



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

» Harmonização de CDG's com Hale

«

DMSI/GEO

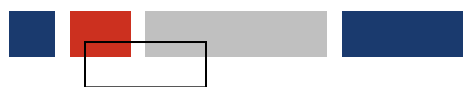
 (05-07-2016)

»





Harmonização

- O que é a Harmonização?
- Para **Hale**
 - “*The alignment is the mapping between source and target schemas. It defines relations between source and target entities (types or properties). Based on the defined relations a transformation is derived.*”
- Segundo **Dean M. Hintz**
 - “*Core to the harmonization workflow is the transformation process which reshapes source schema and geometry to match the required destination structure.*” (Hintz,2012,1)



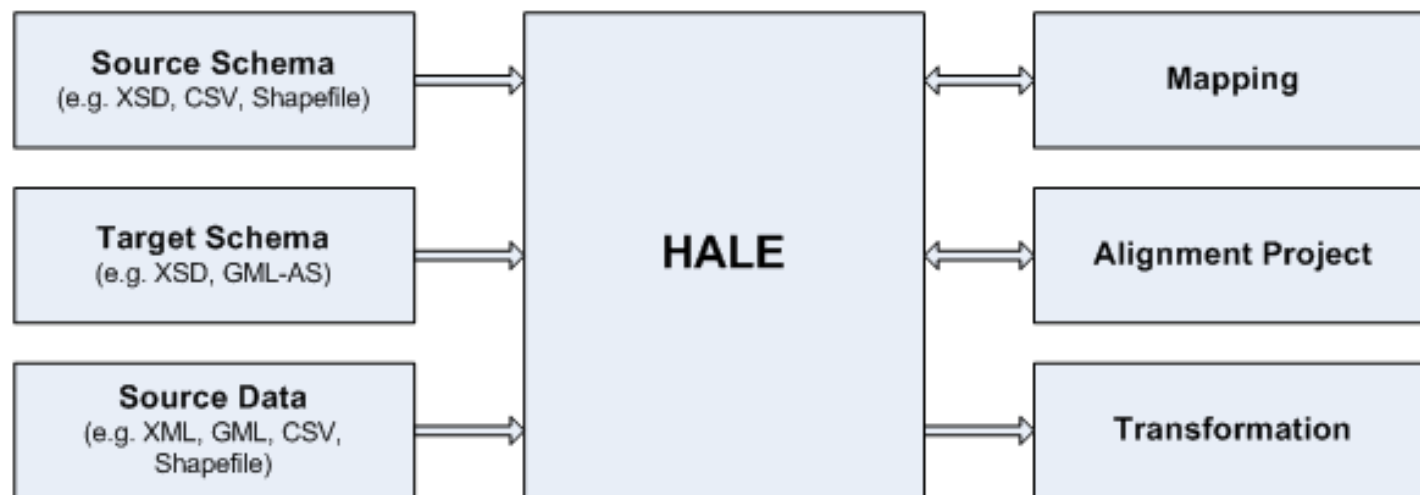
- **Antes de começar a trabalhar com Hale ou outra ferramenta de harmonização**
- Ler (e reler) as especificações dos dados!
 - ***Data Specification - Technical Guideline***
 - Modelos UML
- Preparar com antecedência a Matching Table
- Procurar exemplos de CDG's já harmonizados nos ***Thematic Clusters***.
- Partilhar conhecimentos e experiências com a comunidade!



- **H**UMBOLDT **A**Lignment **E**ditor 
- Software utilizado para criar mapping entre schemas diferentes e aplicar a transformação resultante
- *Software Open Source*
- Fornece uma interface gráfica rica, textual e especificamente aprovado para especialistas em IG
- *Feedback* instantâneo sobre o processo de harmonização de dados 
- [Download](#) versão 2.9.4 (2015-11-01) versão 32 e 64 bit para windows, Mac OS, Linux

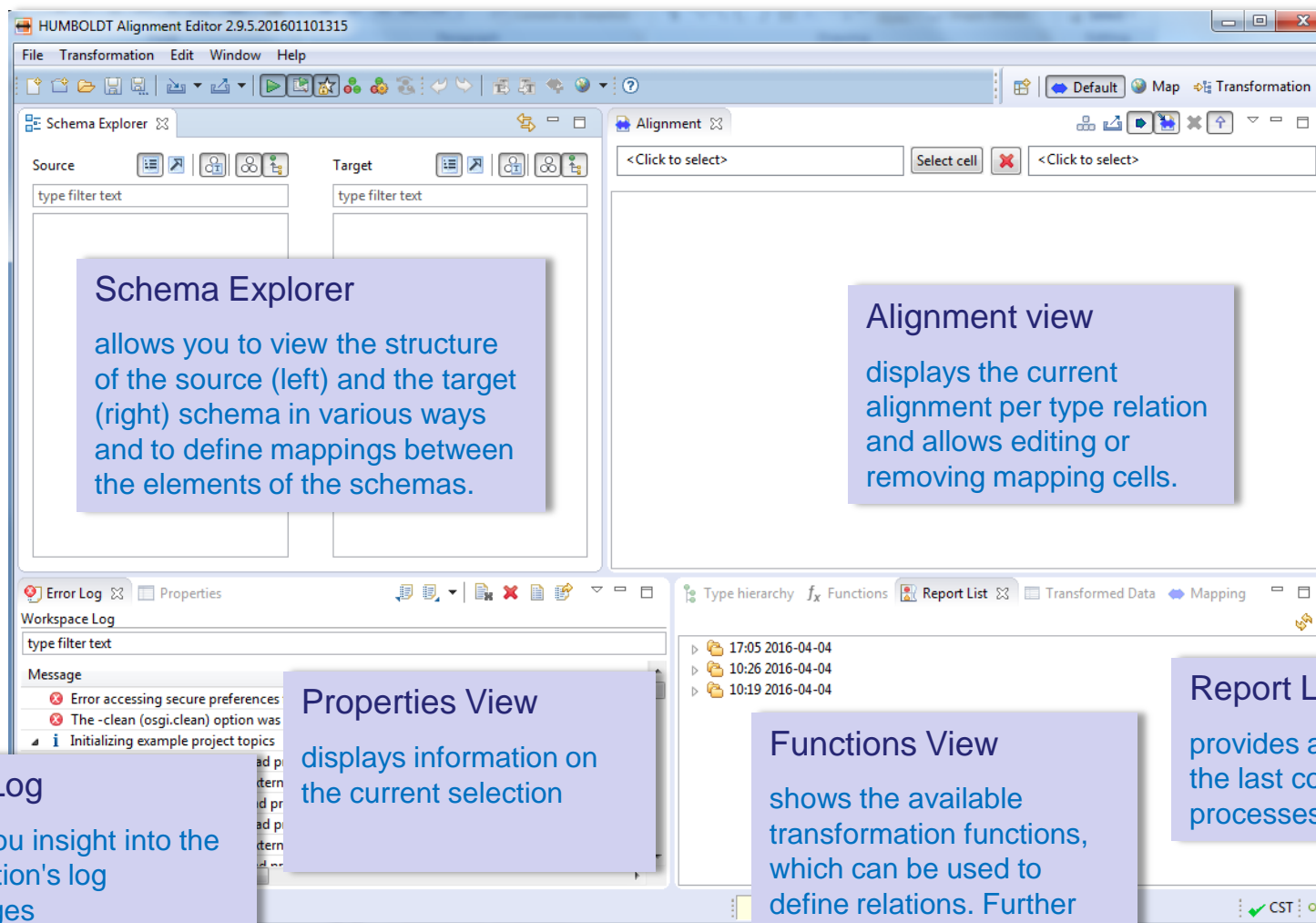


Hale





Hale Interface



Schema Explorer

allows you to view the structure of the source (left) and the target (right) schema in various ways and to define mappings between the elements of the schemas.

Alignment view

displays the current alignment per type relation and allows editing or removing mapping cells.

Properties View

displays information on the current selection

Error Log

gives you insight into the application's log messages

Functions View

shows the available transformation functions, which can be used to define relations. Further information on a selected function will be displayed in the *Properties* view.

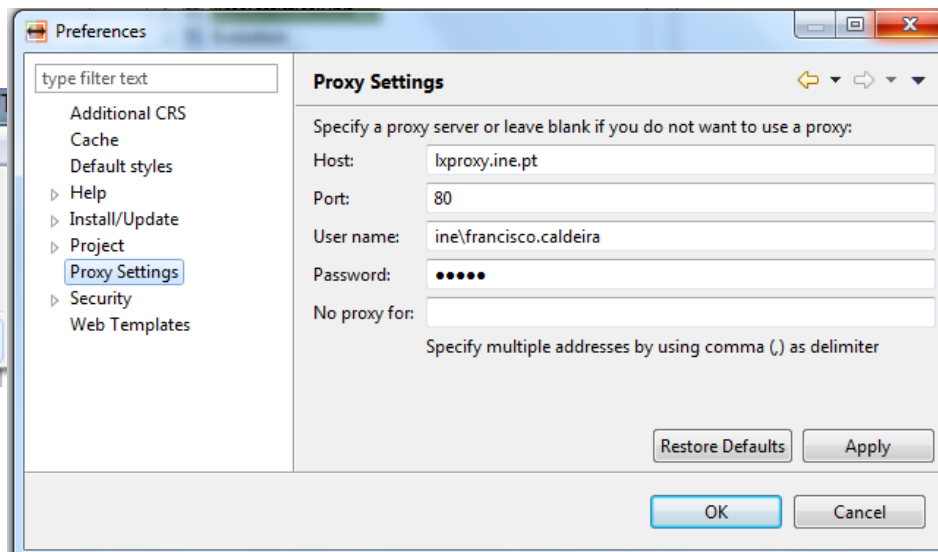
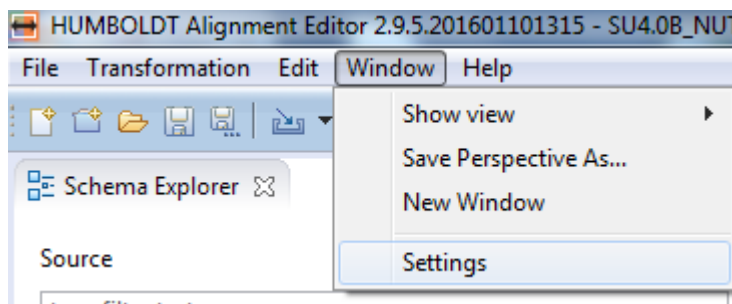
Report List

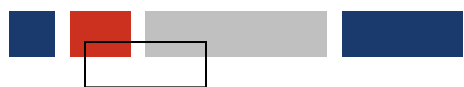
provides an overview of the last completed processes



Proxy Server

- A versão 2.9.4 de **Hale** parece não funcionar bem com proxy
- Fundamental para acesso a recursos na Web (**Codelists**, ...)
- Como definir *proxy*?





