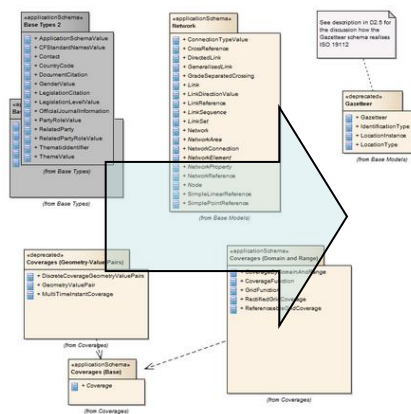


- Não se pode fugir das responsabilidades e da lei...
 - Mais cedo ou mais tarde terá de se fazer...
 - Ninguém vai fazer por nós...
 - Transversal a toda a Administração Pública!!!

Dados novos (ou extensamente reestruturados)
harmonizados e disponíveis através de serviços de rede:
 - Anexo I (23-11-2012)
 - Anexo II e III (21-10-2015)

Dados existentes harmonizados e disponíveis através de serviços de rede:
 - Anexo I (23-11-2017)
 - Anexo II e III (21-10-2020)

- Desafio
 - Organização dos dados



- Harmonização de dados na Directiva INSPIRE

“o processo que permite desenvolver especificações para conjuntos de dados, de modo a que seja possível aceder a estes dados através de serviços, numa representação que permite

combinar esses dados com outros **dados harmonizados de forma coerente**”

- “Sítio” INSPIRE
 - Modelos UML (HTML, projecto EA)
 - XML schemas
 - Feature catalog
 - Matching tables
 - ...

INSPIRE
Infrastructure for spatial information in Europe

European Commission > INSPIRE > Implement > Data Models

Home Learn **Implement** Participate Use Toolkit

Implement

- Guide for implementers
- Roadmap
- Data Specifications**
- Monitoring & Reporting
- Metadata
- Network Services
- Data and Service Sharing
- Spatial Data Services
- Maintenance and Implementation Framework

Data Specifications

- Overview
- Technical Guidelines
- Legislation
- Roadmap
- Themes
- Data Models
- XML Schemas
- Library
- News
- Events
- Training
- MIG Work Programme
- Experts
- Tools

Data Models

INSPIRE data models

The INSPIRE Implementing Rules on interoperability of spatial data sets and services and the data specification guidance documents are based on the UML data models developed by the INSPIRE Thematic Working Groups. These data models are managed in a common UML repository, which also stores older revisions of the models.

This page makes different revisions of the INSPIRE UML models available in different formats and views (see below). Each of these revisions corresponds to a specific set of (draft or approved) Data Specification Technical Guidance (TG) documents and/or Implementing Rules.

Revision	Corresponding TG and IRS	Status	Feature HTML Mapping	EA	SVN	GML & code lists
4618	This version corresponds to the content of the Implementing Rules (EU) No 1089/2010, No 102/2011, No 1253/2013 and the latest publicly available version of the data specifications of Annex I, II+III.		catalogue	view	Tables	project / XMI

Quick search

Implementing Rules
(Disposições de execução = Legalidade)

Data Specifications
(Guias técnicos: <http://inspire.ec.europa.eu/data-specifications/2892>)

INSPIRE LEGISLATION

legally binding / not legally binding

abstract specification / implementation specification

What Member States must implement

How Member States might implement it

INSPIRE Implementing Directive 2007/2/EC

INSPIRE Implementing Rules for Metadata, Network Services Interoperability Spatial data sets and services

INSPIRE Technical Guidance for the Implementation of Discovery Services

INSPIRE Technical Guidance Data Specification on Addresses

Implementation Requirements & Recommendations

- Formas de harmonização dos CDG
 - “On-the-fly”, usando serviços de transformação.
 - “Offline”, usando ferramentas de harmonização.

Início processo
de harmonização



Identificação
do CDG



Modelo
de dados?