Hale - Workflow

Workflow
genérico para
transformar
CDG de
acordo com
os requisitos
do target
Schema

1.Importar Source/TargetSchemas

2.Importar dados

3.Definir mapping rules

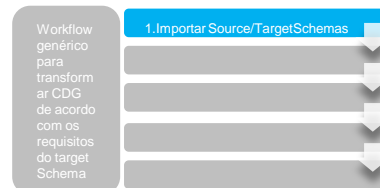
4.Exportar dados transformados

5.Validar dados transformados





Importar Source Schema



■ **Source Schema**

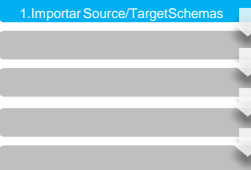
- Define a estrutura dos dados que desejamos transformar
- Várias fontes para dados, incluindo *online* (URL, WFS) , base de dados (PostgreSQL/PostGIS, SpatiaLite)

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



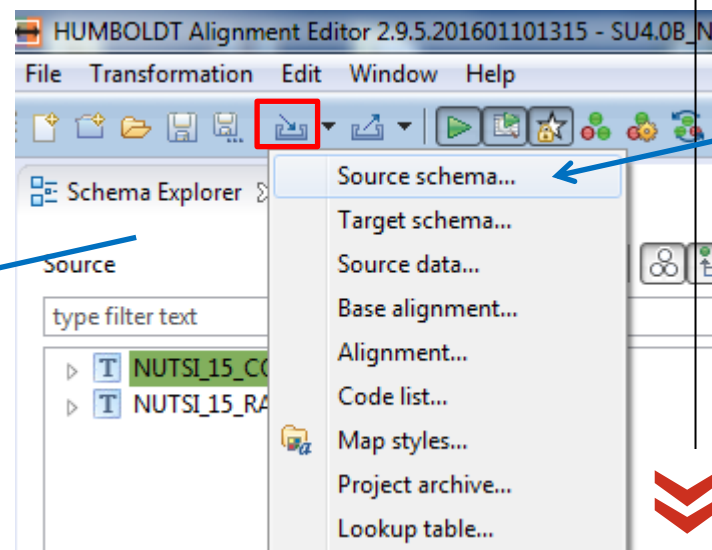
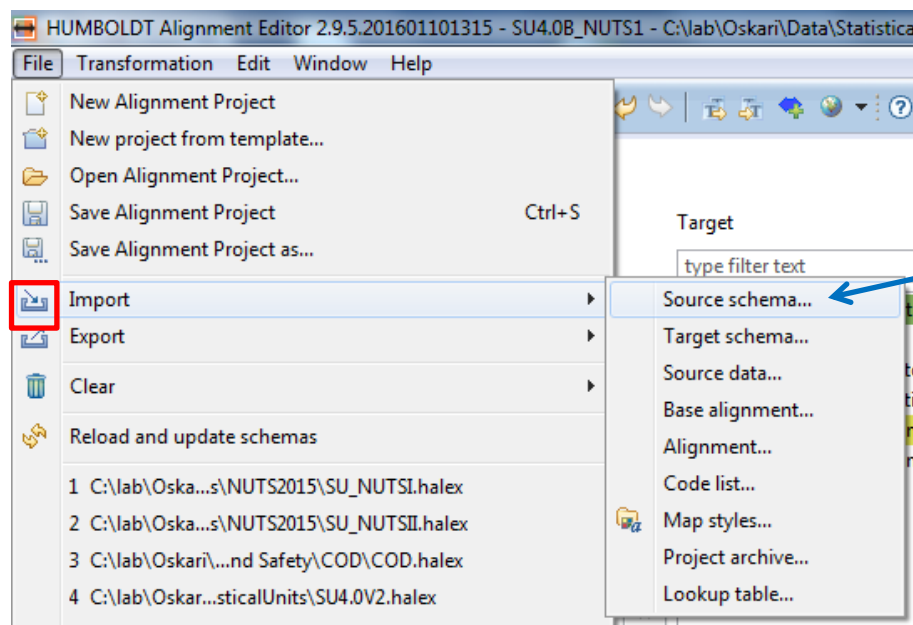
Importar Source Schema

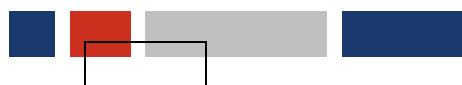
Workflow genérico para transformar CDG de acordo com os requisitos do target Schema



■ **Source Schema**

■ Várias formas de invocar a funcionalidade





Importar Source Schema

Workflow genérico para transformar o CDG de acordo com os requisitos do target Schema

1. Importar Source/Target Schemas

■ importar Source Schema



■ Cuidado com os caracteres *PT*

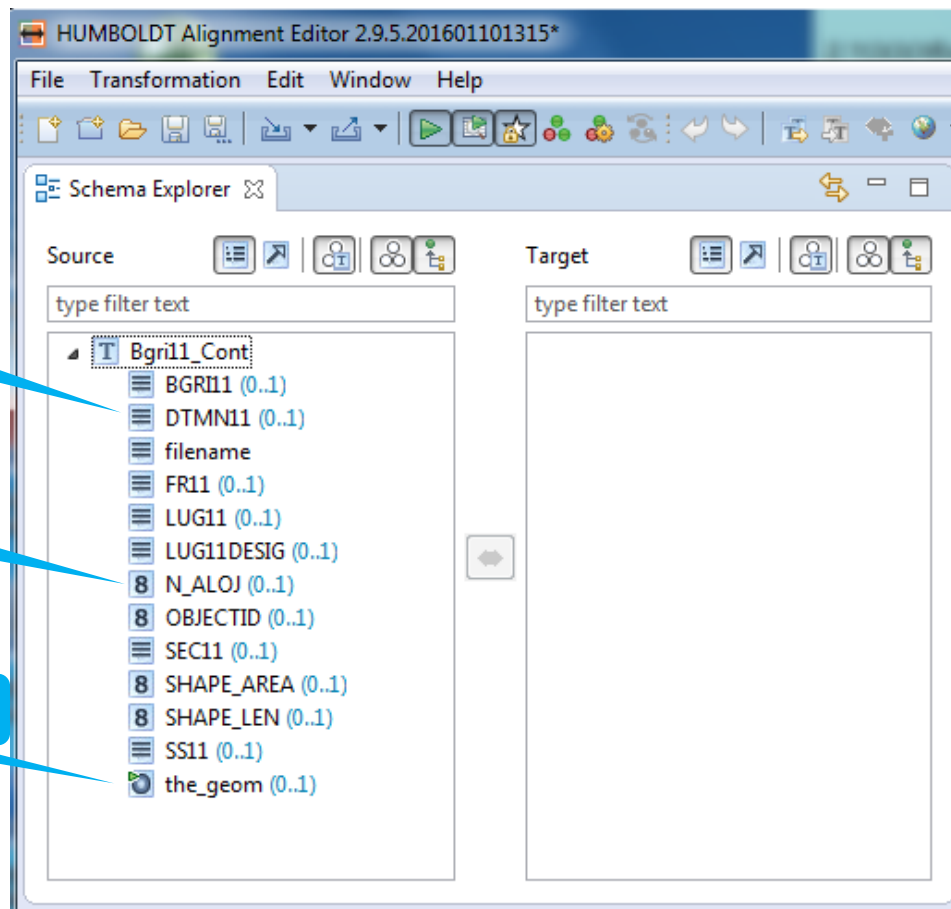


Importar Source Schema

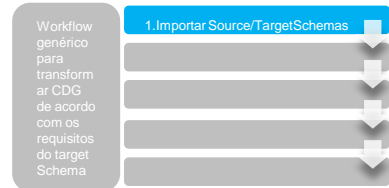
Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

1. Importar Source/Target Schemas

■ importar *Source Schema* - Simbologia

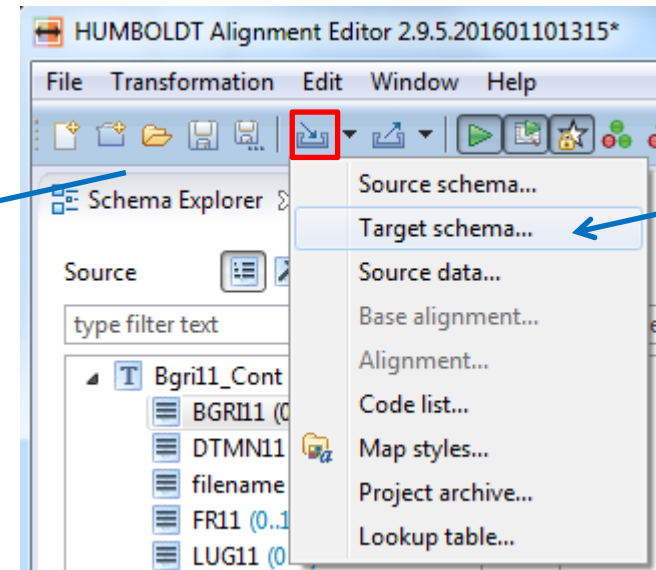
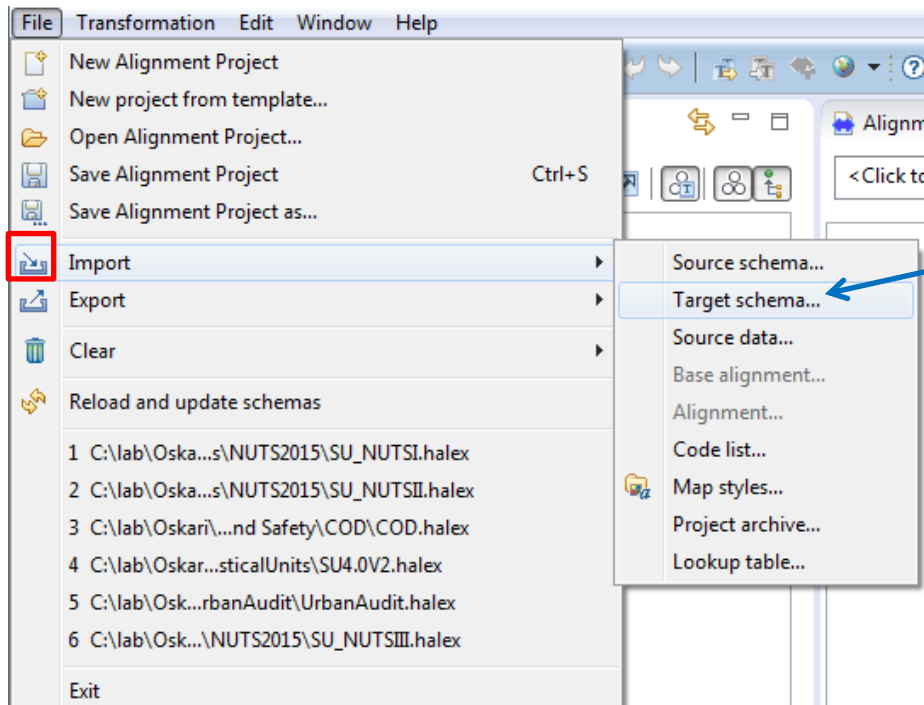


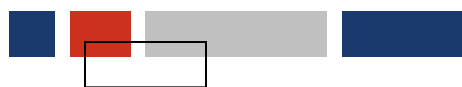
Importar target Schema



■ **Target Schema**

- Define a estrutura para a qual desejamos transformar os dados
- Schemas definidos nas **Data Specifications Inspire**





Importar Target Schema

Workflow genérico para transformar o CDG de acordo com os requisitos do target Schema

1. Importar Source/TargetSchemas

■ Target Schema

- Várias fontes, semelhante, ao **Source Schema** (File, Url, From Preset, WFS, BD)

Import target schema

Import location

Please select a source for the import

From file From URL From preset From WFS

Select preset:

Import as

< Back Next >

Select a schema

Statistical

INSPIRE Statistical Units Base 3.0
INSPIRE Statistical Units Base 4.0
INSPIRE Statistical Units Grid 3.0
INSPIRE Statistical Units Grid 4.0
INSPIRE Statistical Units Vector 3.0
INSPIRE Statistical Units Vector 4.0

Tip!

cial Governmental Services 3.0
cial Governmental Services 4.0
0
0 (Bundled)
0
culture Facilities Model 3.0
culture Facilities Model 4.0
3.0
3.0 (Bundled)
4.0
striction and Regulation Zones

Cancel None

OK Cancel None



Importar target Schema

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

1. Importar Source/Target Schemas

■ Target Schema

HUMBOLDT Alignment Editor 2.9.5.201601101315*

File Transformation Edit Window Help

Schema Explorer

Source

type filter text

- Bgri11_Cont
 - BGR11 (0..1)
 - DTMN11 (0..1)
 - filename
 - FR11 (0..1)
 - LUG11 (0..1)
 - LUG11DESIG (0..1)
 - N_ALOJ (0..1)
 - OBJECTID (0..1)
 - SEC11 (0..1)
 - SHAPE_AREA (0..1)
 - SHAPE_LEN (0..1)
 - SS11 (0..1)
 - the_geom (0..1)

Target

type filter text

- AreaStatisticalUnit
- Evolution
- GeometryDescriptor
- StatisticalTessellation
- VectorStatisticalUnit
- VectorStatisticalUnitGeometry

Alignment

<Click to select: Select cell X <Click to select

Error Log Properties

Workspace Log

type filter text

Message

- XML schema import
 - [14:33:19] Creating types for schema at http://inspire.ec.europa.
 - [14:33:19] Creating types for schema at http://www.w3.org/2001
 - [14:33:19] Creating types for schema at bundleresource://130.fw
 - [14:33:19] Loading includes and imports for schema at http://w
 - [14:33:19] White space facet not supported

Type... f_x Func... Repo... Tran... Mapp...

- 11:55 2016-04-07
 - ✓ XML schema import 14:32.48
 - ✓ Shapefile import 11:55.24
- 17:05 2016-04-04
- 10:26 2016-04-04
- 10:19 2016-04-04

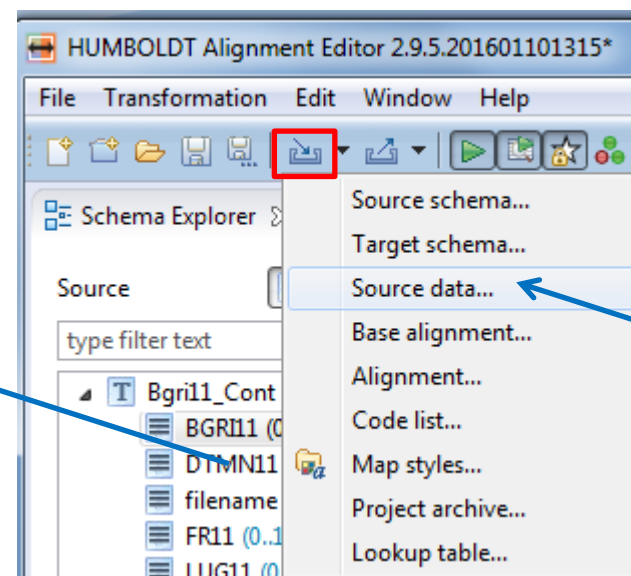
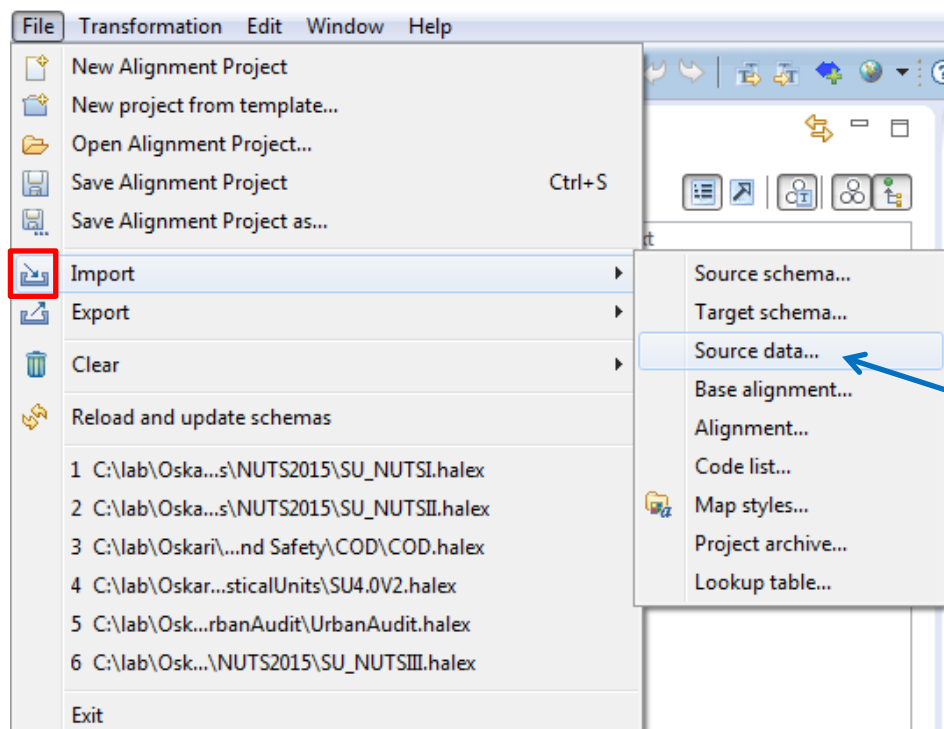


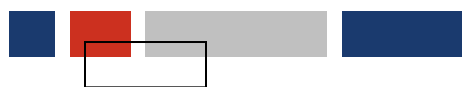
Importar Source data



■ Importar Source Data

- Importar o **source Schema** não importa as *features* (polígonos, linhas ou pontos).

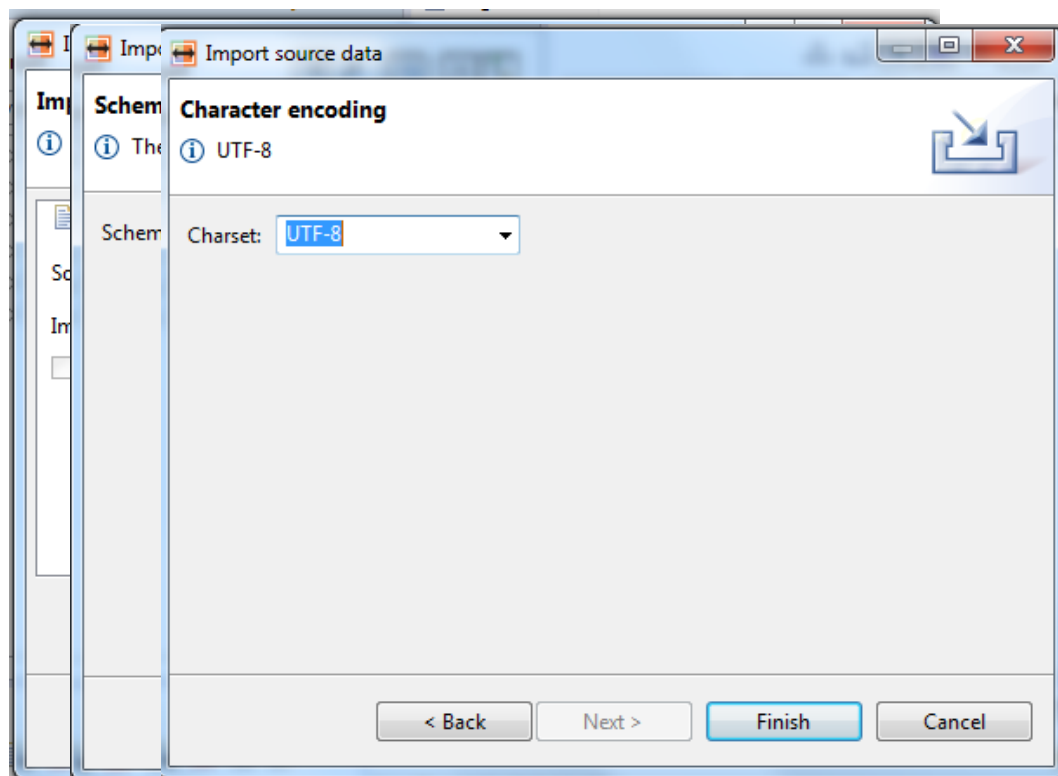




Importar Source data



■ Importar *Source Data*



■ Cuidado com os caracteres **PT**

Tip!





Importar Source data

■ Importar Source Data

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

2.Importar dados

HUMBOLDT Alignment Editor 2.9.5.201601101315*

File Transformation Edit Window Help

Schema Explorer

Source

type filter text

- Bgri11_AC26 ×4470
 - AREA (0..1) ×4470
 - BGRI11 (0..1) ×4470
 - DTMN11 (0..1) ×4470
 - filename ×4470
 - FR11 (0..1) ×4470
 - LEN (0..1) ×4470
 - LUG11 (0..1) ×4470
 - LUG11DESIG (0..1) ×4470
 - N_ALOJ (0..1) ×4470
 - OBJECTID (0..1) ×4470
 - SEC11 (0..1) ×4470
 - SHAPE_AREA (0..1) ×4470
 - SHAPE_LEN (0..1) ×4470
 - SS11 (0..1) ×4470
 - the_geom (0..1) ×4470

Target

type filter text

- AreaStatisticalUnit
- Evolution
- GeometryDescriptor
- StatisticalTessellation
- VectorStatisticalUnit
- VectorStatisticalUnitGeometry

Alignment

<Click to select: Select cell X <Click to select

Error Log

Properties

Workspace Log

type filter text

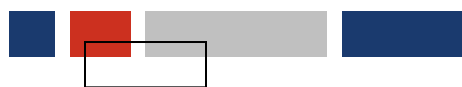
Message

- Extracting Type Definition for Bgri11_AC26
- Extracting Type Definitions.
- Starting task 'Analysing shapefile.'...
- Shapefile import
 - [14:57:48] Created cached schema representation

Type... f_x Func... Repo... Tran... Mapp...