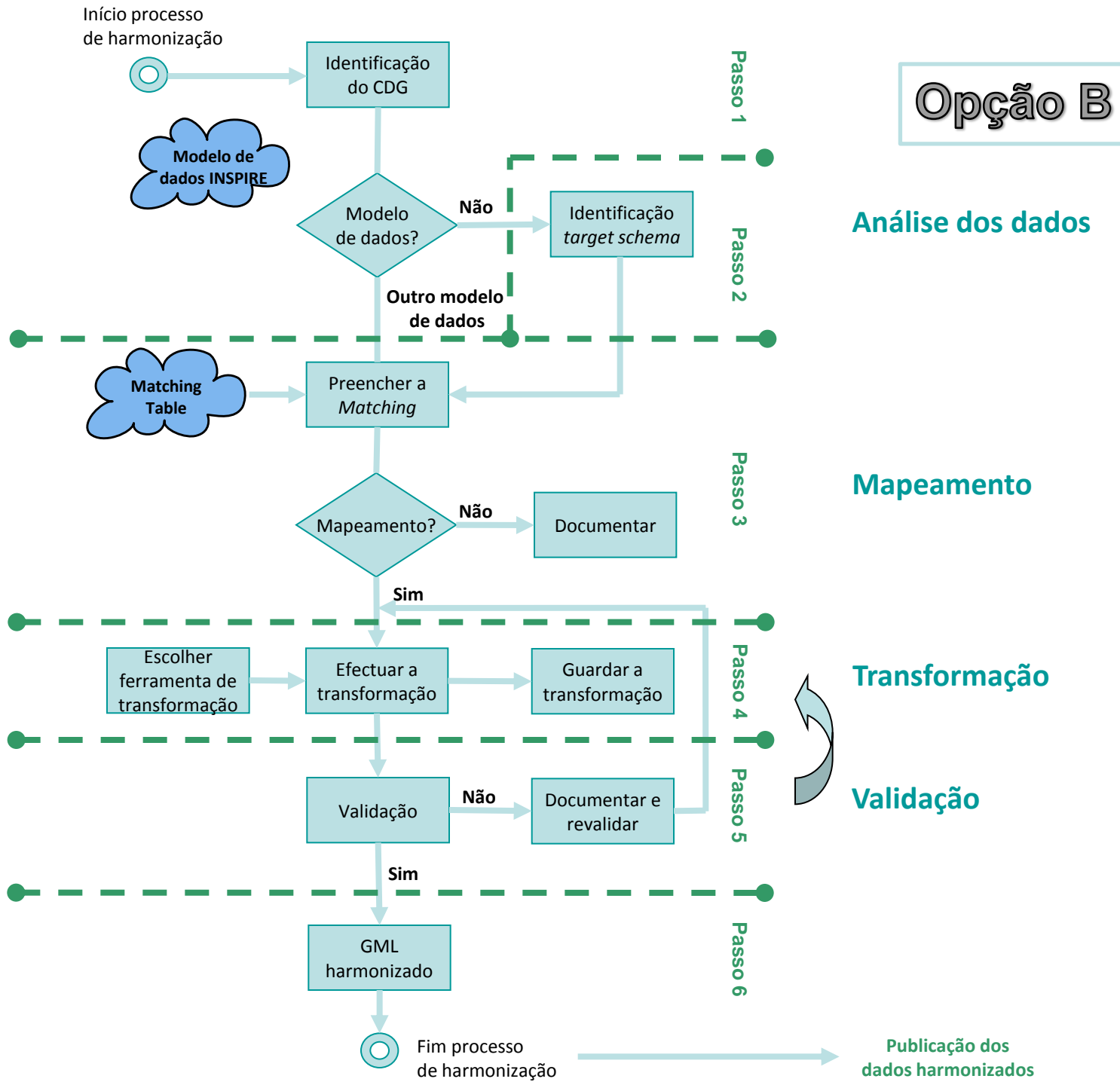
GML
harmonizado

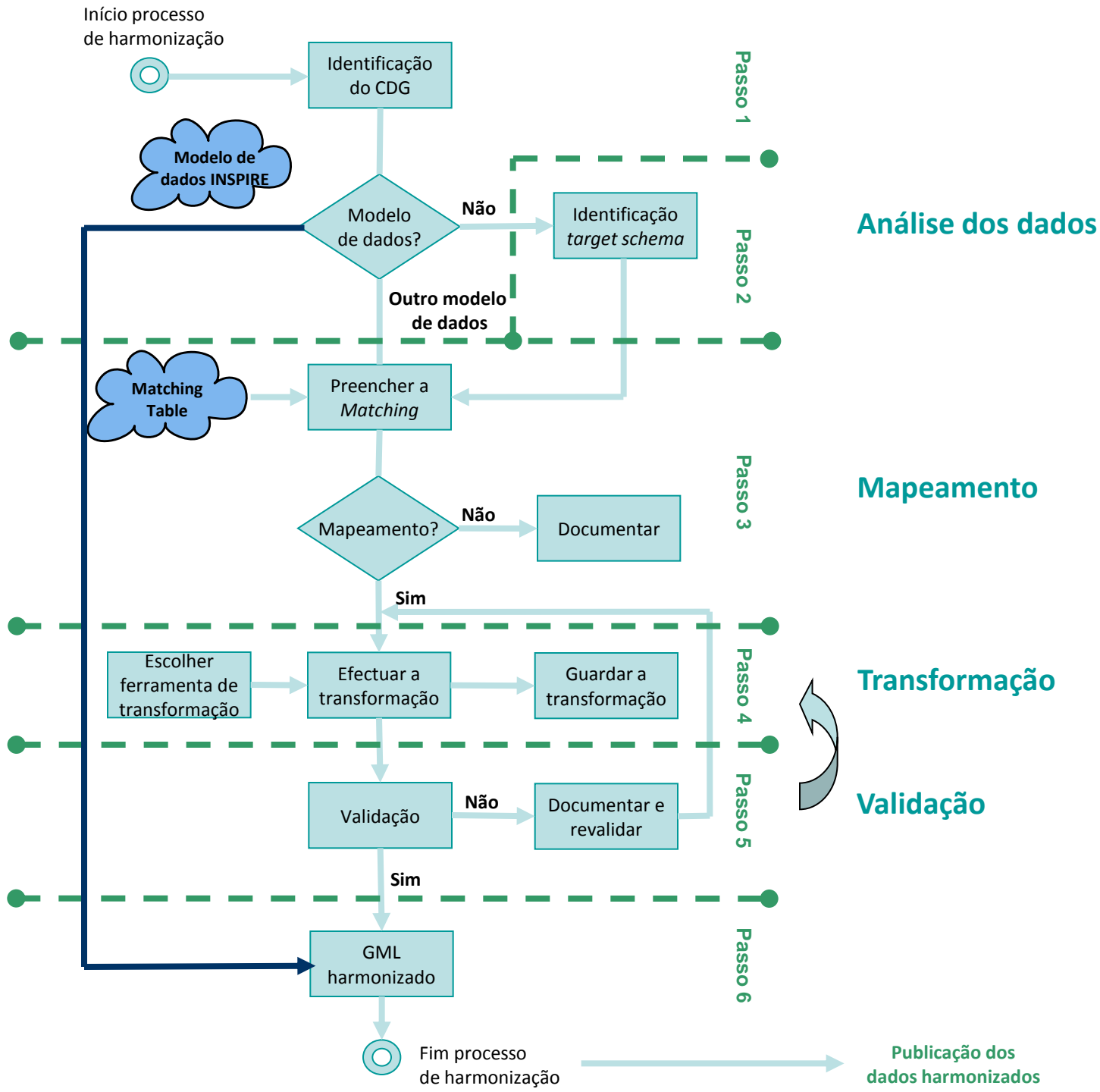


Fim processo
de harmonização

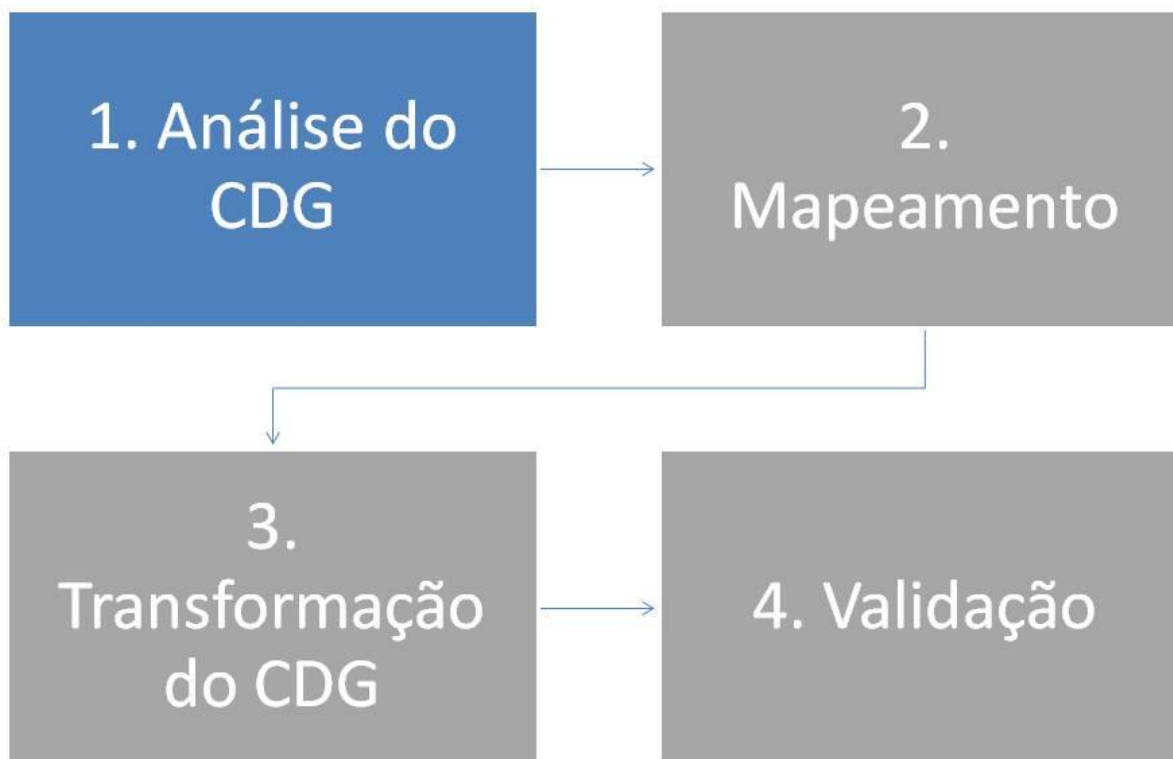
Opção A

Publicação dos
dados harmonizados





- Processo 1

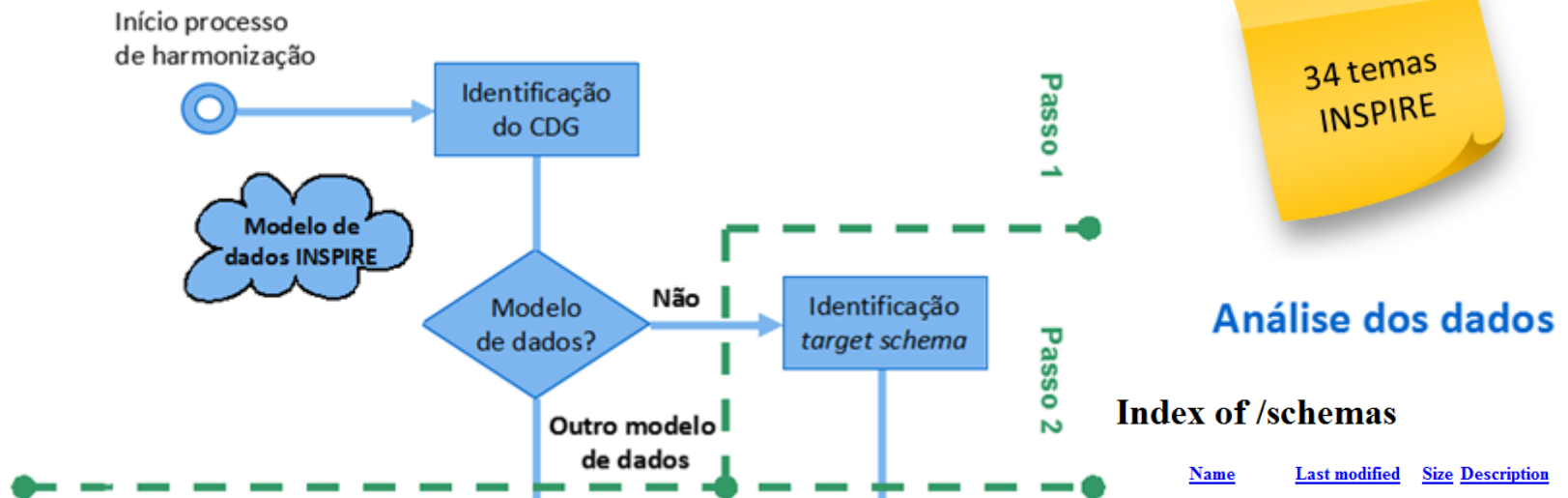


Processo de Harmonização

1. Análise do CDG

- **Passo 1:** Identificar e caracterizar a informação original (*source*)
- **Passo 2:** Identificar o modelo de dados INSPIRE (*target*)

<https://inspire.ec.europa.eu/Themes/Data-Specifications/2892>



Name	Last modified	Size	Description
Parent Directory		-	
ac-mf/	2015-04-29 10:03	-	
act-core/	2015-04-29 10:03	-	
ad/	2015-04-29 10:03	-	
af/	2015-04-29 10:03	-	
am/	2015-04-29 10:03	-	
...			

Processo de Harmonização

1. Análise do CDG

Como devo fazer?

1) Análise dos dados (esquema fonte):

Identificar e caracterizar a informação geográfica existente:

- Formato dos dados / Geometria
- Representação espacial
- Atributos
- Sistema de Coordenadas
- Metadados, etc.

- Existem problemas a montante?
 - Como se obtiveram estes dados?
 - Consigo com estes dados dar resposta ao entendimento que a minha instituição tem sobre a sua missão em PT e também na CE?

Exemplo: Carta de Ocupação do Solo 2010

Modelo de representação	Vetorial
Formato dos dados / Geometria	Shapefile / 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 226 classes

Designação dos atributos	Terminologia/ Formato
Área (ha)	AREA/ Float (19 algarismos/10 casas decimais)
Identificador único	FID/ OID (4 caracteres)
Classe de Ocupação do Solo	USO/ String (10 caracteres)

REVER:

Workshop sobre Harmonização de Dados Geográficos de acordo com as especificações INSPIRE - abril de 2016, DGT

<http://snig.igeo.pt/Inspire/documents/workshop-COS/TZ-INSPIRE-COS.pdf>

Vídeos:

Harmonização da COS de acordo com as especificações INSPIRE - Teresa Zuna – Parte 1

<https://www.youtube.com/watch?v=61m15WdDHEQ>

Harmonização da COS de acordo com as especificações INSPIRE - Teresa Zuna – Parte 2

<https://www.youtube.com/watch?v=azcRtbgYZzA>

Processo de Harmonização

1. Análise do CDG

Como devo fazer?

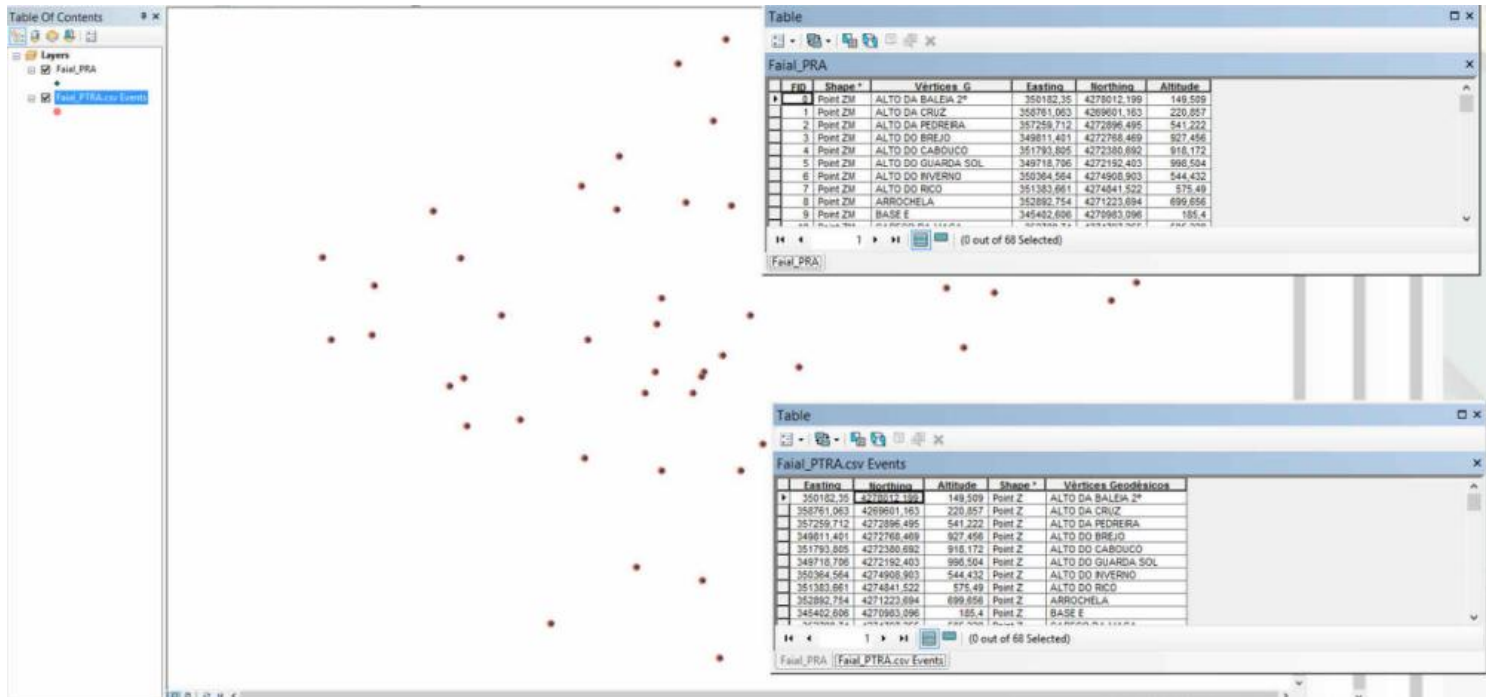
1) Análise dos dados (esquema fonte):

- Ainda tenho os dados no formato CAD (dgn, dwg)? **Sim.**
- Existem “construídos” processos automáticos de extração/conversão para “formato SIG”? **Não.**

Então, por exemplo, converter “manualmente” CAD→SIG (ArcGIS, QGIS, outro)

Recolha de informação e Conversão de dados para SIG (1ª Sessão Técnica RAA, 13 Abril 2016, LREC):

http://www.ideia.azores.gov.pt/docs/Documentos/1Sessao_Apresentacao_MarleneAntunes.pdf



Processo de Harmonização

1. Análise do CDG

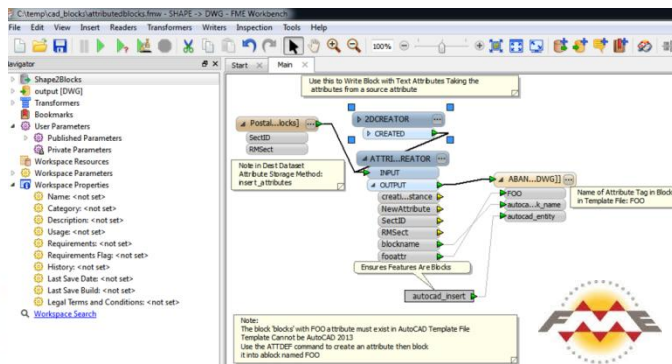
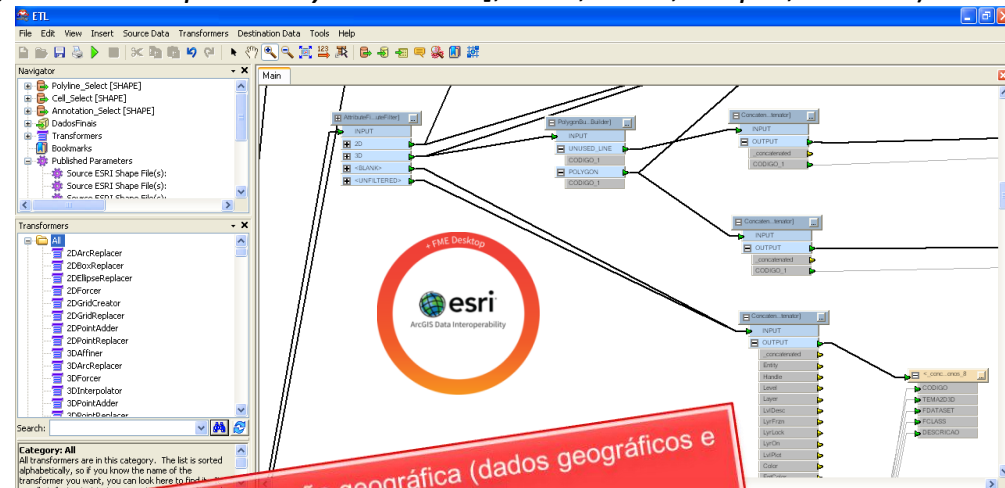
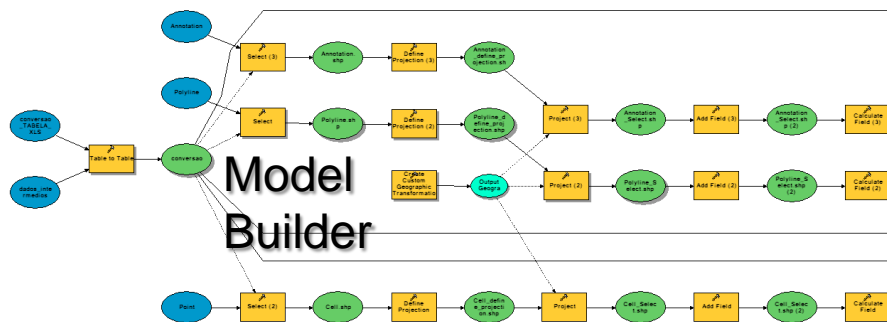
□ **II.2 Land Cover – Carregamento de dados da SRAA no Modelo de Dados INSPIRE RAA [partilha de ecrã] (3ª Sessão Técnica RAA, 18 e 24 de maio 2016):**

Como devo fazer?

1) Análise dos dados (esquema fonte):

- Ainda tenho os dados no formato CAD (dxf, dgn, dwg)? **Sim / Não.**
- Existem “construídos” processos automáticos de extração/conversão para “formato SIG”? **Não.**

Então, por exemplo, arranjar processos automáticos (*Workflows*) “CAD→SIG” ou “SIG1Entrada→SIG2Saída”.
Ferramentas: (ArcGIS Desktop [*model builder, data interoperability extension*], FME, QGIS, *scripts*, outro...)