- 11:55 2016-04-07
 - Load data into database 15:11.58
 - Shapefile import 15:11.58
 - Shapefile import 14:57.47
 - XML schema import 14:32.48
 - Shapefile import 11:55.24
- 17:05 2016-04-04
- 10:26 2016-04-04
- 10:19 2016-04-04





Importar Source data

- Data perspective

Workflow genérico para transformar o CDG de acordo com os requisitos do target Schema

2.Importar dados

601101315*

Help

Select cell <Click to select>

Source Data

Bgri11_AC26

Bgri11_AC26	1	2	3
Bgri11_AC26	+	+	+
8 AREA	0.0	0.0	0.0
BGRI11	42051400101	42030400205	42030400403
DTMN11	4205	4203	4203
filename	Bgri11_AC26	Bgri11_AC26	Bgri11_AC26
FR11	14	04	04
8 LEN	0.0	0.0	0.0
LUG11	027246	027185	027185
LUG11DESIG	São Brás	Capelas	Capelas
8 N_ALOJ	16	0	78
8 OBJECTID	36174	36175	36176
SEC11	001	002	004
8 SHAPE_AREA	310286.019904	7039.7060865	665119.270372
8 SHAPE_LEN	3291.09965573	350.115823455	3838.44559463
SS11	01	05	03
the_geom	{CRS=ITRF93_UTM_Zone	{CRS=ITRF93_UTM_Zone	{CRS=ITRF93_UTM_Zone
Metadata	+	+	+
Identifier	9b06e996-ee7c-467b-9ef	124b5166-0ccb-401d-af9	a56a3bd3-e2e8-444d-bb6

Mapping

11.58
5:11.58
4:57.47



Importar Source data

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

2.Importar dados

■ *Map perspective*

HUMBOLDT Alignment Editor 2.9.5.201601101315 - Formacao - C:\lab\Formacao\formacao.halex

File Transformation Edit Map Window Help

Map

Data, imagery and map information provided by MapQuest, Open Street Map and contributors, CC-BY-SA

Transformed data

ER 1-1A

via rapida

Nordeste

Pico da Vara

Pedreira

Pico Barilomeu

Agua Retorta

EN 1-1A

Povoação

Faial da Terra

Source data

Source Data

Bgri11_AC26

Bgri11_AC26

Bgri11_AC26	
Bgri11_AC26	1
AREA	0.0
BGR11	42050300211
DTMN11	4205
filename	Bgri11_AC26
FR11	03

Transformed Data

Report List

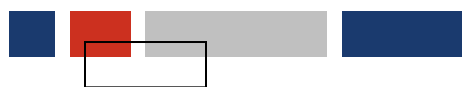
Mapping

- 15:57 2016-04-07
 - ✓ Load data into database 15:57.26
 - ✓ Shapefile import 15:57.26
 - ✓ Shapefile import 15:57.24
 - ✓ XML schema import 15:57.23
 - ✓ HALE XML project import 15:57.14
- 11:55 2016-04-07
- 16:39 2016-04-05
- 17:05 2016-04-04

11:58

15:11.58

14:57.47



Importar Source data



■ Importar *Source Data*

- Importar cdg's com um número elevado de *features* pode trazer problemas de performance porque o **HALE** transforma os dados e valida, por defeito, sempre que existe mapeamento de algum campo.

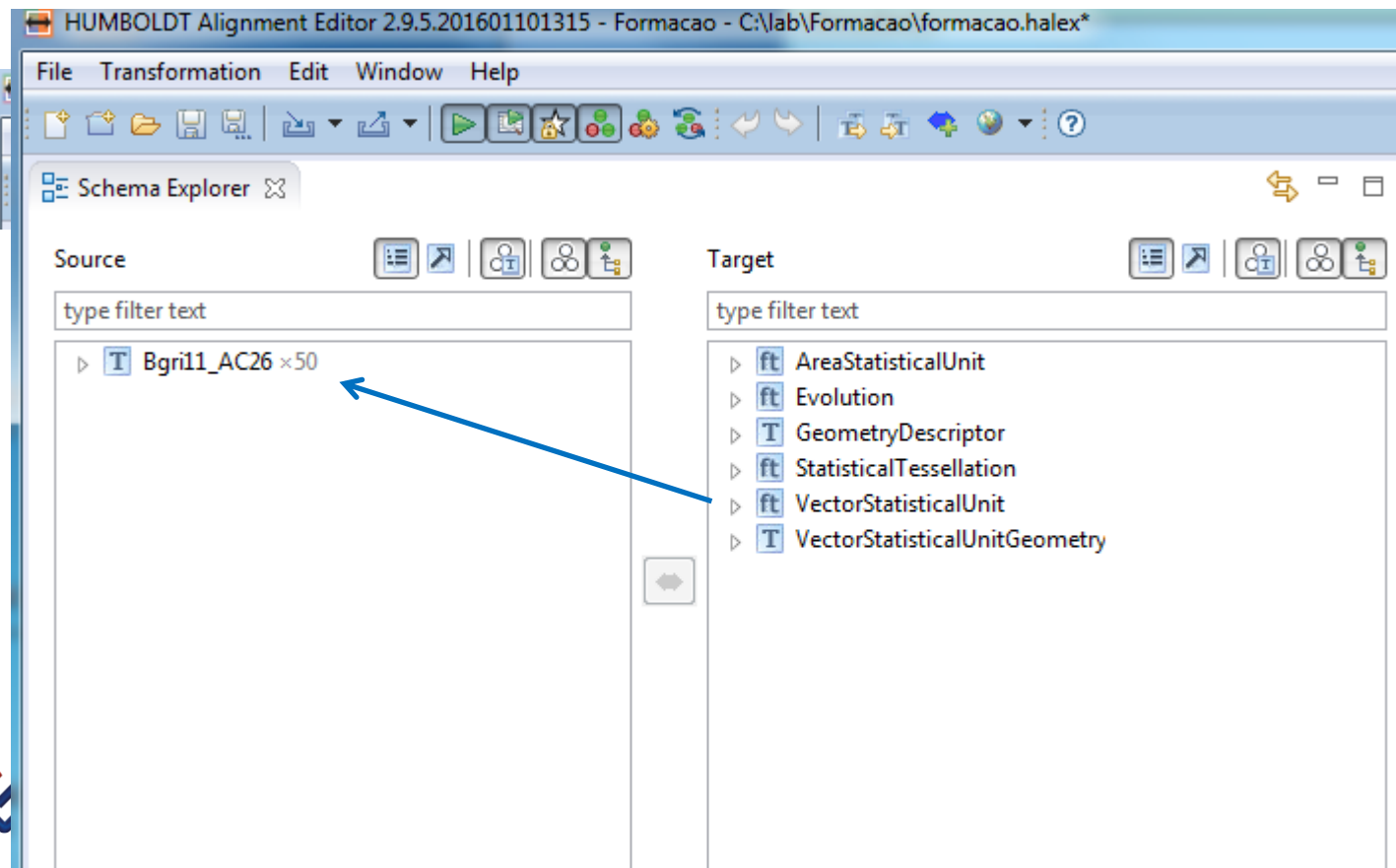
Importar Source data

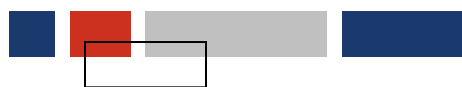
Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

2.Importar dados

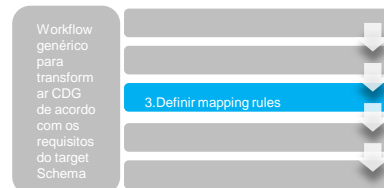
- É possível definir uma seleção de *features* para trabalhar com menos registos.

Tip!

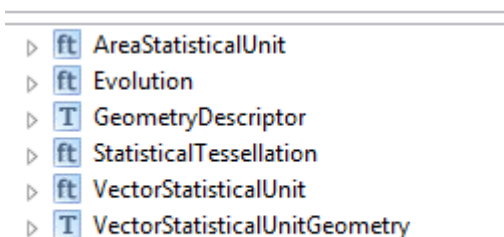




Definir Mapping rules



- No **Target Schema** é necessário identificar **tipo** relevante para o cdg que pretendemos harmonizar.



- As data specifications muitas vezes indicam claramente o tipo



- The application schema on statistical units is composed of different packages:

- Base: The base package.
- Grid: Classes to represent statistical grids.
- Vector: Classes to represent statistical unit having vector geometries (point, line, surfaces).



Definir Mapping rules

Workflow genérico para transformar o CDG de acordo com os requisitos do target

3. Definir mapping rules

HUMBOLDT Alignment Editor 2.9.5.201601101315 - Formacao - C:\lab\Formacao\formacao.halex*

File Transformation Edit Window Help

Schema Explorer

Source

type filter text

- Bgri11_AC26 ×50
 - AREA (0..1) ×50
 - BGRI11 (0..1) ×50
 - DTMN11 (0..1) ×50
 - filename ×50
 - FR11 (0..1) ×50
 - LEN (0..1) ×50
 - LUG11 (0..1) ×50
 - LUG11DESIG (0..1) ×50
 - N_ALOJ (0..1) ×50
 - OBJECTID (0..1) ×50
 - SEC11 (0..1) ×50
 - SHAPE_AREA (0..1) ×50
 - SHAPE_LEN (0..1) ×50
 - SS11 (0..1) ×50
 - the_geom (0..1) ×50

Target

type filter text

- ft AreaStatisticalUnit
- ft Evolution
- ft GeometryDescriptor
- ft StatisticalTessellation
- ft VectorStatisticalUnit ×50
- ft VectorStatisticalUnitGeometry

Alignment

<Click to select> Select cell <Click to select>

T Bgri11_AC26 → T Retype → ft VectorStatisticalUnit

Para cada instância na data source vai criar uma instância no tipo identificado do *target* Schema

Error Log

Properties

Instance validation

Report

Warnings

- ft VectorStatisticalUnit (250 warnings)
 - country (50 warnings)
 - Cardinality (showing 5 of 50)
 - Not enough values for the property present: 0 < 1
 - Not enough values for the property present: 0 < 1
 - Not enough values for the property present: 0 < 1
 - Not enough values for the property present: 0 < 1
 - Not enough values for the property present: 0 < 1
 - geometry (50 warnings)
 - Cardinality (showing 5 of 50)

Instance validation

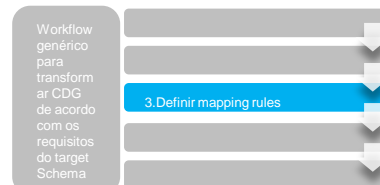
- Instance transformation 10:10.10
- Load data into database 10:10.09
- Shapefile import 16:07.03
- Load data into database 16:07.03
- Shapefile import 16:02.13
- Load data into database 16:02.13
- Shapefile import 16:01.11
- Load data into database 16:01.11
- Shapefile import 16:00.32
- Load data into database 16:00.32
- Shapefile import 15:57.26
- Load data into database

191M of 372M

CST



Definir Mapping rules



■ Schema elements

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

Target

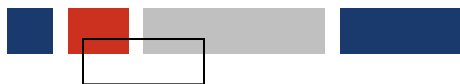
type filter text

- AreaStatisticalUnit
- Evolution
- GeometryDescriptor
- StatisticalTessellation
- VectorStatisticalUnit ×50
 - location (0..1)
 - beginLifespanVersion
 - boundedBy (0..1)
 - country
 - description (0..1)
 - descriptionReference (0..1)
 - endLifespanVersion (0..1)
 - evolutions (0..n)
 - geographicalName (0..n)
 - geometry (1..n)
 - id
 - identifier (0..1)
 - inspireId
 - metaDataProperty (0..n)
 - name (0..n)
 - referencePeriod
 - thematicId (0..n)
 - validityPeriod (0..1)

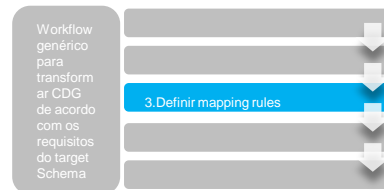
■ Propriedades obrigatórias são facilmente identificadas no error log quando não estão preenchidas

Tip!





Definir Mapping rules



■ Cardinalidade

- Para cada elemento existe informação sobre a sua cardinalidade
- O valor n ($0..n$) representa um número ilimitado máximo de ocorrências
- Se não existe informação sobre a cardinalidade significa que a propriedade ocorre apenas uma vez
- Quando a cardinalidade é ($0..1$) significa que podem ou não existir ocorrências, ou seja é um elemento opcional.

Target

type filter text

- ▶ ft AreaStatisticalUnit
- ▶ ft Evolution
- ▶ T GeometryDescriptor
- ▶ ft StatisticalTessellation
- ▶ ft VectorStatisticalUnit ×50
 - ▶ location (0..1)
 - ▶ beginLifespanVersion
 - ▶ boundedBy (0..1)
 - ▶ country
 - ▶ description (0..1)
 - ▶ descriptionReference (0..1)
 - ▶ endLifespanVersion (0..1)
 - ▶ evolutions (0..n) ←
 - ▶ geographicalName (0..n)
 - ▶ geometry (1..n)
 - ▶ id ←
 - ▶ identifier (0..1) ←
 - ▶ inspireId
 - ▶ metaDataProperty (0..n)
 - ▶ name (0..n)
 - ▶ referencePeriod
 - ▶ thematicId (0..n)
 - ▶ validityPeriod (0..1)

Definir Mapping rules

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

3. Definir mapping rules

HUMBOLDT Alignment Editor 2.9.5.201601101315 - Formacao - C:\lab\Formacao\formacao.halex*

File Transformation Edit Window Help

Schema Explorer

Source

type filter text

- Bgri11_AC26 ×50
 - AREA (0..1) ×50
 - BGRI11 (0..1) ×50
 - DTMN11 (0..1) ×50
 - filename ×50
 - FR11 (0..1) ×50
 - LEN (0..1) ×50
 - LUG11 (0..1) ×50
 - LUG11DESIG (0..1) ×50
 - N_ALOJ (0..1) ×50
 - OBJECTID (0..1) ×50
 - SEC11 (0..1) ×50
 - SHAPE_AREA (0..1) ×50
 - SHAPE_LEN (0..1) ×50
 - SS11 (0..1) ×50
 - the_geom (0..1) ×50

Target

type filter text

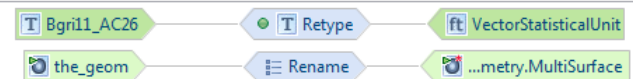
- endLifespanVersion (0..1)
- evolutions (0..n)
- geographicalName (0..n)
- geometry (1..n) ×50
 - VectorStatisticalUnitGeometry ×50
 - geometry ×50
 - AbstractGeometry (0..1) ×50
 - AbstractSolid
 - CompositeCurve
 - CompositeSolid
 - CompositeSurface
 - Curve
 - GeometricComplex
 - Grid
 - LineString
 - MultiCurve
 - MultiGeometry
 - MultiPoint
 - MultiSolid
 - MultiSurface ×50
 - OrientableCurve
 - OrientableSurface
 - Point
 - Polygon

Alignment

T Bgri11_AC26 ×50

Retype

ft VectorStatisticalUnit ×50



Error Log Properties

Workspace Log

type filter text

Message

Instance transformation

Type hierarchy Functions Report List Transformed Data Mapping

15:57 2016-04-07

- Instance validation 11:40:57
- Instance transformation 11:40:56
- Instance validation 11:34:05
- Instance transformation 11:34:04

Help

< Back

Next >

Finish

Cancel



INSTITUTO NACIONAL
DE ESTATÍSTICA
PORTUGAL



Definir Mapping rules

▪ Geometry

Tip!

▪ Porque *Multisurface* e não *polygon*?

Se na
existirem
apenas
harmoniza
ignorados



Table

NUTS1_15_AC26

FID	Shape *	OBJECTID	NUTS1_15	NUTS1_15DE	SHAPE AREA	SHAPE LEN
0	Polygon	1	PT2	Região Autónoma dos Açores	2163885709,67	849427,366641

(1 out of 1 Selected)

NUTS1_15_AC26



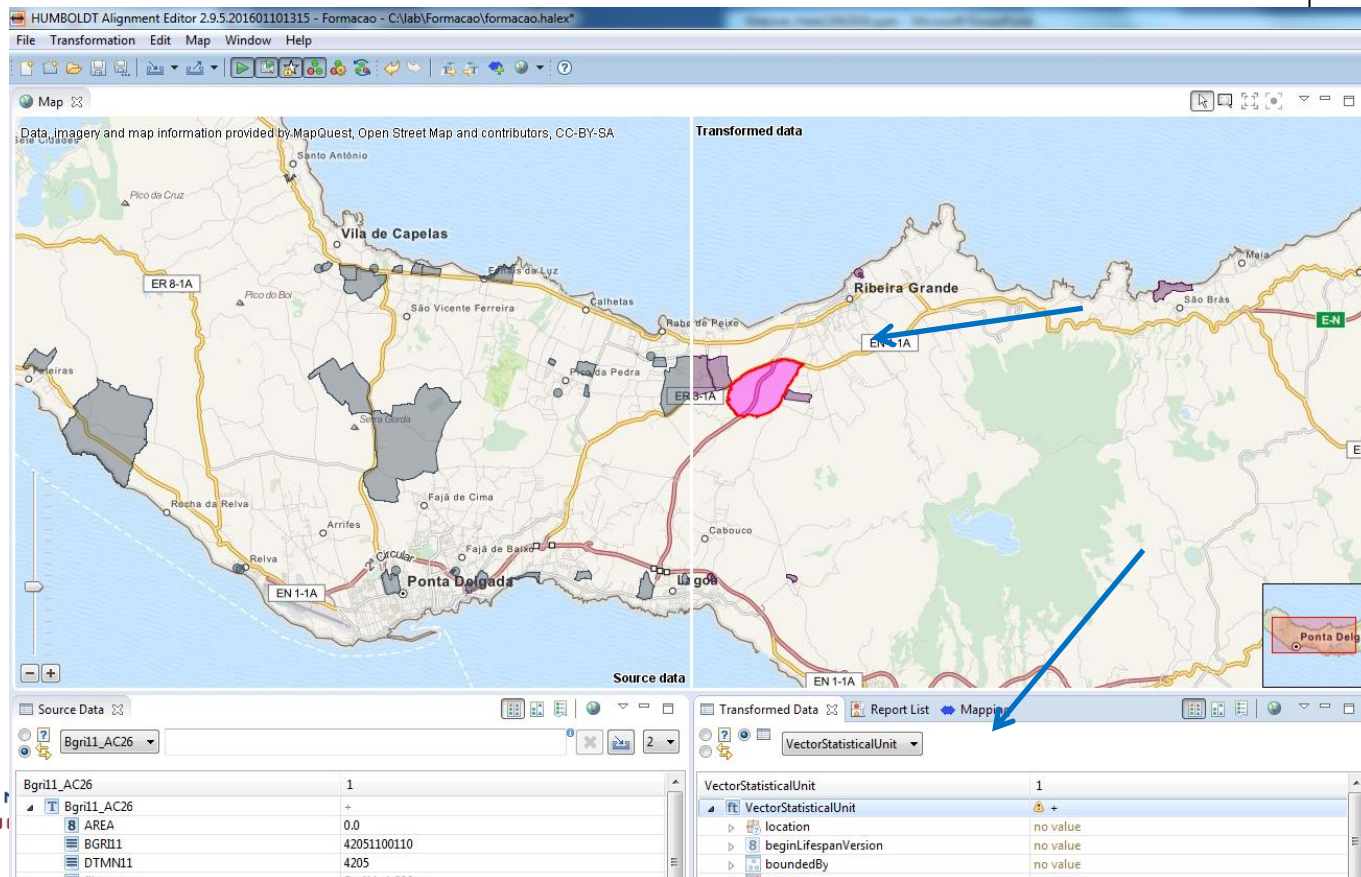
Definir Mapping rules

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

3. Definir mapping rules

■ Map View

- É possível ver a **transform data**, geometria e atributos



INSTITUTO NACIONAL
DE ESTATÍSTICA

Definir Mapping rules

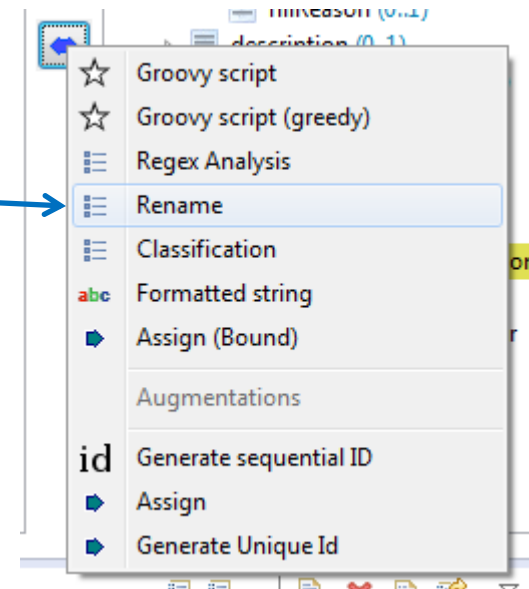
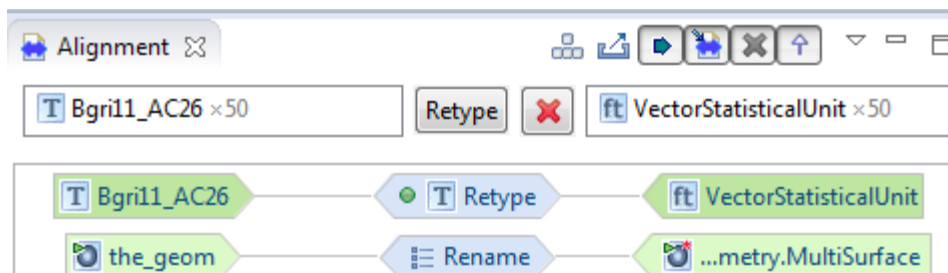
Workflow genérico para transformar o CDG de acordo com os requisitos do target Schema

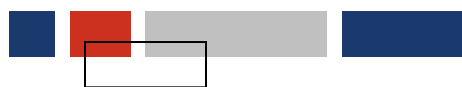
3. Definir mapping rules

Geoserver

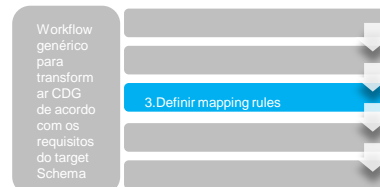
▪ Função *rename*

- Permite copiar uma propriedade qualquer do **Source Schema** para o **target Schema**
- Foi utilizada na geometria, mas pode ser utilizada em qualquer propriedade





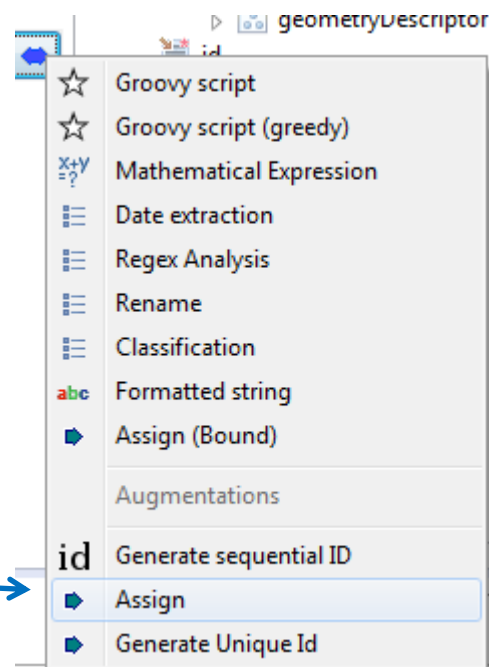
Definir Mapping rules

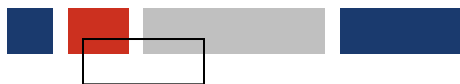


Geoserver

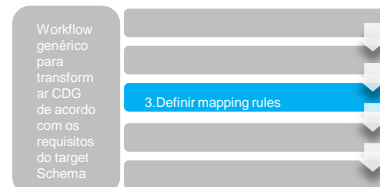
▪ Função *Assign*

- Permite atribuir um valor a uma propriedade do **target Schema**, que não exista no **Source Schema**





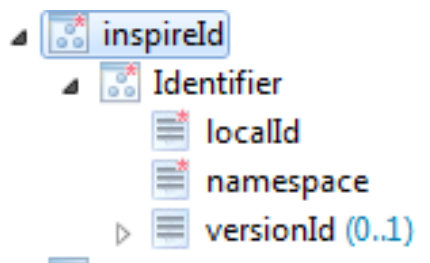
Definir Mapping rules



Geoserver

▪ 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.