- É necessário garantir que a informação geográfica (dados geográficos e alfanuméricos) foi devidamente convertida.
- Não se consegue automatizar tudo da maneira que pretenderíamos.
 - Equipa SIG a jusante para:
 - verificar automatismos (de "caso para caso" pode variar)
 - garantir processos de controlo de qualidade aos dados
 - completar "manualmente" o que não for possível fazer automaticamente (Nota: Por vezes dispense-se nesta fase muito tempo)

Processo de Harmonização

1. Análise do CDG

Como devo fazer?

1) Análise dos dados (esquema alvo):

Análise das Disposições de Execução (DE):

<http://inspire-regadmin.jrc.ec.europa.eu/dataspecification/>



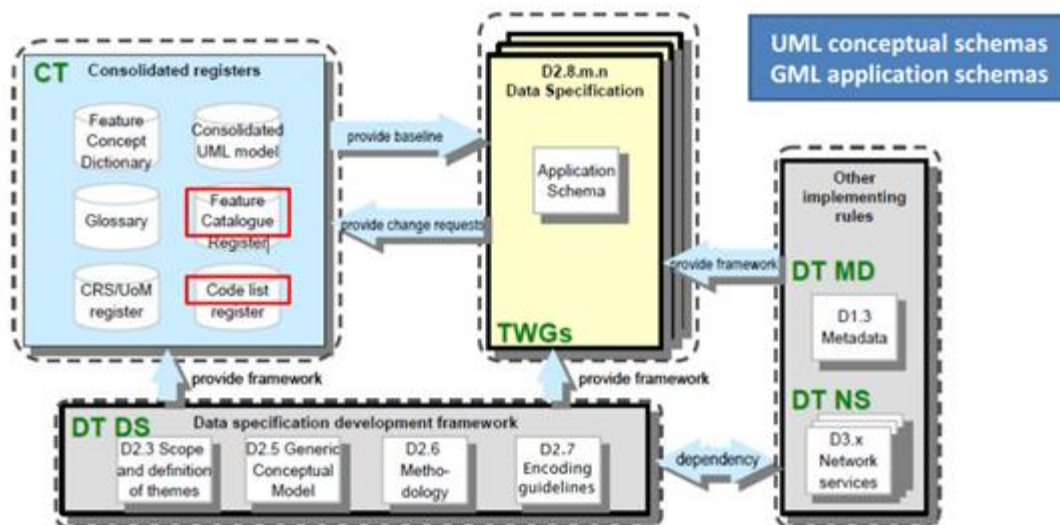
INSPIRE
Infrastructure for Spatial Information in Europe

D2.8.1.7 INSPIRE Data Specification on Transport Networks – Guidelines

Title	D2.8.1.7 INSPIRE Data Specification on Transport Networks – Guidelines
Creator	INSPIRE Thematic Working Group Transport Networks
Date	2010-04-26
Subject	INSPIRE Data Specification for the spatial data theme Transport Networks
Publisher	INSPIRE Thematic Working Group Transport Networks
Type	Text
Description	This document describes the INSPIRE Data Specification for the spatial data theme Transport Networks
Contributor	Members of the INSPIRE Thematic Working Group Transport Networks
Format	Portable Document Format (PDF)
Source	
Rights	Public
Identifier	INSPIRE_DataSpecification_TN_v3.1.pdf
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

Diretiva INSPIRE

XML schema files (.xsd) são fornecidos pelo INSPIRE



Esta figura ilustra as relações e as dependências existentes entre os documentos INSPIRE. As caixas representam as disposições de execução ou outros documentos relevantes; os cilindros representam os registros para certos elementos constituintes da IDE e há a obrigatoriedade de possuírem uma identificação única através de um http URI. As setas demonstram as relações de dependência entre os diversos documentos

INSPIRE Thematic Clusters

<https://themes.jrc.ec.europa.eu/>

Processo de Harmonização

1. Análise do CDG

Como devo fazer?

1) Análise dos dados (esquema alvo):

- Especificações de Dados INSPIRE



INSPIRE
Infrastructure for Spatial Information in Europe

D2.8.1.4 INSPIRE Data Specification on Administrative units – Guidelines

Title	D2.8.1.4 INSPIRE Data Specification on <i>Administrative units</i> – Guidelines
Creator	INSPIRE Thematic Working Group <i>Administrative units</i>
Date	2010-04-28
Subject	INSPIRE Data Specification for the spatial data theme <i>Administrative units</i>
Publisher	INSPIRE Thematic Working Group <i>Administrative units</i>
Type	Text
Description	This document describes the INSPIRE Data Specification for the theme <i>Administrative units</i>
Contributor	Members of the INSPIRE Thematic Working Group <i>Administrative units</i>
Format	Portable Document Format (pdf)
Source	
Rights	public
Identifier	INSPIRE_DataSpecification_AU_v3.0.1.pdf
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

Theme Overview

[Executive Summary](#)

[Detailed description](#)

[Data content and structure](#)

[Data quality](#)

[Metadata](#)

[Delivery](#)

[Data capture](#)

[Portrayal](#)

[Abstract Test Suite](#)

[Use cases](#)

[Code list values](#)

[Additional information](#)

Processo de Harmonização

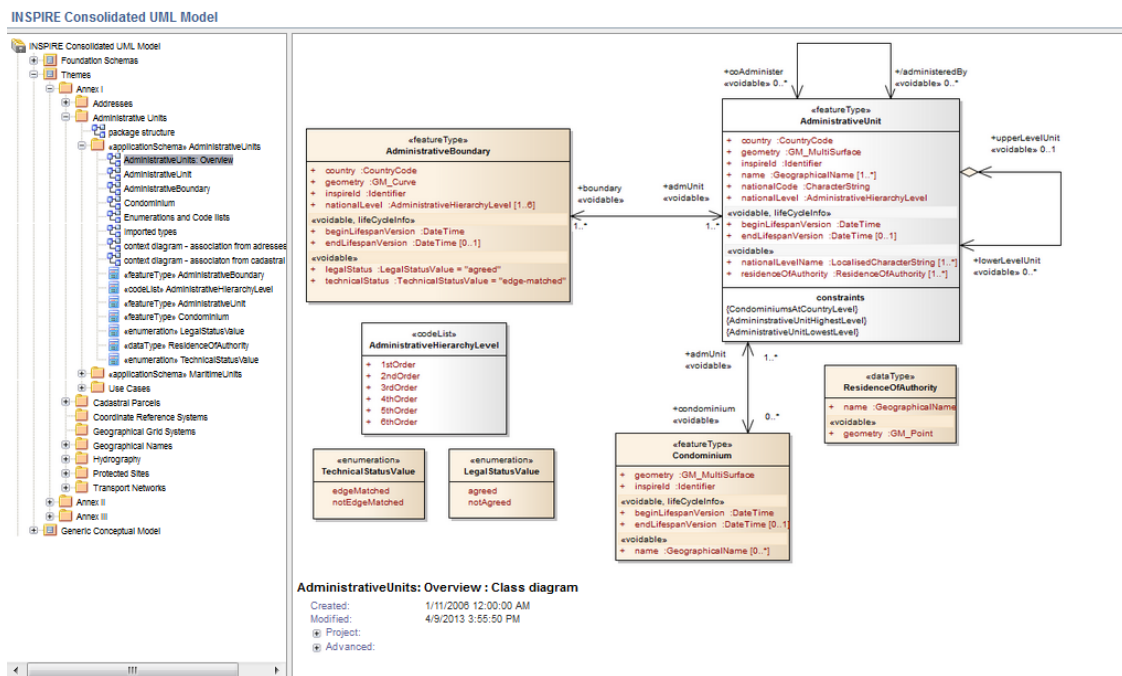
1. Análise do CDG

Como devo fazer?

1) Análise dos dados (esquema alvo):

- Diagramas UML

<http://inspire.ec.europa.eu/data-model/approved/r4618-ir/html/>



- *Feature Catalog* – tipo de objecto e de dados espaciais definidos nas especificações de dados.

Processo de Harmonização

1. Análise do CDG

Análise das Disposições de Execução (DE) dos temas do GT temático e de outra documentação relevante:

- Disposições de execução (pdf, ferramenta interativa de pdfs, etc)
 - Data models: <http://inspire.ec.europa.eu/index.cfm/pageid/2/list/datamodels>
 - Matching tables (XML ⇒ abrir com o Excel) : <http://inspire.ec.europa.eu/data-model/approved/r4618-ir/mapping/>
 - GML Application Schemas: (XSD ⇒ abrir por exemplo com o HALE): <http://inspire.ec.europa.eu/schemas/>

Links úteis importantes:

Implementing Rules EN vs PT	http://eur-lex.europa.eu/legal-content/EN-PT/TXT/?uri=CELEX:02010R1089-20131230&from=EN
Implementing Rules (errata)	http://inspire.ec.europa.eu/documents/commission-regulation-eu-no-10892010-23-november-2010-implementing-directive-20072ec-0
Catálogo (objetos, codelists, etc)	http://inspire-regadmin.jrc.ec.europa.eu/dataspecification/CatalogueINSPIREObjects.action
Layers (WMS)	http://inspire.ec.europa.eu/layer
Glossário	http://inspire.ec.europa.eu/glossary

GT	Tema do Anexo DE pdf	DE interativa	Matching tables (XML ⇒ abrir com o Excel)	GML-AS (XSD)
GTI-TE-9 - Topografia e Cadastro	I.3 Toponímia [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
	I.4 Unidades administrativas [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
	I.5 Endereços [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
	I.6 Prédios (Nota: em PT é diferente de Parcelas cadastrais) [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
	I.7 Redes de transporte [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
			[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
			[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
	I.8 Hidrografia [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
	III.2 Edifícios [Link]	[Link]	[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]
			[Link XML] = [Link XML PCT]	[Link XSD v4.0] = [Link XSD PCT]

Processo de Harmonização

1. Análise do CDG

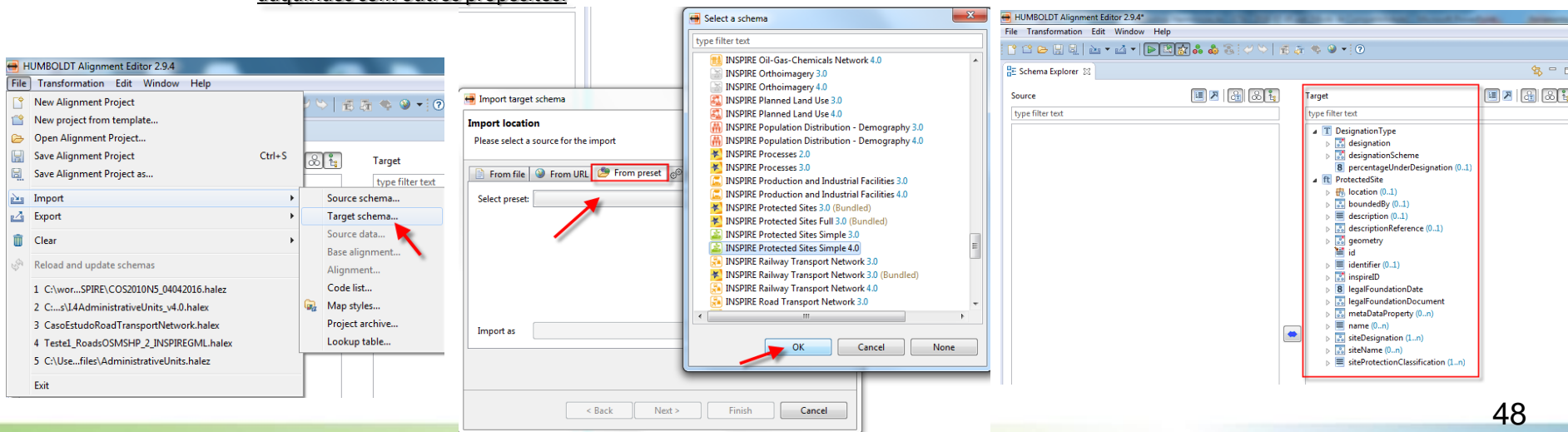
Como devo fazer?

1) Análise dos dados (esquema alvo):

- Identificação do tema da diretiva INSPIRE
- Interpretação dos documentos INSPIRE
 - *General Conceptual Model* http://inspire.ec.europa.eu/documents/Data_Specifications/D2.5_v3.4rc3.pdf
 - *Data Specifications* <http://inspire.ec.europa.eu/index.cfm/pageid/2>
 - Feature catalog (Catálogo de objectos)
 - Codelists, etc...
 - Diagrama UML
 - Matching table (Tabela de correspondências)
 - Application Schema XSD, XML, etc...

Com o software Hale é possível:

- Analisar o esquema alvo (para saber o que se pretende e como vamos lá chegar a partir dos dados fonte).
- Repensar estratégias, e novos *workflows* (a montante), para analisar de que forma os dados fonte conseguem “encaixar” no “esquema alvo”.
 - Dificuldades: Pode não ser possível “encaixar” os dados existentes neste modelo de dados (de nenhuma forma) já que os dados foram adquiridos com outros propósitos.