Definir Mapping rules

▪ Função Assign

Geoserver

- O atributo **Namespace** não existe na **Source data** deve ter o valor <http://id.igeo.pt/so/SU/VectorStatisticalUnit>

▪ Ona

Source

type filter text

ht

são

▪ In

▪ Es

▪ Si

▪ Ol

▪ lo

orig

▪ ve

http

http

IN

S-

th

8

8

8

8

8

8

8

8

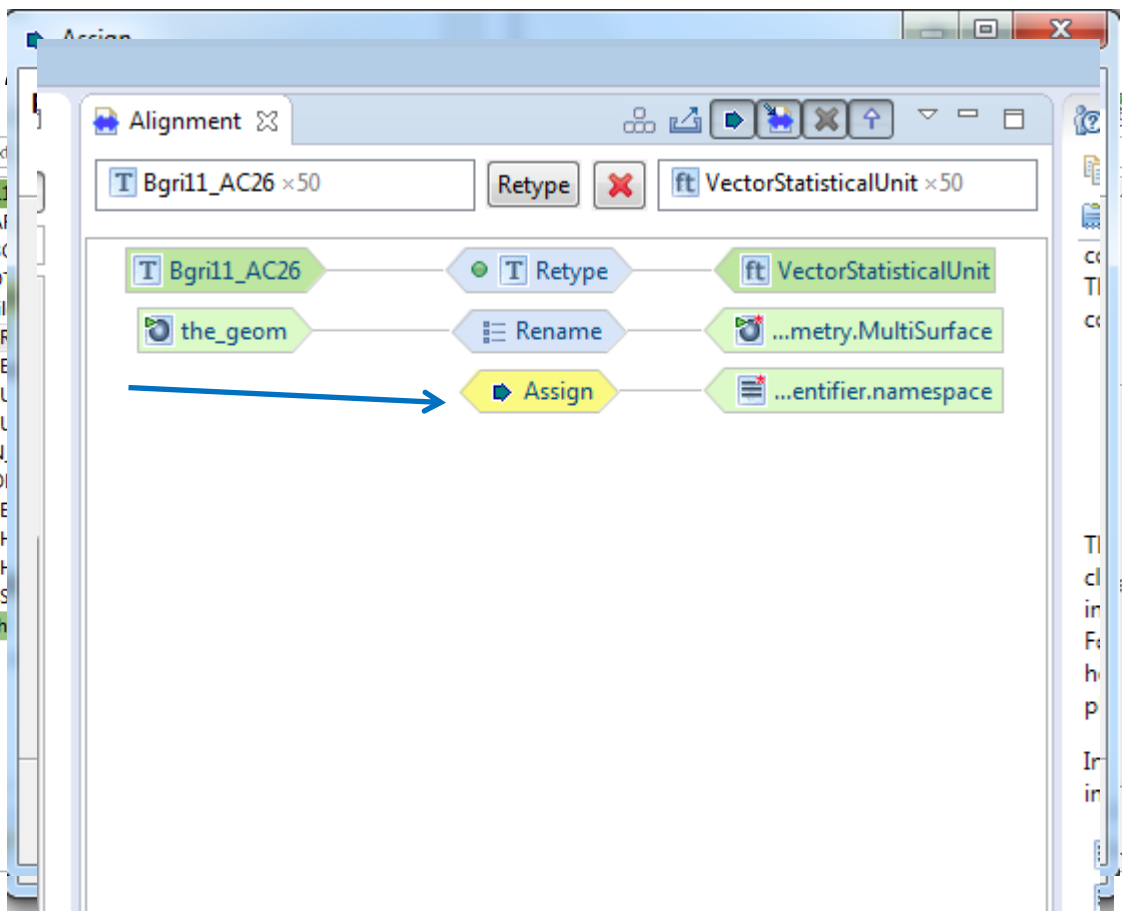
8

8

8

8

8



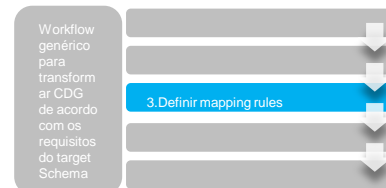
alld}/{ver

Inspire}

geográfico



Definir Mapping rules



▪ InspireId

▪ Os restantes campos:

▪ ***LocalId rename*** de um campo com código único ou designação do recurso original + “_” + código do objecto geográfico

▪ -- Definition --

▪ A local identifier, assigned by the data provider. The local identifier is unique within the namespace, that is no other spatial object carries the same unique identifier.

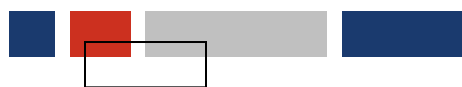
▪ -- Description --

▪ NOTE It is the responsibility of the data provider to guarantee uniqueness of the local identifier within the namespace.

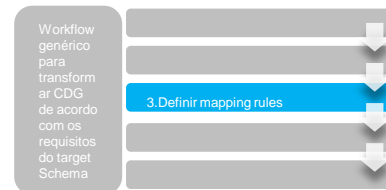
▪ ***VersionId assign*** de um valor da versão

▪ -- Definition --

▪ The identifier of the particular version of the spatial object, with a maximum length of **25 characters**. If the specification of a spatial object type with an external object identifier includes life-cycle information, the version identifier is used to distinguish between the different versions of a spatial object. Within the set of all versions of a spatial object, the version identifier is unique.



Definir Mapping rules

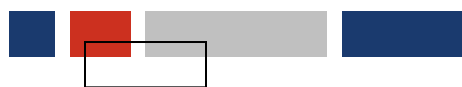


■ **Codelists**

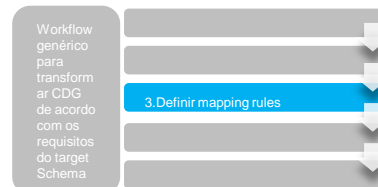


- São listas de códigos e seus valores, conforme definido nas regras de implementação **INSPIRE** relativas à interoperabilidade dos conjuntos e serviços de dados espaciais (Regulamento (UE) no 1089/2010 da Comissão).