Processo de Harmonização

1. Análise do CDG

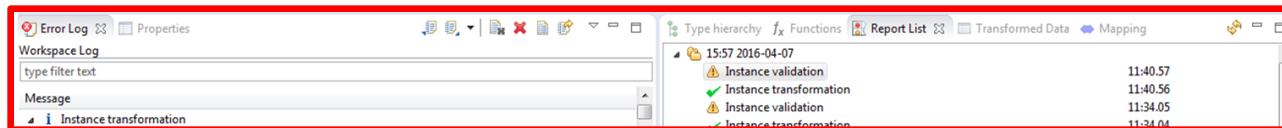
Como devo fazer?

1) Análise dos dados (esquema alvo):

Objetivo principal (nesta fase):

Identificar e tentar perceber se com os dados que temos é possível “dar resposta” às propriedades obrigatórias (*mandatory*):

- facilmente identificadas no *error log* quando não estão preenchidas



Groups

Normal group containing a set of properties.

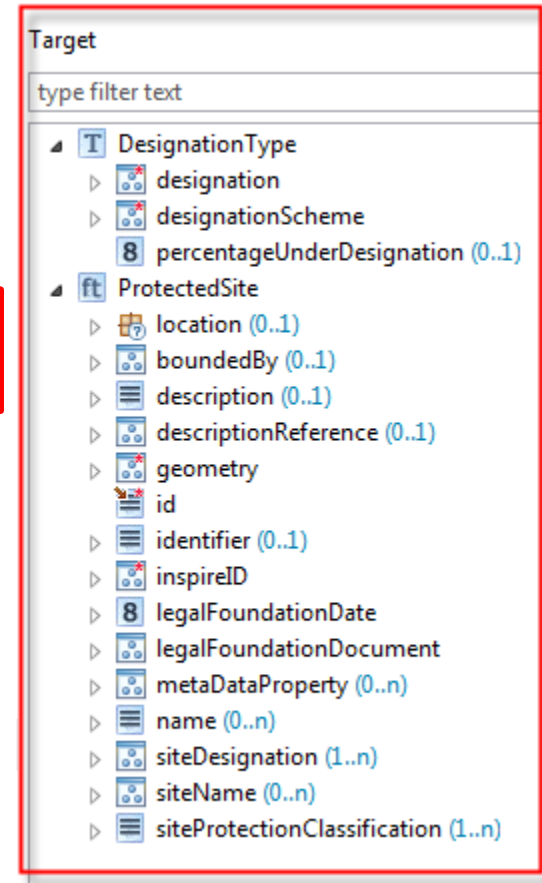
Choice group, where only one of the specified properties is allowed as a child.

A red asterisk marks properties that are mandatory, i.e. they occur exactly once and must have a value (and the value may not be null).

Please note that if the parent of such a property is a choice () the choice takes precedence, i.e. only one of its children may be present in an object, but the child that is present may still not hold a null value if marked with a red asterisk.

A small brown error in the top left corner marks a property from a XML schema as being defined as a XML attribute.

A property that is deemed to hold the main geometry of a type is marked with a small green triangle. Per type, you can set one property as the default geometry property. This property is then used when retrieving geometries for display in the map.



Properties

- String property
- Numeric property
- Geometry property
- Other (complex) property

Processo de Harmonização

1. Análise do CDG

Como devo fazer?

1) Análise dos dados (esquema alvo): propriedade obrigatória → inspireId

Grupo Transversal (GTI-TR):

O inspireId é constituído pelos seguintes campos:

- localID – refere-se ao objectID de cada elemento do CDG;
- namespace – espaço de nomes que define o âmbito do conjunto de códigos. Este deverá ser registado no INSPIRE External Object Identifier Namespaces Register se o inspireId não tiver a forma de um URI.
- versionID – versão do CDG;

Qualquer actualização a que este CDG seja sujeito, deverá conter o mesmo localID e o mesmo namespace, neste caso só o número da versão será alterado.

<http://id.igeo.pt/so/{SiglaTemaInspire}/{ObjectoGeograficoInspire}/{localId}/{versão}>

InspireId = Espaço de Nomes + localId + versão

Espaço de Nomes = <http://id.igeo.pt/so/{SiglaTemaInspire}/{ObjectoGeograficoInspire}>

SiglaTemaInspire = Sigla do Tema INSPIRE, 2 caracteres. Ver Anexo.

ObjectoGeograficoInspire = Objecto geográfico INSPIRE. Por extenso. Ver anexo.

localId = código ou designação do recurso original + “_” + código do objecto geográfico original

versão = versão do recurso original

Exemplos:

http://id.igeo.pt/so/GN/NamedPlace/topon200k_1/2010

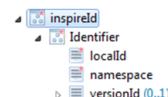
http://id.igeo.pt/so/AU/AdministrativeUnit/caop_010101/2010

A infraestrutura INSPIRE exige que qualquer informação contida nos temas, seja descrita com clareza, gerida ao longo do tempo e que esteja disponível online. Exige ainda que esta seja referenciada através de identificadores únicos, de forma a permitir uma correcta e permanente identificação de cada item constituinte de cada tema.

Com o software Hale:

Função Assign

- Exemplo da função Assign com campo inspireId
- inspireId é um campo complexo constituído por 3 campos localId, namespace e versionId, em que os dois primeiros são obrigatórios.



GEOGRAPHICAL NAMES GN

WMSNomeLayer	WMSTituloLayer	SiglaTemaInspire	ObjectoGeograficoInspire	Geometria	Tipo
GN.GeographicalNames	Geographical Names	GN	NamedPlace	GM_Curve	«featureType»

GM_Object

ADMINISTRATIVE UNITS AU

WMSNomeLayer	WMSTituloLayer	SiglaTemaInspire	ObjectoGeograficoInspire	Geometria	Tipo
AU.AdministrativeBoundary	Administrative boundary	AU	AdministrativeBoundary	GM_Curve	«featureType»
AU.AdministrativeUnit	Administrative unit	AU	AdministrativeUnit	GM_MultiSurface	«featureType»
AU.Condominium	Condominium	AU	Condominium	GM_MultiSurface	«featureType»
AU.Baseline	Baseline	AU	Baseline	GM_Curve	«featureType»
AU.MaritimeBoundary	Maritime boundary	AU	MaritimeBoundary	GM_Curve	«featureType»
AU.<CodeListValue>	<human readable name>	AU	MaritimeZone	GM_MultiSurface	«featureType»

Processo de Harmonização

1. Análise do CDG

Como devo fazer?

1) Análise dos dados (esquema alvo): *codelists* = listas de códigos

GTI-TR:

Identificadores para as Listas de Códigos e Códigos

Autor	GT Transversal
Data de criação	2016-06-16
Data de alteração	
Assunto	
Publicação	SNIG 2020
Descrição	Padrões a utilizar em Portugal para identificadores permanentes de listas de códigos e códigos não definidos no INSPIRE Registry, sob a forma de um URI e mantidos a nível nacional através da aplicação FOSS Re3gistry.

Contribuição/ Revisão

Estado Versão 0.9

Padrões para os identificadores das listas de códigos e códigos não definidos no INSPIRE Registry:

1. <http://registro.igeo.pt/codelist/{NomeListaCodigos}Value>
2. <http://registro.igeo.pt/codelist/{NomeListaCodigos}Value/{ValorCodigo}>

NomeListaCodigos: nome definido pela entidade responsável pela manutenção da lista (*CamelCase*, letra maiúscula inicial, PT ou EN)

ValorCodigo: nome definido pela entidade responsável pela manutenção da lista respectiva (*CamelCase*, letra minúscula inicial, PT ou EN)

Exemplos:

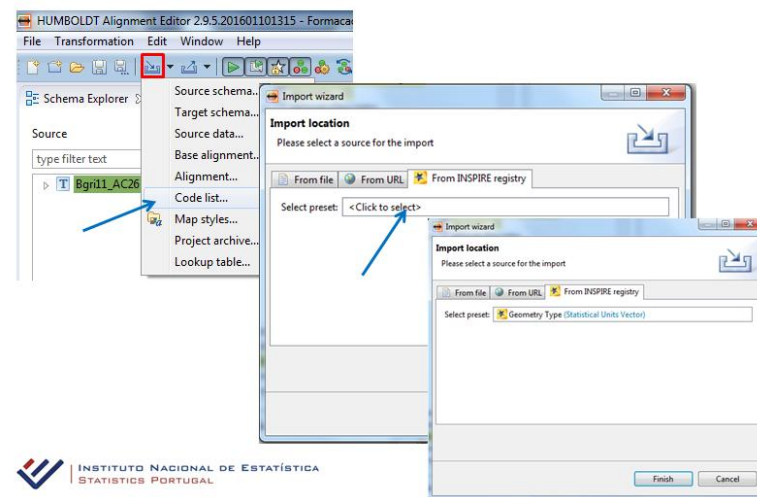
<http://registro.igeo.pt/codelist/HumanHealthAndSecurityAggregationUnitValue>

<http://registro.igeo.pt/codelist/HumanHealthAndSecurityAggregationUnitValue/NUTSI>

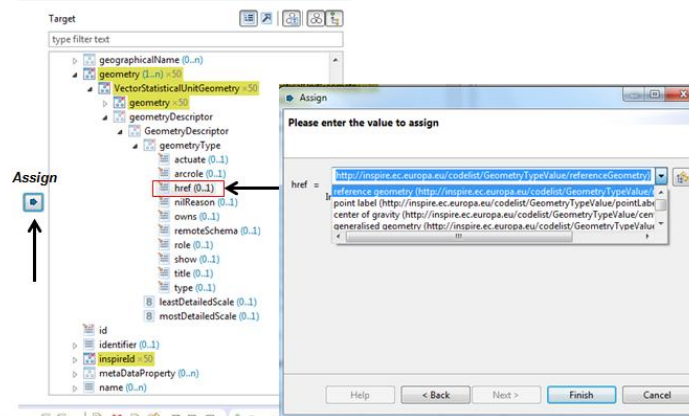
<http://registro.igeo.pt/codelist/CartaOcupacaoSoloValue/1.1.1.01.1>

Com o software Hale:

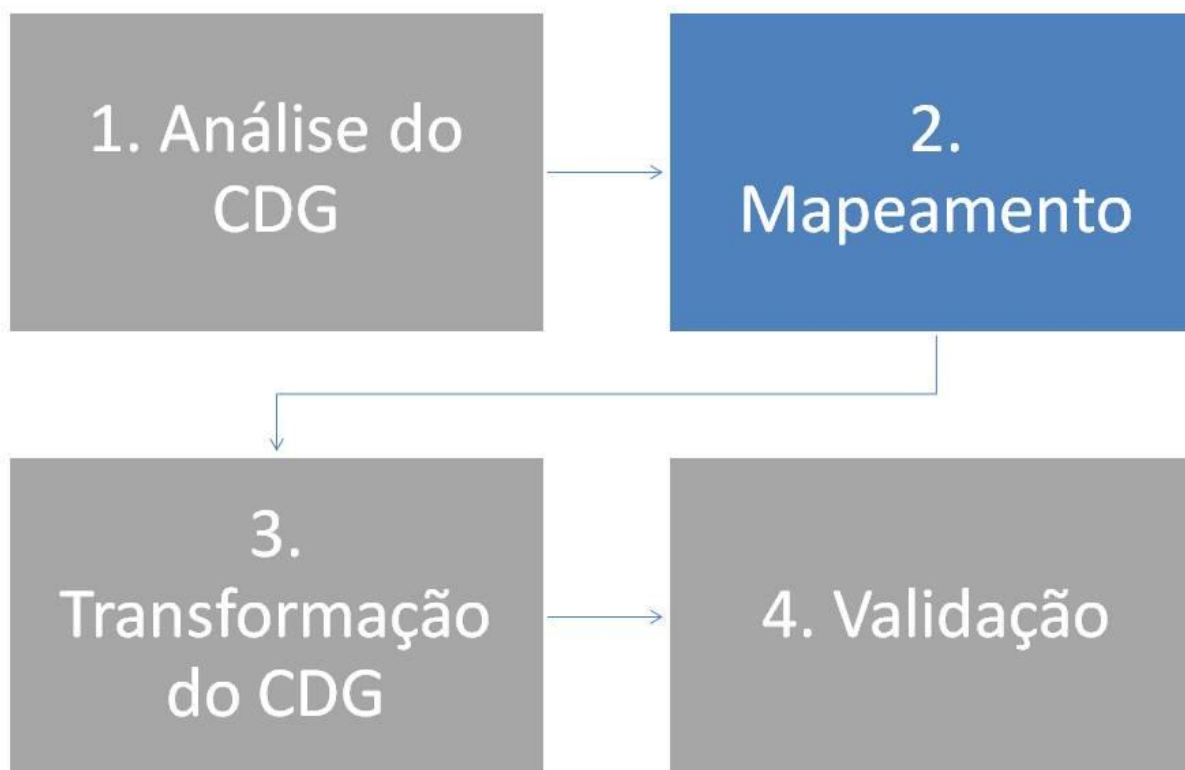
Como importar *Codelists*



Como utilizar *Codelists*



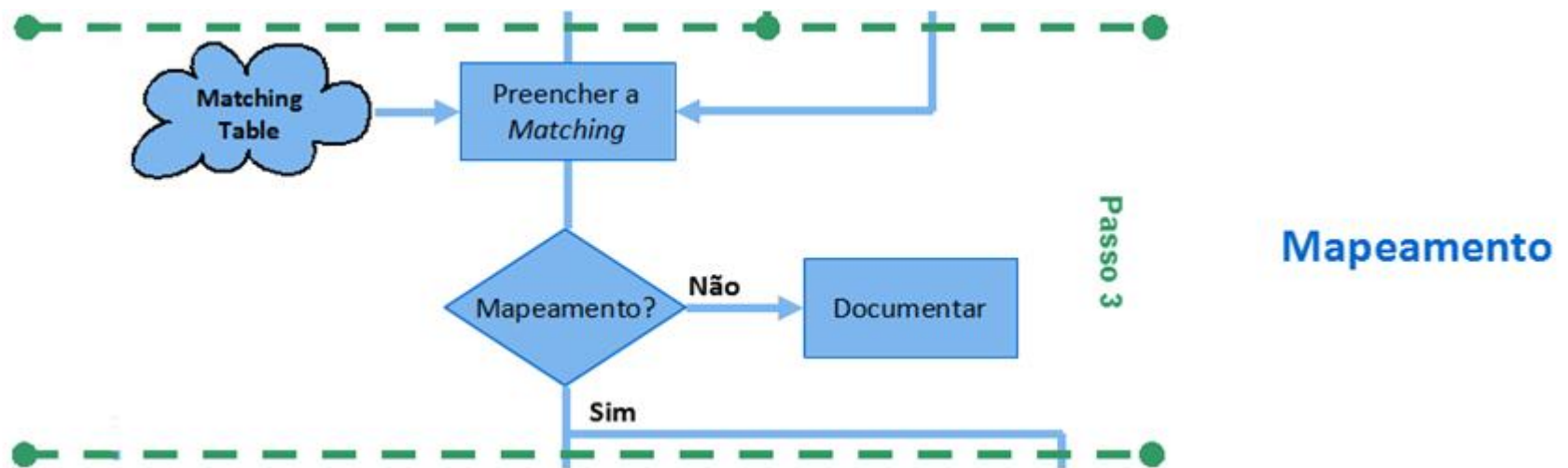
- Processo 2



Processo de Harmonização

2. Mapeamento

- **Passo 3:** Comparar e documentar a informação inicial e final



Processo de Harmonização

2. Mapeamento

Como devo fazer?

2) Mapeamento


- Deve-se analisar a documentação existente para o efeito:

- [Find your scope](http://inspire-regadmin.jrc.ec.europa.eu/dataspecification/FindYourScope.action) <http://inspire-regadmin.jrc.ec.europa.eu/dataspecification/FindYourScope.action>

O principal resultado desta aplicação é a lista dos objetos INSPIRE, incluindo suas propriedades - atributos, valores de listas de código, etc, que são relevantes para um conjunto de dados. A lista final inclui também todos os objetos associados e suas propriedades.

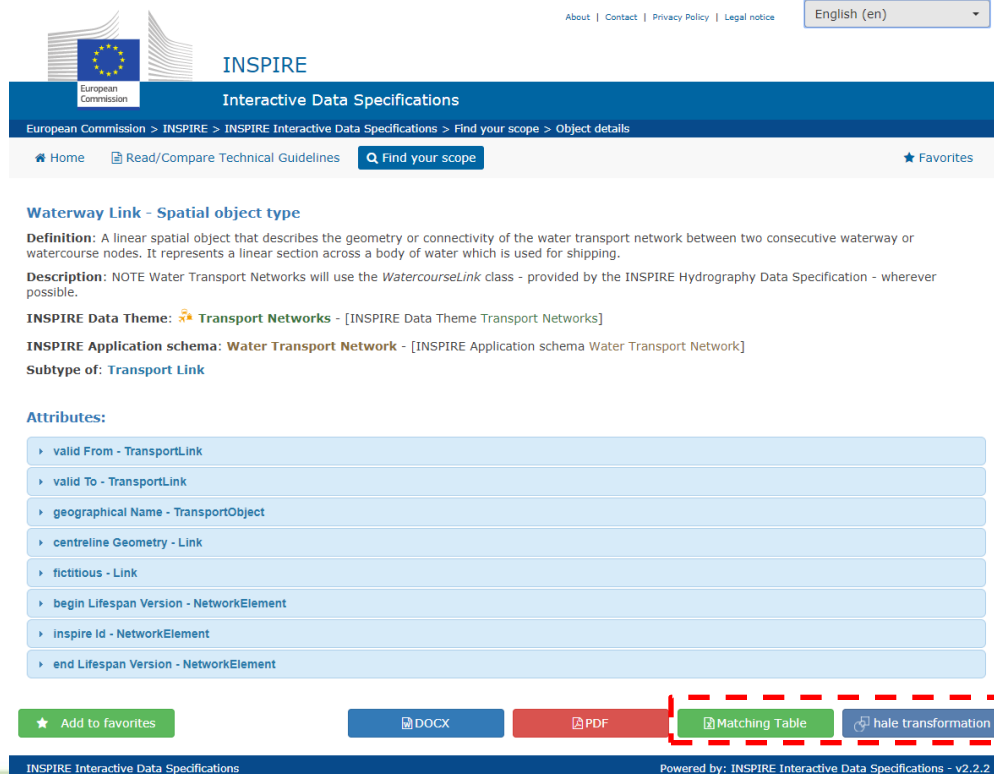
- *Matching table*

- *hale transformation* (permite abrir e ver as propriedades de determinado *Spatial object type*)



CATALOGUE OF INSPIRE OBJECTS

a catalogue of all spatial objects and their properties defined by INSPIRE in the alphabetic order. The user still can select one or more objects and continue with e.g. comparison with the data he/she administer.



The screenshot shows the INSPIRE Interactive Data Specifications web application. The header includes the European Commission logo, the INSPIRE logo, and navigation links: About, Contact, Privacy Policy, Legal notice. A language dropdown menu is set to English (en). The main navigation bar contains links for Home, Read/Compare Technical Guidelines, Find your scope (highlighted), and Favorites. The content area displays details for the 'Waterway Link - Spatial object type'. It includes a definition, a description, the INSPIRE Data Theme (Transport Networks), the INSPIRE Application schema (Water Transport Network), and the Subtype (Transport Link). A list of attributes is shown, including valid From, valid To, geographical Name, centreline Geometry, fictitious, begin Lifespan Version, inspire Id, and end Lifespan Version. At the bottom, there are buttons for 'Add to favorites', 'DOCX', 'PDF', 'Matching Table' (highlighted with a red dashed box), and 'hale transformation' (also highlighted with a red dashed box). The footer indicates the application is powered by INSPIRE Interactive Data Specifications - v2.2.2.

Processo de Harmonização

2. Mapeamento

- **Passo 3:** Comparar e documentar a informação inicial e final
 - *Matching table*

[illegible]

DS INSPIRE
"sitename"

Target Data model

Source Data model

Communication/agreements

Source schema

Modelo Inspire
*.xsd

CDG da entidade

Processo de Harmonização

2. Mapeamento

Como devo fazer?

2) Mapeamento

- Preenchimento da *matching table*

gml: identificador único

Application
schemas

Descrição
application schema

Multiplicity

Voidable

Atributos

Descrição dos atributos

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 spatial object. NOTE: An Namespace uniquely identifying the data source of the spatial object.	localId		
			The identifier of the particular version of the spatial object.	namespace		
				version	1	
		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 Namespace uniquely identifying the data source of the spatial object.	localId		
			The identifier of the particular version of the spatial object, with a maximum length of 25	namespace		
				version	1	
		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.					

2. Mapeamento

Matching Mapping tables

Schema mapping - regras de transformação a aplicar entre objetos mapeados. Esta fase pode incluir a reclassificação de valores e a conversão de tipos de dados (números, textos, geometrias) ou de sistemas de referência de coordenadas.

Schema transformation - processo de extração, transformação e carregamento dos dados. Este processo move os dados desde a sua origem no “esquema fonte” para o destino de acordo com o “esquema alvo”. Este processo é conhecido como ETL (*Extract, Transform & Load*).

Esquema alvo - Modelo *INSPIRE*

Application Schema 'Road Transport Network' (version 3.0)						
Type	Documentation	Attribute / Association role	Association role / Constraint documentation	Data type/Values / Code List/Enumerations	Multiplicity	Validable / Non-Validable
Road <small>SuperTypes: TransportLinkSetTransportObjectLinkSetNetworkElement</small>	A collection of road link sequences and/or individual road links that are characterized by one or more thematic identifiers and/or properties. EXAMPLE: Examples are roads characterized by a specific identification code, used by road management authorities or tourist routes, identified by a specific name.					
		geographicalName	A geographical name that is used to identify the road.	GeographicalName	0..1	validable
		beginLifeSpanVersion	Date and time at which this version of the road is created.	DateTime	1	validable
		inspireId	External object identifier of the spatial object NOTE: An external object identifier is a unique value.	Identifier	0..1	
		endLifeSpanVersion	Date and time at which this version of the road is deleted.	DateTime	0..1	validable
		inNetwork	The new links in which this road link is used.	Network	1..*	validable
		link	The set of links and link sequences that are associated with the road link set.	GeneralisedLink	1..*	
		validFrom	The time from which the transport link set is valid.	DateTime	1	validable
		validTo	The time from which the transport link set is no longer valid.	DateTime	0..1	validable
		post	Marker post along a route in a transport network.	MarkerPost	0..*	validable
		localRoadCode	Identification code assigned to the road.	CharacterString	0..1	validable
		nationalRoadCode	The national number of the road.	CharacterString	0..1	validable
RoadLink <small>SuperTypes: TransportLinkTransportObjectLinkGeneralisedLinkNetworkElement</small>	A linear spatial object that describes the geometry and connectivity of a road network between two or more points in the network. Road links can represent paths, bicycle roads, single carriageways, multiple carriageway roads and even fictitious trajectories across traffic squares.					
		geographicalName	A geographical name that is used to identify the road link.	GeographicalName	0..1	validable
		beginLifeSpanVersion	Date and time at which this version of the road link is created.	DateTime	1	validable
		inspireId	External object identifier of the spatial object NOTE: An external object identifier is a unique value.	Identifier	0..1	
		endLifeSpanVersion	Date and time at which this version of the road link is deleted.	DateTime	0..1	validable
		inNetwork	The new links in which this road link is used.	Network	1..*	validable
		centreLineGeometry	The geometry that describes the road link.	GM_Curve	1	
		fictitious	Indicator that the road link is fictitious.	Boolean*	1	
		endNode	The optional end node for the link. The end node is the node at which the road link ends.	Node	0..1	
		startNode	The optional start node for the link.	Node	0..1	
		validFrom	The time from which the transport link set is valid.	DateTime	1	validable
		validTo	The time from which the transport link set is no longer valid.	DateTime	0..1	validable

Transformação

Attribute	Documentation	Data Type / Code Lists /	Multiplicity	Visibility / Note
localId	A local identifier	Character string	1	
name space	Namespace	Character string	1	
versionId	The identifier of the	Character string	0..1	visibility
localId	A local identifier	Character string	1	
name space	Namespace	Character string	1	
versionId	The identifier of the	Character string	0..1	visibility

Esquema fonte

Source Data Model				NOTES
"File name" or URL	Name of attribute	Example of one data source value	Example of one data target value	
Roads.shp	Roads.[STNAME1]	TH_Road_[STNAME1]		Best Practice: the localid value should be the same as gmlid value
MS DataProvider.Product	ProviderName	if ([DataProviderName])RoadDataset		
		1.0		
			#TH_Newark_[Roads.ID]	
Roads.shp	Roads.ID		#TH_RoadLink_[Roads.ID]	
Roads.shp	Roads.ID		TH_RoadLink_[Roads.ID]	Best Practice: the localid value should be the same as gmlid value
MS DataProvider.Product	ProviderName	if ([LocalProviderName])RoadDataset		
		1.0		
			#TH_RoadNode_[Roads.ID]	
			#TH_RoadNode_[Roads.ID]	

Processo de Harmonização

2. Mapeamento

- **Passo 3:** Comparar e documentar a informação inicial e final
 - *Matching table*
 - Cada atributo do *source data* tem de ser mapeado para o atributo mais relevante do *target schema*.
 - Alguns atributos são do tipo "complexo" e por isso poderá ser necessário "expandir" a matching table.