Definir Mapping rules

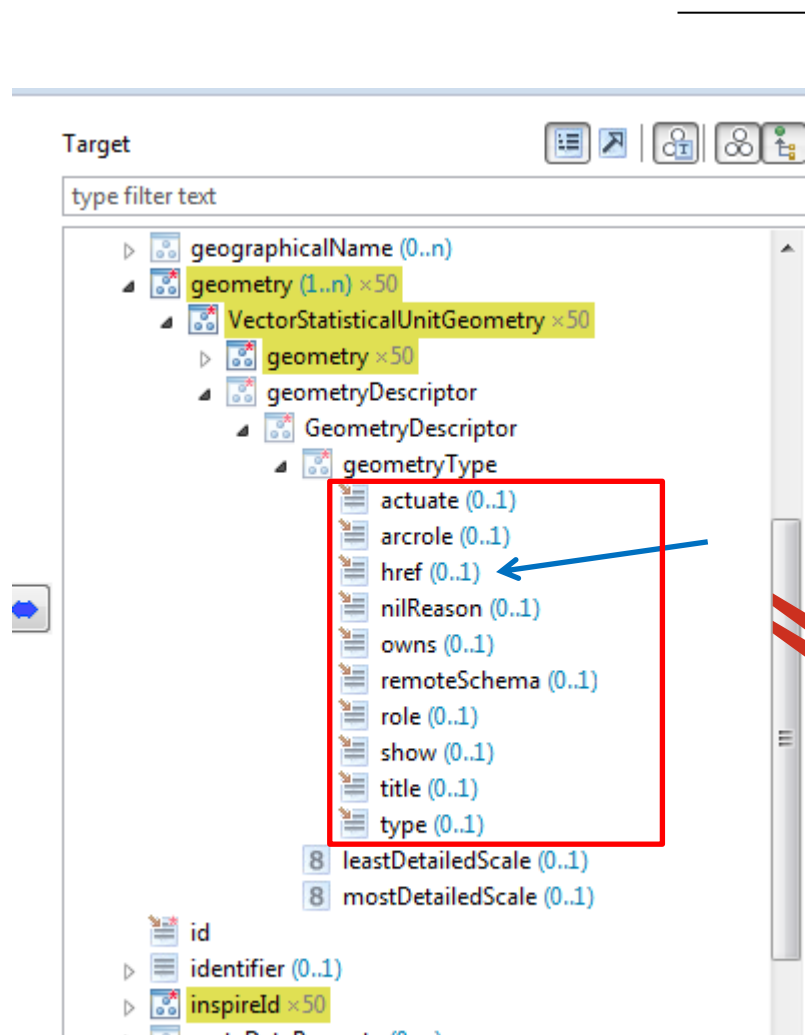


■ **Codelists**



■ No **target Schema** esta estrutura de elementos está associada a **Codelist**

■ O campo **Href** irá receber os valores

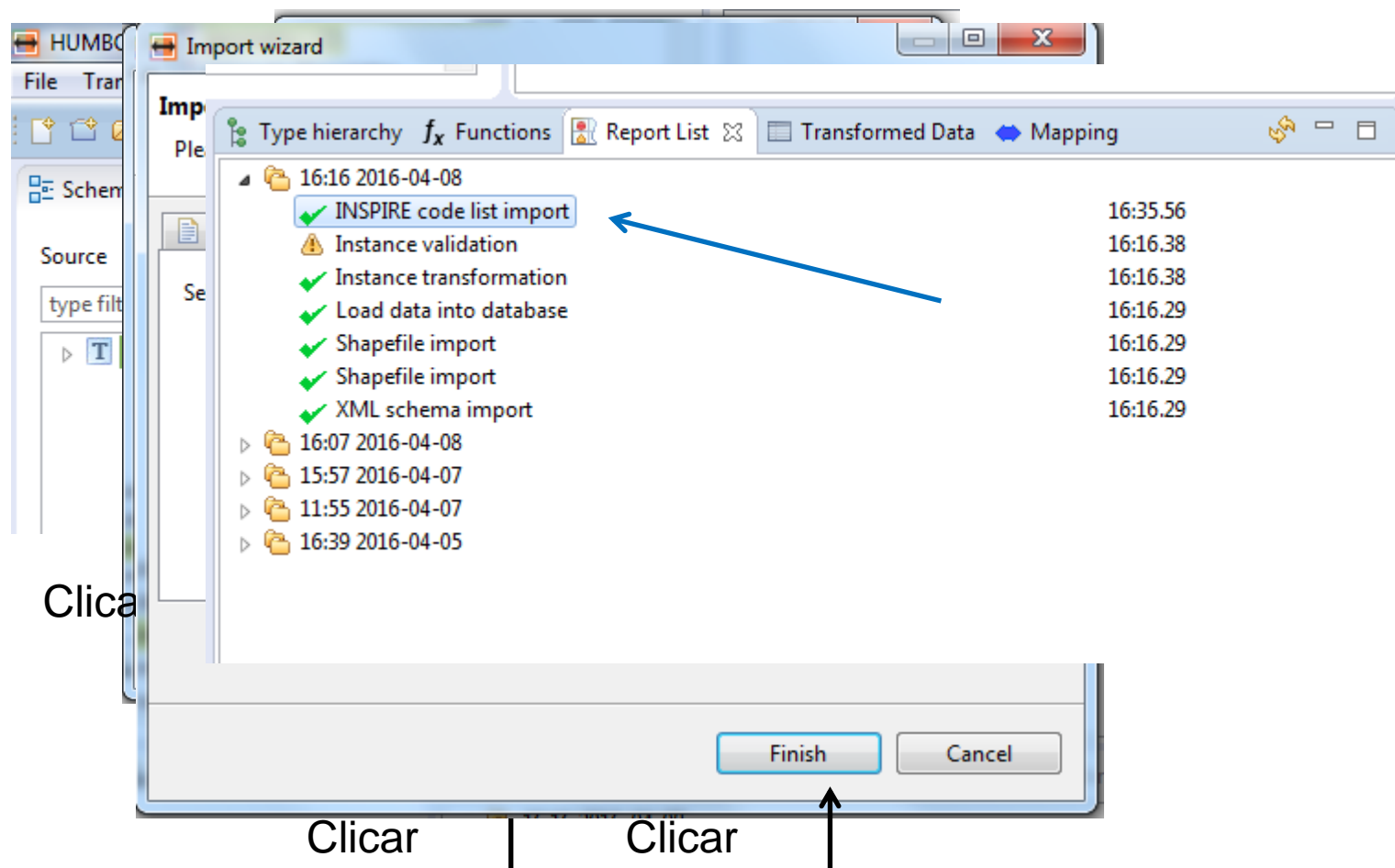


Definir Mapping rules

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

3. Definir mapping rules

■ Como importar **Codelists**

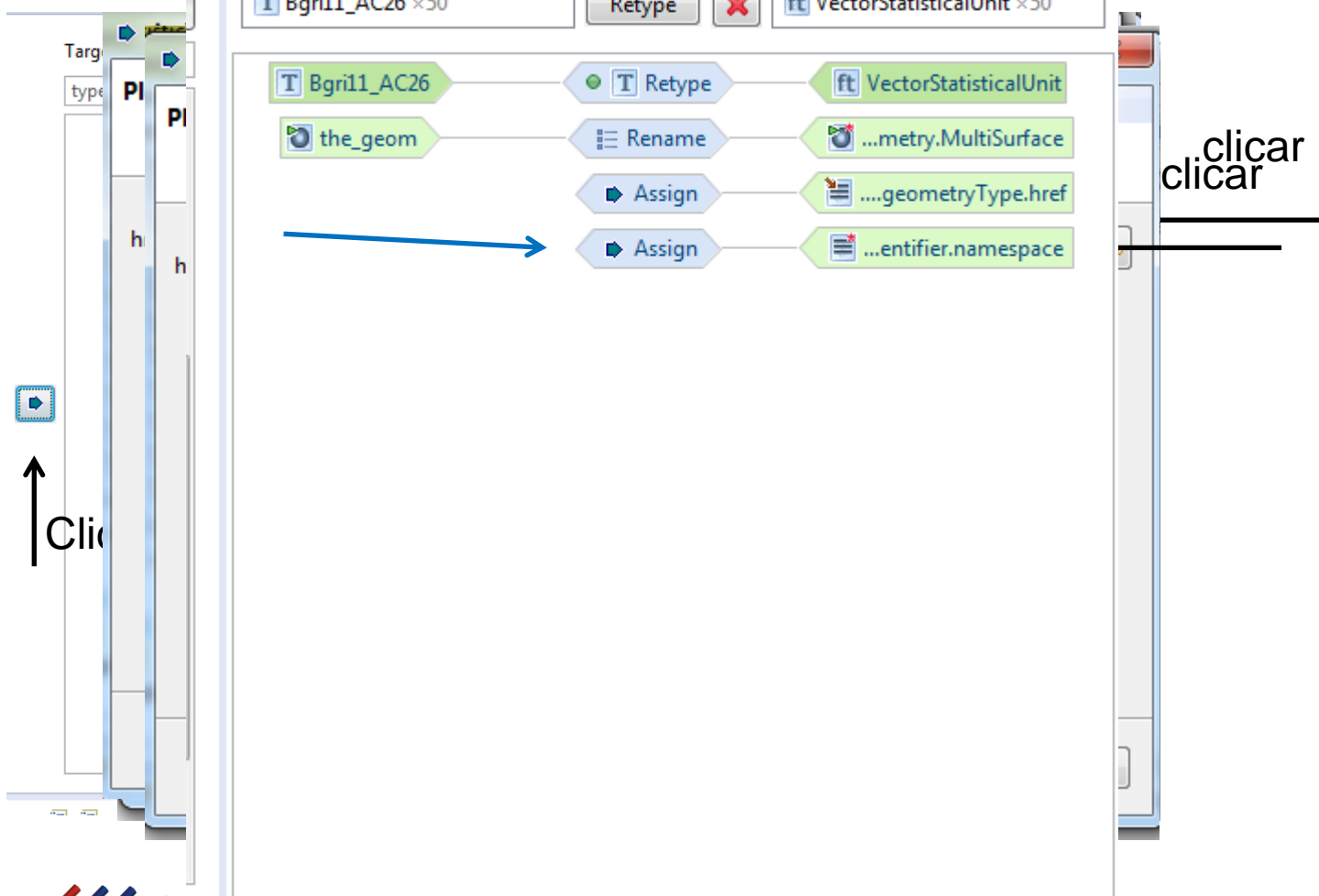


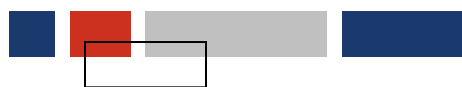
Definir Mapping rules

Workflow genérico para transformar o CDG de acordo com os requisitos do target Schema

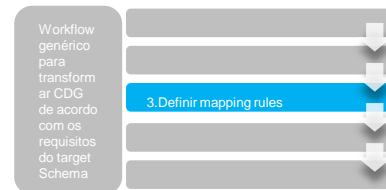
3. Definir mapping rules

Como





Definir Mapping rules

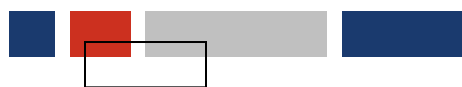


▪ ***Codelists Externas***

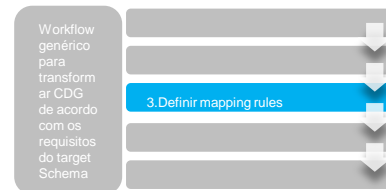
▪ Existem codelist mantidas por organizações externas ao Inspire

▪ Causas de morte Eurostat

▪ Patologia World Health Organization WHO



Definir Mapping rules



▪ ***Codelist ainda sem valores***

▪ Existem algumas *codelist* sem códigos definidos, por vezes completamente vazias

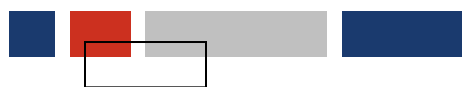
▪ *INSPIRE Registry*:

▪ Padrão em estudo

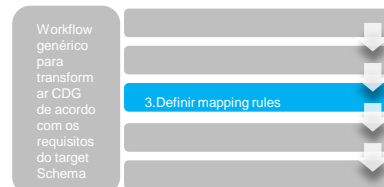
▪ 1. <http://registo.igeo.pt/codelist/{NomeListaCodigos}Value>

▪ 2. <http://registo.igeo.pt/codelist/{NomeListaCodigos}Value/{ValorCodigo}>

▪ <http://registo.igeo.pt/codelist/AggregationUnitValue/NUTSIII>



Definir Mapping rules



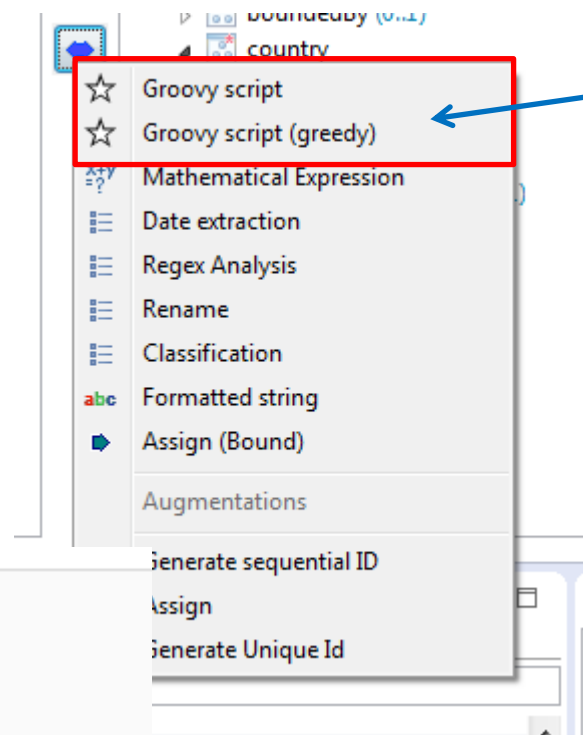
Geoserver



■ Outras funções

■ **Groovy Script**

- Permite o cálculo através de código de um valor de uma propriedade no **target schema**

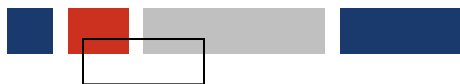


```
// retrieve all values of "result" in the merged object
def list = _source.p.result.values()

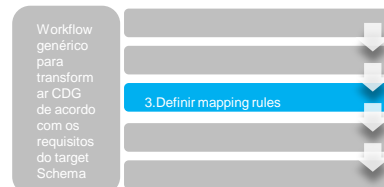
// concatenate the values, separated by comma
def str = list.join(',')

// assign the string to the property "resultString" in the target object
_target {
  resultString(str)
}
```