14	AdministrativeUnit	-- Name -- administrative unit Unit of administration where a Member State has and/or exercises jurisdictional rights, for local, regional and national governance.	beginLifespanVersion	-- Name -- begin lifespan version Date and time when the administrative unit began to exist	DateTime	1	voidable
15			country	-- Name -- country Two character country code	CountryCode* BE* DE* FR* IT* NL* ES*	1	
16			endLifespanVersion	-- Name -- end lifespan version Date and time when the administrative unit ended to exist	DateTime	0..1	voidable
17			geometry	-- Name -- geometry Geometrical representation of the administrative unit	GM_MultiSurface	1	
18			inspireId	-- Name -- inspire id Unique identifier of the administrative unit	Identifier	1	
19			name	-- Name -- name Official name of the administrative unit	GeographicalName	1..*	
20							

Application Schema 'Base Types' (version 3.3rc3)						
Type	Documentation	Attribute Association role / Constraint	Attribute / Association role / Constraint	Values / Enumerations	Multiplicity	Voidable / Non-Voidable
SpatialData Set	Identifiable collection of spatial data. NOTE The type SpatialDataSet is offered as a pre-defined type for spatial data sets	identifier	Identifier of the spatial data set	Identifier	1	
		metadata	Metadata of the spatial data set	MD_Metadata	1	voidable
Identifier	External unique object identifier published by the responsible body, which may be used by external applications to reference	localId	A local identifier, assigned by the data provider. The Namespace uniquely identifies the data source	CharacterString	1	
		namespace		CharacterString	1	
		versionId	The identifier of the version of the data	CharacterString	0..1	voidable

Processo de Harmonização

2. Mapeamento

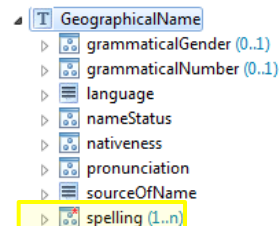
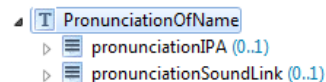
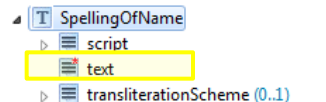
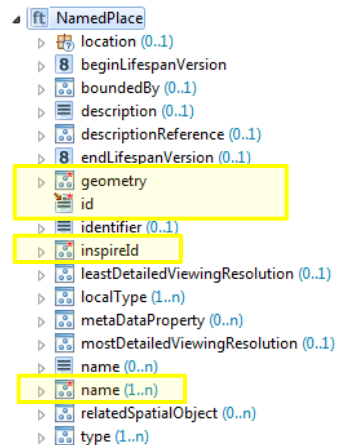
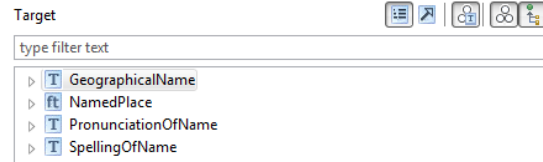
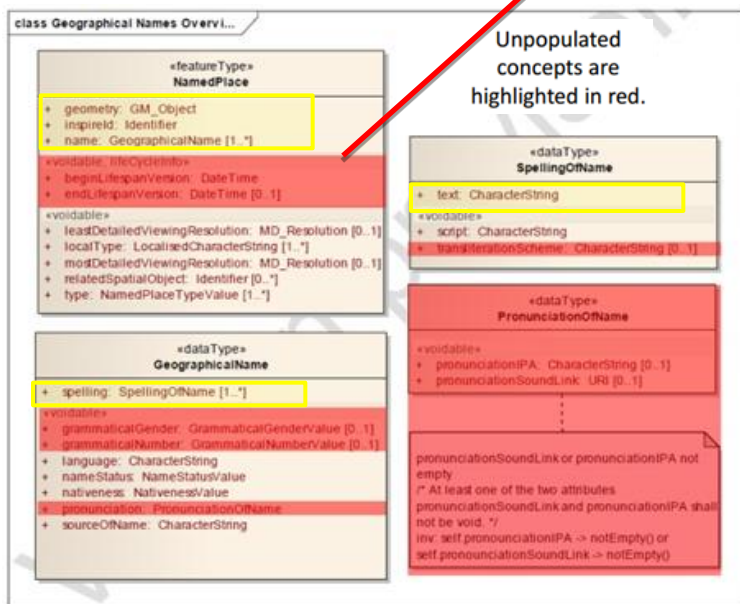
Como devo fazer?

1) Mapeamento (exemplo)

- No âmbito dos GTI-TE: Discutir assuntos “Passo a passo”, propriedade a propriedade, atributo a atributo, ou seja, no propriedade XPTO todos temos de falar a mesma linguagem (mesma semântica) para determinado valor do atributo para os dados reportados. P.S: Cuidado porque a representação de um objeto geográfico em determinado intervalo de escala pode levar a interpretações diferentes ou a *outputs* errados.

Obrigatório

<voidable>
unpopulated
unknown
withheld



Geographical Names
Target Schema - UML

Geographical Names
Target Schema – Software HALE

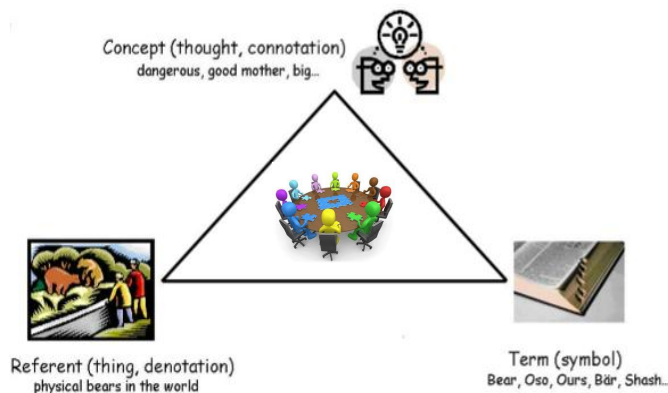
Processo de Harmonização

2. Mapeamento

Como devo fazer?

1) Mapeamento (exemplo)

- No âmbito dos GTI-TE: Harmonizar conceitos, definições e terminologias harmonizadas por todas as entidades e que tenham uma correspondência biunívoca com o modelo de dados do INSPIRE, ou seja, analisar se é possível “encaixar” valores (*codelists*) de determinada propriedade / atributo.



Semantics:

Semantics – the study of what something means

Terms – the basic semantic units for conveying concepts. Usually single-word nouns.

Dictionary – contains definitions and pronunciations

Thesaurus – list words grouped together according to similarity of meaning (synonyms and antonyms)

Ontology's – to describe the meanings in a systematic way

- o *There is a philosophical approach and a computer science approach*

Processo de Harmonização

2. Mapeamento

- Preencher a *matching table*. Por exemplo, comparativamente pode-se avaliar o modelo de dados elaborado pela RAA

- Identificar o modelo de dados INSPIRE (*target*)

<http://inspire.ec.europa.eu/schemas/>

- Identificar dados de preenchimento obrigatório (*mandatory*)

- Identificar propriedades que podem ser *voidable*

- Identificar as correspondências possíveis

- Identificar as lacunas

O modelo de dados INSPIRE define uma lista de enumerações a aplicar quando o objecto espacial não contém o atributo esperado. Os motivos *voidable* podem ser *unpopulated*, *unknown* e *withheld*.

unpopulated - A propriedade não faz parte do CDG mantido pelo produtor de dados, embora esta característica possa existir no mundo real. Esta propriedade será atribuída a todos os objetos espaciais desse CDG;

unknown - O valor correcto não é conhecido e não é possível produzi-lo, embora este possa existir;

withheld - A propriedade pode existir, mas é confidencial e não é divulgada pelo fornecedor de dados.

K9										
A	B	C	D	E	F	G	H	I	J	K
Address (Endereço)	Tabela	Uma identificação do local fixo da propriedade mediante uma composição estruturada de identificadores e topónimos.	status	Validade do endereço no ciclo de vida (versão) do objecto geográfico do endereço.		voidable	retired proposed reserved alternative	Esta lista de códigos deve ser gerida num registo comum de listas de códigos.	Um endereço deve ter um objecto geográfico cuja componente de endereço correspondente à unidade administrativa é de nível 1 (país). Um endereço deve ter exactamente uma posição geográfica predefinida (o atributo «default» do objecto geográfico «GeographicPosition» deve ser «true» (verdadeiro)).	
			validFrom	data e hora de que esta versão do endereço foi ou será válida no mundo real.	DateTime	voidable				
			validTo	data e hora em que esta versão do endereço deixou de existir ou deixará de existir no mundo real.	DateTime	voidable				
			beginLifespanVersion	Data e hora em que esta versão do objecto geográfico foi inserida ou alterada no conjunto de dados geográficos.	DateTime	voidable				
			endLifespanVersion	Data e hora em que esta versão do objecto geográfico foi substituída no conjunto de dados geográficos ou dele retirada.	DateTime	voidable				
			building	Edifício ao qual o endereço está atribuído ou associado.	AbstractConstruction	voidable		Relação com AbstractConstruction (Building /BuildingPart)		
			component	Indica que a componente do endereço faz parte do endereço.	AddressComponent			Relação com AddressComponent (AdminUnitName/ PostalDescriptor/AddressAreaName/ThoroughfareName)		
			parcel	Parcela cadastral à qual este endereço está atribuído ou associado.	CadastralParcel	voidable		Relação com CadastralParcel		
			parentAddress	Endereço principal (endereço-mãe) com o qual este (sub)endereço está estreitamente ligado.	Address	voidable		Relação com Address		

Processo de Harmonização

2. Mapeamento

• A harmonização de dados é difícil porque, por exemplo (neste caso), conseguimos visualizar que pretende-se que haja ligações do tema dos Endereços (*Addresses*) com o tema dos Prédios (*Cadastral Parcels*) e dos Edifícios (*Buildings*)... mas não é só, porque este tema também se relaciona com a Toponímia (*Geographical Names*) e “a rede” de Transporte Rodoviário (*Road Transport Network*) bem como as Unidades Administrativas (*Administrative Units*).

• Como existe a possibilidade dos atributos serem *Voidable* pode-se preencher o mínimo possível (nesta fase) em dados que ainda existam no no nosso país. A médio prazo quando muitos conceitos estiverem assimilados e cimentados, os GTI-TE poderão abordar estas questões e avaliar a possibilidade das entidades cooperarem para discutir como tudo isto se poderia interligar e conjugar na respetiva missão de cada entidade.

E146				
A	B	C	D	E
41		carVithTrailer	Veículo particular com reboque atrelado.	
42		deliveryTruck	Camião de dimensões relativamente reduzidas, principalmente utilizado na distribuição de bens e materiais.	
43		emergencyVehicle	Veículo utilizado para intervenções de emergência, nomeadamente, mas não exclusivamente, os veículos da polícia, ambulâncias e veículos de bombeiros.	
44		employeeVehicle	Veículo conduzido por um empregado de uma organização, que é utilizado de acordo com os procedimentos previstos pela mesma.	
45		facilityVehicle	Veículo destinado a ser utilizado unicamente numa área localizada, numa propriedade privada ou de acesso restrito.	
46		farmVehicle	Veículo geralmente associado a actividades agrícolas.	
47		highOccupancyVehicle	Veículo lotado com um número de ocupantes igual (ou superior) ao número mínimo de passageiros especificado.	
48		lightRail	Veículo de transporte de tipo comboio, que circula apenas numa rede ferroviária num perímetro limitado.	
49		mailVehicle	Veículo utilizado na recolha, transporte ou distribuição de correio.	
50		militaryVehicle	Veículo autorizado por uma autoridade militar.	
51			Velocidade de curso controlada, associada com motor de combustão interna, de	
K9				
A	B	C	D	E
52			status	Validade do endereço no ciclo de vida (versão) do objecto geográfico do endereço.
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				



Processo de Harmonização

2. Mapeamento

FuncionalRoadClass (Classe funcional da estrada)	Tabela	Uma classificação baseada na importância da estrada na rede rodoviária. Este tipo é um subtipo de «TransportProperty».	networkRef	Referência geográfica da propriedade relacionada com a rede.	NetworkReference	voidable	Tipo de dados que implica relação com SimplePointReference
			validFrom	O momento em que a propriedade de transporte começou a existir no mundo real.	DateTime	voidable	
			validTo	O momento em que a propriedade de transporte deixou de existir no mundo real.	DateTime	voidable	
			functionalClass	Classificação funcional do segmento da estrada na rede rodoviária.	FunctionalRoadClassValue	mainRoad firstClass secondClass thirdClass fourthClass fifthClass sixthClass seventhClass eighthClass ninthClass	Valores autorizados para a enumeração «FunctionalRoadClassValue» Esta propriedade só pode ser associada a um objecto geográfico que faz parte de uma rede de transporte rodoviário.



FunctionalRoadClassValue - Classe funcional da estrada	mainRoad	As estradas mais importantes numa determinada rede.
	firstClass	As segundas estradas mais importantes numa determinada rede.
	secondClass	As terceiras estradas mais importantes numa determinada rede.
	thirdClass	As quartas estradas mais importantes numa determinada rede.
	fourthClass	As quintas estradas mais importantes numa determinada rede.
	fifthClass	As sextas estradas mais importantes numa determinada rede.
	sixthClass	As sétimas estradas mais importantes numa determinada rede.
	seventhClass	As oitavas estradas mais importantes numa determinada rede.
	eighthClass	As nonas estradas mais importantes numa determinada rede.
	ninthClass	As estradas menos importantes numa determinada rede.

Correspondência com a Base de Dados em uso na DRIG

Corresponde à tabela TipoClassif1.

De notar que internamente, ou seja, no front-end com o utilizador, pretende-se continuar a usar os termos definidos no TipoClassif1.

Enumerações - 7.7.2.1. Valores autorizados para a enumeração «FunctionalRoadClassValue»

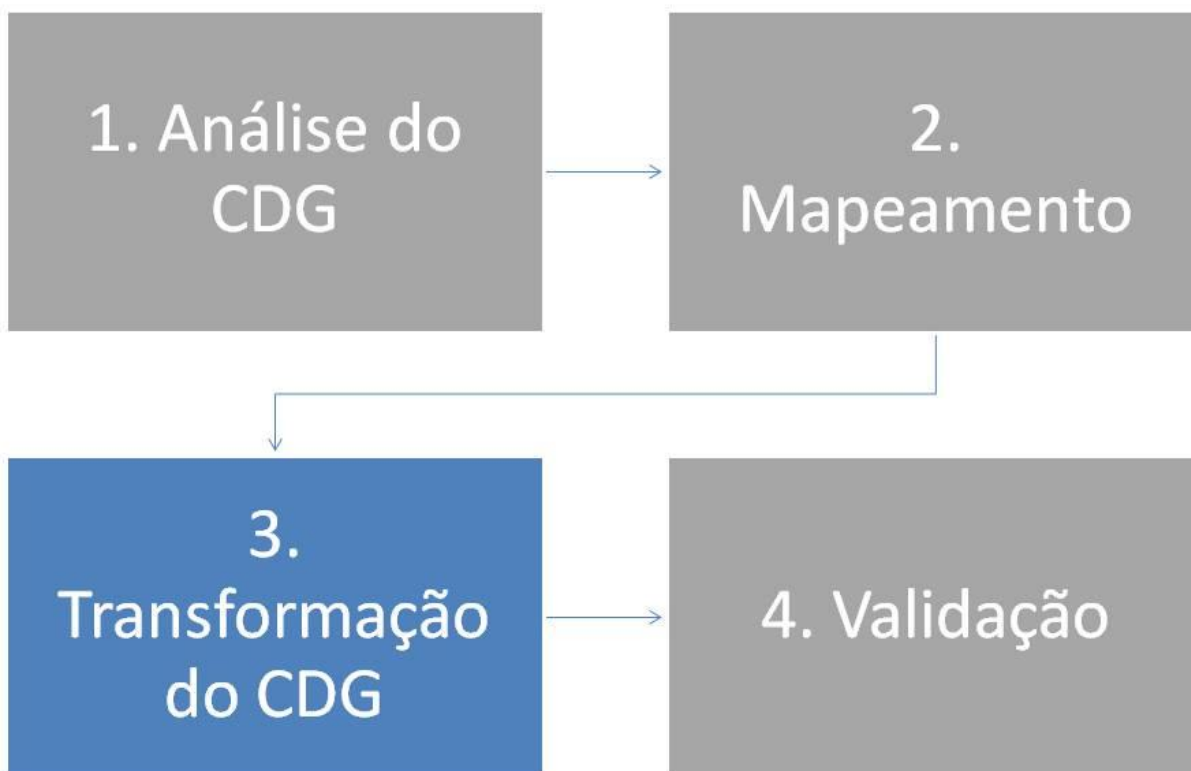
VALOR	DEFINIÇÃO	TipoClassif1
mainRoad	As estradas mais importantes numa determinada rede.	Estrada Regional Principal
firstClass	As segundas estradas mais importantes numa determinada rede	Estrada Regional Complementar
secondClass	As terceiras estradas mais importantes numa determinada rede.	Caminho Municipal
thirdClass	As quartas estradas mais importantes numa determinada rede.	Rede Viária Florestal
fourthClass	As quintas estradas mais importantes numa determinada rede.	Rede Florestal Divisional
fifthClass	As sextas estradas mais importantes numa determinada rede.	



- Adoção de uma linguagem universal para todas as instituições em que se consiga identificar univocamente cada objeto no contexto português e adaptado ao contexto europeu.

Exemplo: perceber se uma *mainroad* vai ser uma estrada com características de Auto-estrada e mais algum tipo de estrada (caracterizar). Como tal, é preciso harmonizar o conceito de “Estradas mais importantes numa determinada rede”.

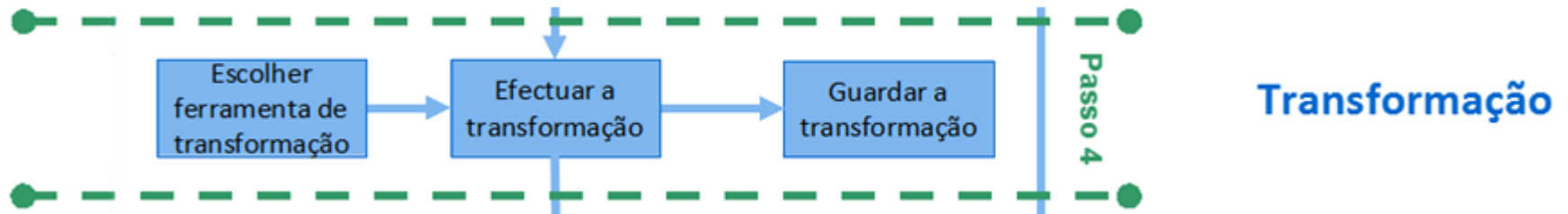
- Processo 3



Processo de Harmonização

3. Transformação do CDG

- **Passo 4:** Definir a ferramenta a utilizar e efetuar a transformação



Processo de Harmonização

3. Transformação do CDG

Harmonização CDG – E-learning, Workshop	LINKVIT (GISIG E-learning Platform)	INSPIRE Workshop - Data harmonization 2015 (HALE)
	a) Data Harmonisation [Link PCT]	Transformation Workshop [Link PCT]
	b) Procedures for Data and Metadata Harmonization [Link PCT]	
	c) Examples of Data Transformation [Link PCT]	

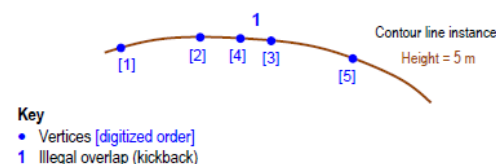
Topics

- Schema matching
- Matching table
- Schema mapping
- Translation rules
- Quality of schema mapping and matching

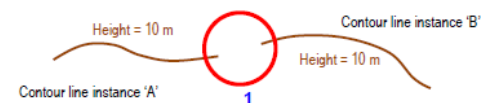
Exemplo: Varchar -> Date

"2006-09-12"	2006-09-12
"13/2-05"	2005-02-13
"2007/20/07"	2007-07-20
"Christmas eve, 1996"	?

Example – Self-overlap of a contour line instance.



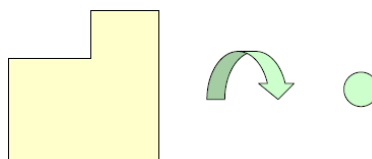
Example – Connectivity between two consecutive contour line instances with the same height which have to be connected.



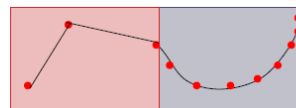
Matching process

- To start matching process you need to:
 - Identify feature types in both the source schema and the target schema
 - Identify structural properties of the feature types
 - Identify attribute names in both schemas
 - Identify data-value types and characteristics
- The matching process can be performed manually as a desk study or using automated tools that uses intelligent techniques.
- The result of schema matching is to make sure that features and attributes in both schemas are semantically related.
- The matching table will be used during mapping.

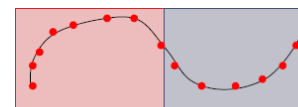
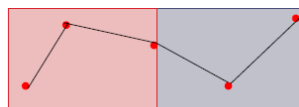
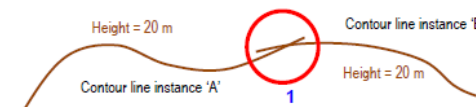
Exemplo: Polígono -> Ponto



Resampling or filtering?



Example – Connectivity between two consecutive contour line instances with the same height which have to be connected.



Processo de Harmonização

3. Transformação do CDG

Aparte: A transformação de dados por vezes terá que ter em conta um ambiente de produção de dados e também uma perspetiva de generalização dinâmica dos dados para que estes sejam visíveis com diferentes representações geográficas, alfanuméricas e regras de simbologia, consoante os níveis de zoom (intervalos de escala) de disponibilização dos dados.



GEO tools ELF within WP's (2):

WP#	Tool	Description
4	GoPublisher	Transformation solution to harmonise and serve INSPIRE data
4	HALE	Open Source Transformation Tool
4	1Spatial Cloud	Tool to report on quality of data based on different quality parameters
4	ArcGIS Data Reviewer	Tool to report on quality of data based on different quality parameters
4	PP Repair	Check, find errors and automatically repair data
4	IGN France Change Detection Tool	detection of differences between 2 releases of the same vector database
4	Geo Product Finder	interface to find, view, compare and access the geo-information
4	1Integrate	Edge matching tool of 1Spatial
4	ArcGIS Edge Match tool	Edge matching tool of Esri
4	SLD Editor	Tool to define symbology for geo-objects
4	IGN France Generalisation tool	Generalize ELF Regional data (medium scale) in order to produce ELF Global data (small scale). All themes and feature classes of the ELF Global level should be processed with this tool.
4	1Generalise	Generalisation tool of 1Spatial
4	Gen.tool KADNL & Esri	Tool and Procedure for generalisation from Master 1 to 2 based on ArcGIS Modelbuilder from Esri

http://inspire.ec.europa.eu/events/conferences/inspire_2014/pdfs/20.06_2_09.00_Dorus_Kruse.pdf

Processo de Harmonização formato dos dados *versus* ferramentas

FERRAMENTAS E APLICAÇÕES:

UML CASE Tools:

- Eclipse (open source)
- Enterprise Architect (comercial)
- Visio (comercial)
- Altova [XMLSpy] (comercial)



Schema matching
Schema mapping
Schema transformation

Schema translation tools

(Desktop / Server / Cloud):

- Hale Humboldt (open source)
- Altova [Mapforce] (comercial)
- FME (comercial)
- ArcGIS for INSPIRE (Data Interoperability extension)



The ecosystem of tools



Tool	Data Transformation	Metadata management	Network Services Publishing	Notes
HALE	*		*	Exports to GML
FME	*	*	*	Commercial
Geokettle	*	*		Some functionality not mature enough
Geoserver	* app-schema extension		*	INSPIRE compliant Services → Extension → Complex feature types (limited)
Mapserver			*	INSPIRE compliant Services (view, discovery, partial download)
Deegree	*	*	*	INSPIRE compliant Services
Geonetwork		*	*(CSW)	INSPIRE compliant Services & metadata
Geoportal Server		*	*	INSPIRE compliant Services & metadata
Geoconverter	*			
ArcGIS for INSPIRE	*	*	*	Commercial
Snowflake	*		*	Commercial

Functionality	HALE	GoPublisher
Format of source dataset	<ul style="list-style-type: none"> • Shape file • WFS • PostGIS • CSV (for non-spatial data) 	<ul style="list-style-type: none"> • ORACLE • PostGIS • SQL Server and MS Access or Excel file (for non-spatial data)
Conversion of values	<ul style="list-style-type: none"> • Predefined classification function 	<ul style="list-style-type: none"> • SQL scripting
Mapping of INSPIRE complex data type	<ul style="list-style-type: none"> • Predefined function for "inspireld" and "Geographical name" 	<ul style="list-style-type: none"> • Manual creation of the data type structure

ArcGIS for INSPIRE:

<http://geospatialworldforum.org/speaker/SpeakersImages/Roberto%20Lucchi.pdf>

Processo de Harmonização

3. Transformação do CDG

Como devo fazer?

3) Transformação do CDG (nível *desktop*, *server*, *cloud*)

- Escolher ferramenta ETL (*Extract-Transform-Load*)
- Construir *workflow* de mapeamento (Tentativa/Erro)
 - Se não for possível, deve-se rever se é possível alterar procedimentos (passos intermédios) a montante.

Software Open Source:

- hale STUDIO (*desktop*, *server*)
- GeoKettle



GeoKettle
Open Source Spatial ETL



hale
STUDIO

Outras soluções:

hale CONNECT (*cloud*), inspiregis (*server*, *cloud*)



hale
CONNECT



inspiregis



Public Cloud



Private Cloud



On Premise

Software proprietário:

- FME (*desktop*, *server*, *cloud*) (*INSPIRE solution pack*)
- ArcGIS Data Interoperability (pacote *ARCGIS for INSPIRE*)
- Snowflake (Go Loader; Go Publisher)
- ...



Software gratuito:

- Geobide (Geoconverter, etc.)



Geobide

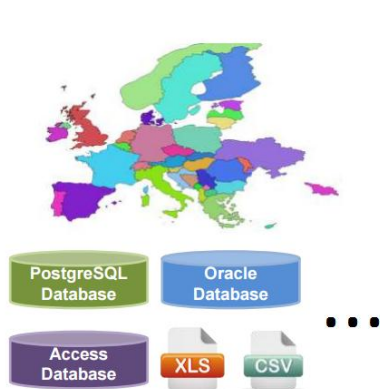
Cuidado: Com os dados existentes podemos nunca conseguir transformar todos os CDGs para o respetivo *Target Schema*!!!
Experimentar (Tentativa/Erro)

Processo de Harmonização

3. Transformação do CDG

Exemplo: Como transformamos / validamos os dados harmonizados?

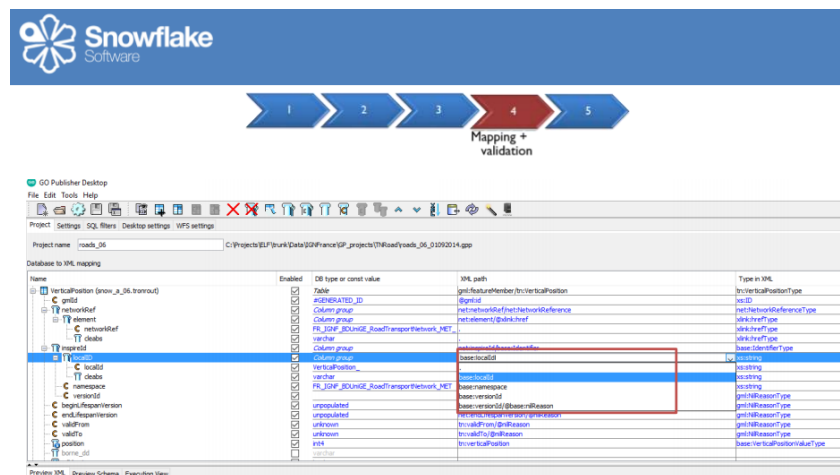
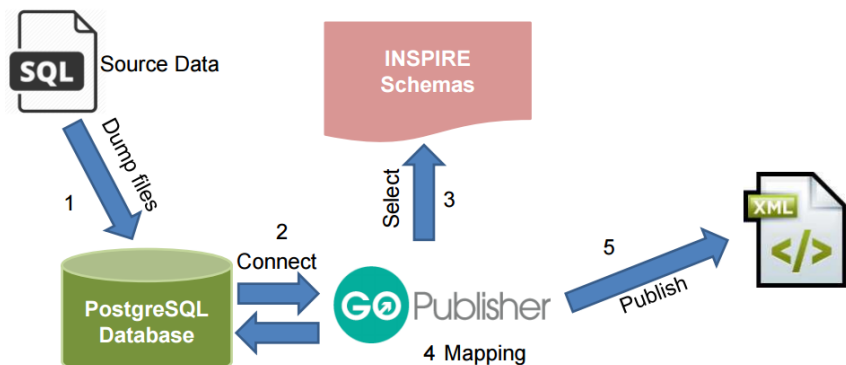
Data Transformation



Difficult to share data



Easy to share data



INSPIRE validation workshop (Junho 2016):

<https://www.youtube.com/watch?v=BMjmjJ8u0Ds&feature=youtu.be>

http://eurogeographics.org/sites/default/files/20160603AlbertoOlivares_SnowflakeSoftware.pdf

Processo de Harmonização

3. Transformação do CDG

Exemplo: Como transformamos / validamos os dados harmonizados?

Coordinate System Extractor



What is FME?

FME transforms data to use and share.

- Convert data
- Transform data
- Share data
- Integrate data
- Validate data
- And more



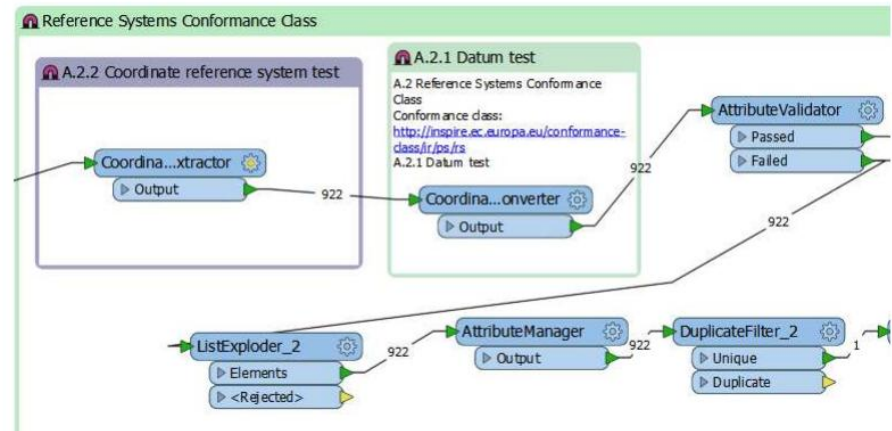
CONNECT. TRANSFORM. AUTOMATE.

#fmewebinar

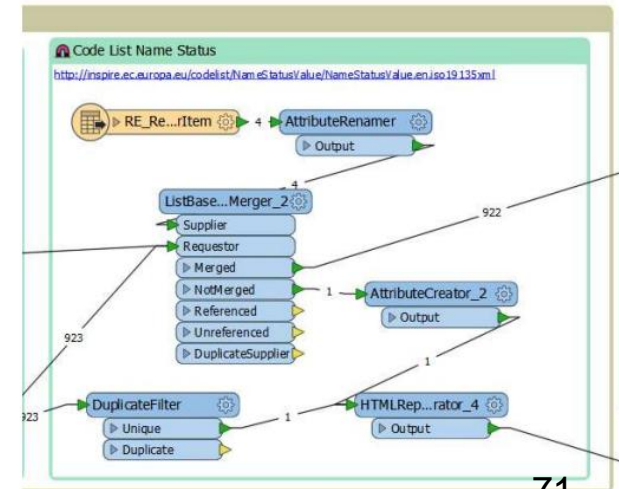
INSPIRE validation workshop (Junho 2016):

<https://www.youtube.com/watch?v=yFRyMdmXCrM&feature=youtu.be>

http://eurogeographics.org/sites/default/files/20160603ValidationFME_KenBragg.pdf



Read Code Lists

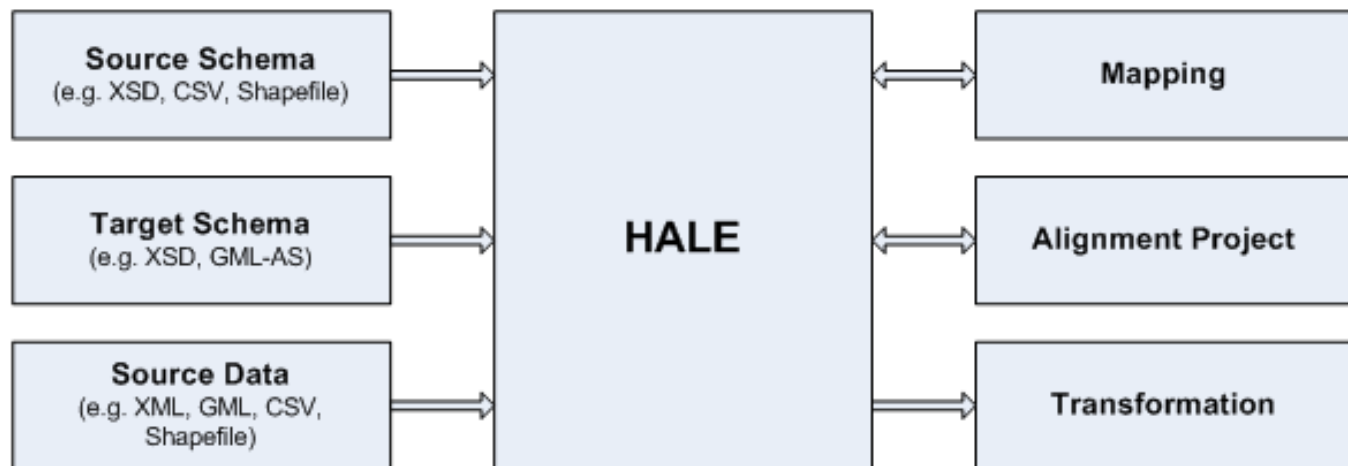


Processo de Harmonização

3. Transformação do CDG

hale STUDIO - Alignment Editor

- Define, avalia e executa mapeamento entre modelos de dados:
 - Ferramenta ETL Open Source
 - Bom suporte de XML/GML
 - Adaptações feitas relativamente à Diretiva INSPIRE (Code Lists e outras características)
 - Transformação em tempo real com feedback
 - Validação online
 - Permite a criação de scripts, etc.



Processo de Harmonização

3. Transformação do CDG

hale STUDIO - Alignment Editor

- *Software* hale STUDIO tem os princípios básicos do que se pretende:
 - Identificação das funções de mapeamento / transformação
 - Definir mapeamento / transformação
 - Exportação para GML 3.2.1

General	Retype	Estabelece o mapeamento entre as <i>feature type</i> (do <i>source schema</i> para o <i>target schema</i>)
	Merge	Funde múltiplos atributos do <i>source schema</i> num único atributo no <i>target schema</i> .
	Join	Junta múltiplos atributos do modelo de origem, num único atributo no <i>target schema</i> .
	Create	Cria objectos espaciais de um tipo de <i>schema</i> específico.
	Date extraction	Extraí a data de um campo do tipo texto (string).
	Rename	Copia a propriedade de um atributo do <i>source schema</i> para um atributo do <i>target schema</i> .
	Assign	Atribui um valor a um atributo do <i>target schema</i> , quando não existe correspondência do lado do <i>source</i> .
	Classification	Permite a utilização de <i>code lista</i> e outros domínios.
	Formatted string	Cria uma <i>string</i> com um formato baseado num padrão ou em variáveis de entrada (quando existe uma correspondência com o valor de entrada da função entre {} a <i>string</i> é validada, caso contrário o texto não é transformado).
	Inline transformation	Usa uma transformação tipo definida no mapeamento para qualquer outra transformação que tenha a mesma correspondência com o <i>source</i> e o <i>target schemas</i>

Geometric	Assign (bound)	Atribui um valor a um atributo do <i>target schema</i> , quando não existe correspondência do lado da <i>source</i> .
	Ordinates to point	Cria um ponto a partir de atributos com coordenadas (valores de X e Y são obrigatórios, Z é opcional)
	Network expansion	Cria um <i>buffer</i> em torno de um objecto geométrico
	Calculate length	Calcula o comprimento de um objecto geométrico
	Calculate area	Calcula área de um objecto geométrico
INSPIRE	Centroid	Calcula o centróide da geometria dos CDG de origem e passa essa informação para a propriedade selecionada do <i>target schema</i> .
	Compute extent	Calcula a extensão geométrica com base em todas as geometrias dos dados originais. Estão disponíveis as opções <i>Bounding box</i> , <i>Convex Hull</i> e <i>União</i> .
	Aggregate	Agrega geometrias semelhantes.
INSPIRE	Reproject Geometry	Projeção do sistema de coordenadas da geometria.
	INSPIRE Identifier	Cria a estrutura do identificador único INSPIRE. Para a parte local do ID a propriedade <i>source</i> é utilizada, enquanto o <i>namespace</i> é derivado da informação fornecida pelo país, fornecedor e pelo produto.
	Geographical Name	Cria a estrutura do INSPIRE Geographical Name, para uma ou mais propriedades <i>source</i> que são usadas como <i>spelling</i> .
Numeric	Mathematical Expression	Define um valor utilizando uma expressão matemática com suporte para variáveis
	Generate sequential ID	Cria uma chave única sequencial para cada objecto. É possível acrescentar texto antes e/ou depois do ID.
Groovy	Groovy retype	Criação de funções de transformação personalizadas, através de linguagem de programação.
	Groovy create	
	Groovy merge	
	Groovy join	
	Groovy script (greedy)	

Quadro 7. Lista das funções de transformação disponíveis no HALE (Fonte: Barreira 2013; User's manual, 2016)

Processo de Harmonização

3. Transformação do CDG

GTI-TE-4



Trabalho desenvolvido no INE

Tip!



Partindo do pressuposto que podemos implementar a diretiva com software Open Source

1. Criação de Metadados (*Gema*)
2. Preparação e análise dos Conjuntos de dados Geográficos (*Qgis* + *Gaia*)
3. Harmonização dos dados Geográficos (*Hale*)
4. Validação do GML harmonizado (*Hale* + *oxygen XML* + *EnvPlus*)
5. Carregamento do GML na Base de dados (*postGres* + *PostGis*)
6. Publicação de serviços Wms (*geoserver* + *plugin Inspire* + *Linux*)
7. Serviço de descarregamento + *GEORSS*

Aprender a “trabalhar” com o HALE [ver manual]

❑ **Humboldt Alignment Editor (HALE) – (Parte 1 e 2) [partilha de ecrã]** (3ª Sessão Técnica RAA, 18 e 24 de maio 2016):

❑ <https://www.youtube.com/watch?v=GeyWAjxPI8&feature=youtu.be> (Parte 1)

❑ https://youtu.be/BI_RJ-gnXQ4 (Parte 2)

❑ http://www.ideia.azores.gov.pt/docs/Documentos/3sessao_Humboldt%20Alignment%20Editor%20%28HALE%29.pdf (Parte 1)



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

❑ **Harmonização de dados de Sondagens Geotécnicas com base no Modelo de Dados INSPIRE RAA e no HALE** (3ª Sessão Técnica RAA, 18 e 24 de maio 2016):

❑ <https://youtu.be/IK9tRG1XFBg>

❑ http://www.ideia.azores.gov.pt/docs/Documentos/3Sessao_Harmonização%20de%20dados%20de%20Sondagens%20Geotécnicas%20com%20base%20no%20Modelo%20de%20Dados%20INSPIRE%20RAA%20e%20no%20HALE.pdf

Processo de Harmonização

3. Transformação do CDG

Como devo fazer?

3) Transformação do CDG

- A estrutura dos dados que pretendemos transformar (*Source Schema*) pode condicionar a escolha da ferramenta ETL (Extract-Transform-Load)

- Por exemplo, com o *software* Hale conseguimos:

▪ Vários formatos para “dados de entrada”, incluindo *online* (URL, WFS), base de dados (*PostgreSQL/PostGIS, SpatiaLite*):

MS Access Database file (*.mdb)
XML schema (*.xsd, *.xml)
hale Schema Definition (JSON) (*.hsd.json, *.json)
CSV file (*.csv)
GZipped hale Schema Definition (*.hsd.gz, *.haleschema.gz)
MS OOXML Format Spreadsheet (XLSX) (*.xlsx)
Shapefile (*.shp)
Excel Spreadsheet (XLS) (*.xls)
SpatiaLite Database (*.sqlite)
hale Schema Definition (*.hsd, *.haleschema)

<http://www.ogcnetwork.net/node/1629>

▪ Vários formatos para “dados de saída”, WFS-T [experimental] !?

CSV file
Database (JDBC) [experimental]
GML (FeatureCollection)
GML (INSPIRE SpatialDataSet) [deprecated by INSPIRE]
GML (WFS 2.0 FeatureCollection)
GeoJSON
JSON
SQLite/SpatiaLite Database [experimental]
WFS-T (Direct upload) [experimental]
WFS-T (Partitioned upload) [experimental]
XLS file
XML (Custom root element)