Definir Mapping rules



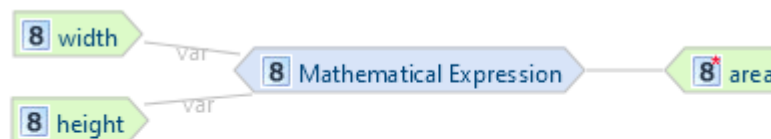
Geoserver



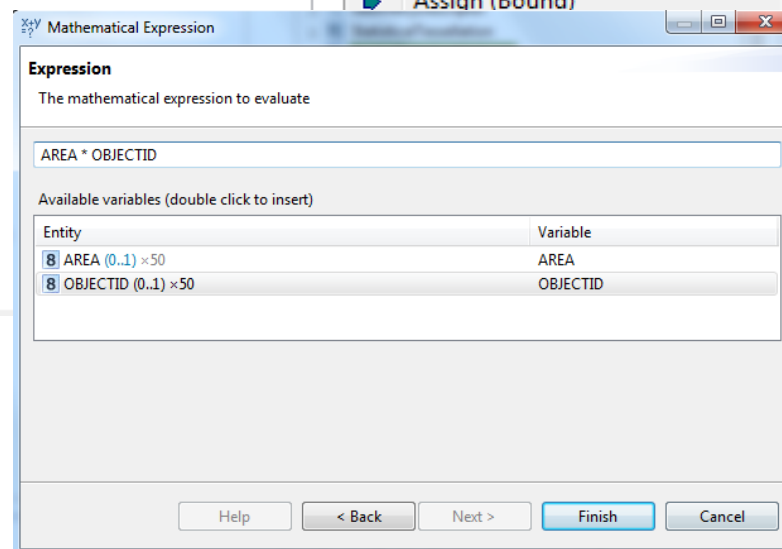
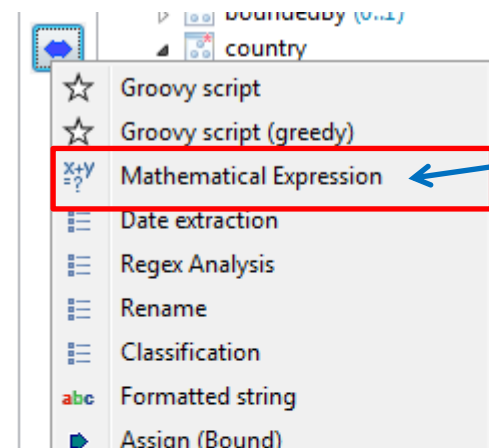
■ Outras funções

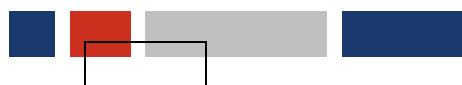
■ *Mathematical Expression*

- Define um valor através de uma expressão matemática
- Pode utilizar variáveis

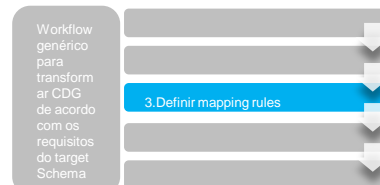


Parameter	Value
expression	width * height





Definir Mapping rules



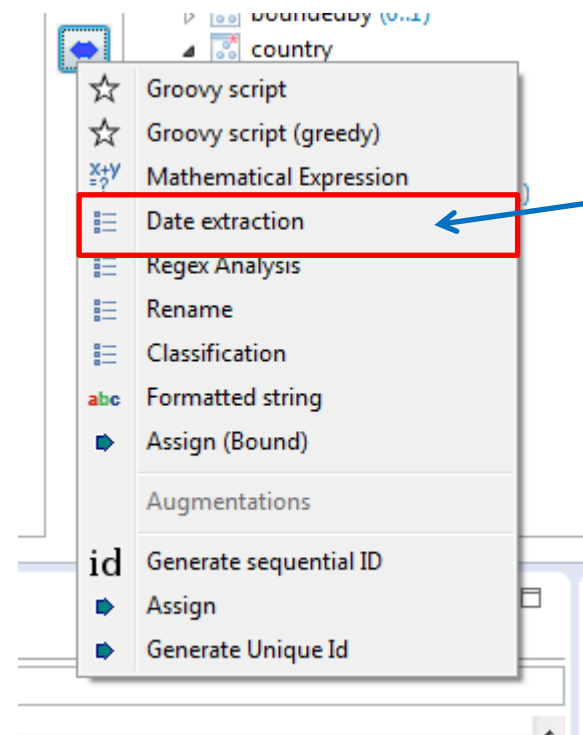
Geoserver

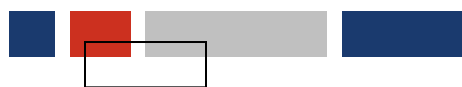


■ Outras funções

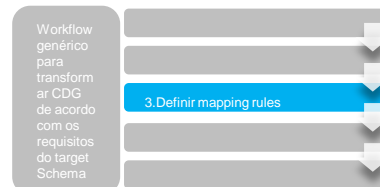
■ ***Date Extraction***

- Extrai uma data de uma string
- **yyyy-MM-dd HH:mm:ss**





Definir Mapping rules



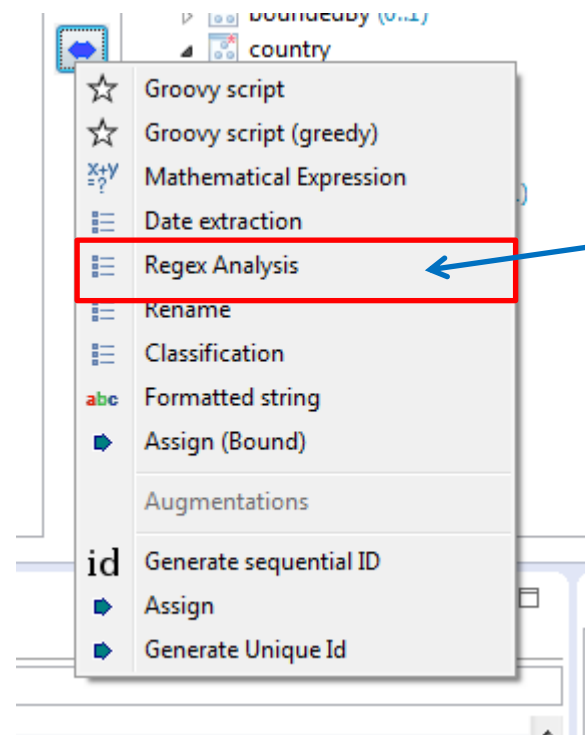
Geoserver

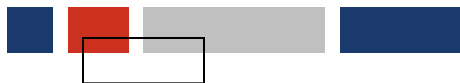


■ Outras funções

■ *Regex Analysis*

- Aplica uma *regular expression* a uma *string*





Definir Mapping rules

Workflow genérico para transformar o CDG de acordo com os requisitos do target Schema

3. Definir mapping rules

Let us assume that the source property represents a date of the format:

20081209

and that the target property needs a date in the format :

YYYY-MM-DD HH:MM:SS

We can use a regular expression like:

([0-9]{4})([0-9]{2})([0-9]{2})

and an output format:

{1}-{2}-{3} 00:00:00

to achieve our goal.

In fact the result of the above example will result in:

2008-12-09 00:00:00

Explained

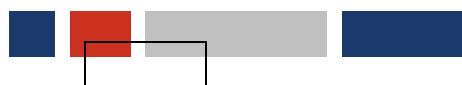
The regular expression used defines 3 groups, which are separated by the round brackets:

Regex Group	Explained	Catched part of 20081209
([0-9]{4})	catches 4 numbers between 0 and 9	2008
([0-9]{2})	catches 2 numbers between 0 and 9	12
([0-9]{2})	catches 2 numbers between 0 and 9	09

The output can then be formatted concatenating groups with any string.

The format **{1}-{2}-{3} 00:00:00** simply defines that the first three groups will be concatenated by a dash and that a default time will be added at the end of the resulting string.





Definir Mapping rules



Geoserver

- Outras funções
 - Classification
 - Reclassifica valores

Classification

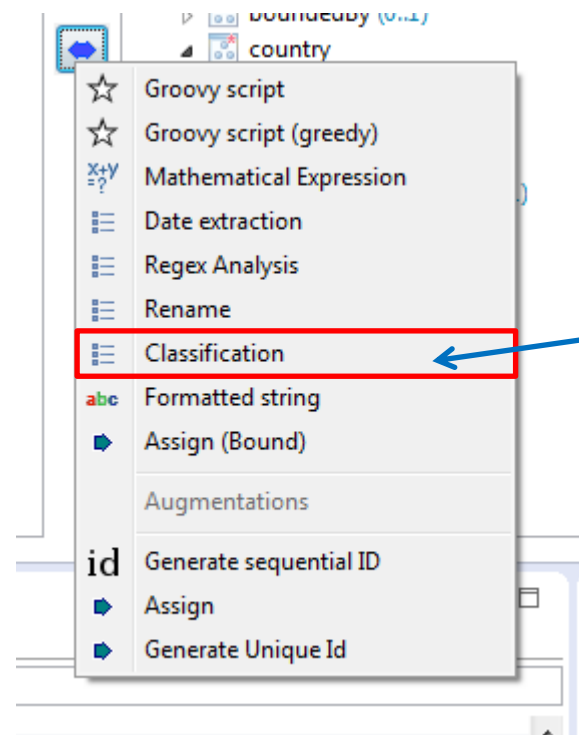
For unmapped source values: assign null

Explicit From file

Source value	Target value
4201	Lagoa
4202	Nordeste
4203	Pontal Delgada
4204	(unmapped)
4205	(unmapped)

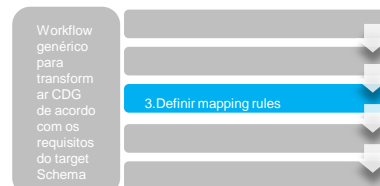
Double click on a table cell to change its value.

Help < Back Next > Finish Cancel





Definir Mapping rules



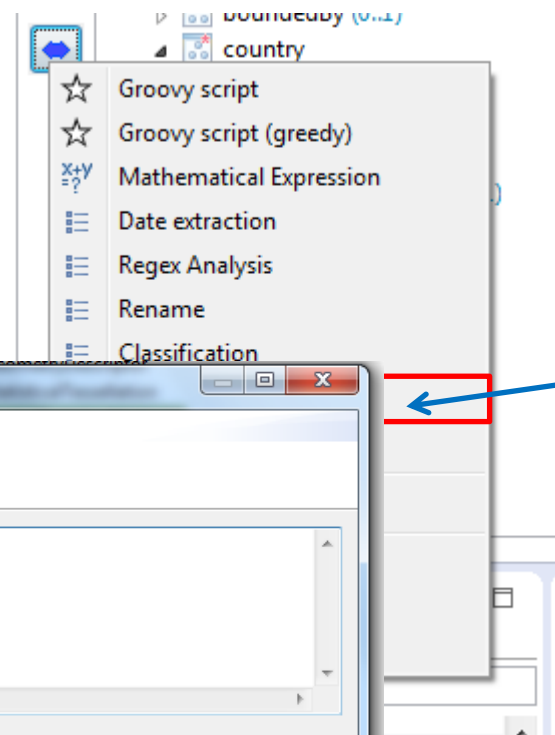
■ Outras funções

Geoserver



■ *Formatted String*

- Semelhante ao Mathematical Expression mas com Strings
- Cria uma string através de um padrão e variáveis



abc Formatted string

Format pattern
Enter a pattern

Código de Município : {DTMN11}

Available variables (double click to insert)

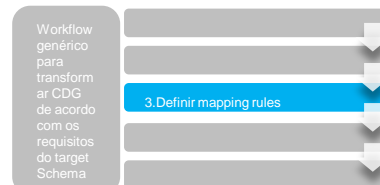
Entity	Variable
DTMN11 (0..1) x50	{DTMN11}

Help < Back Next > Finish Cancel





Definir Mapping rules



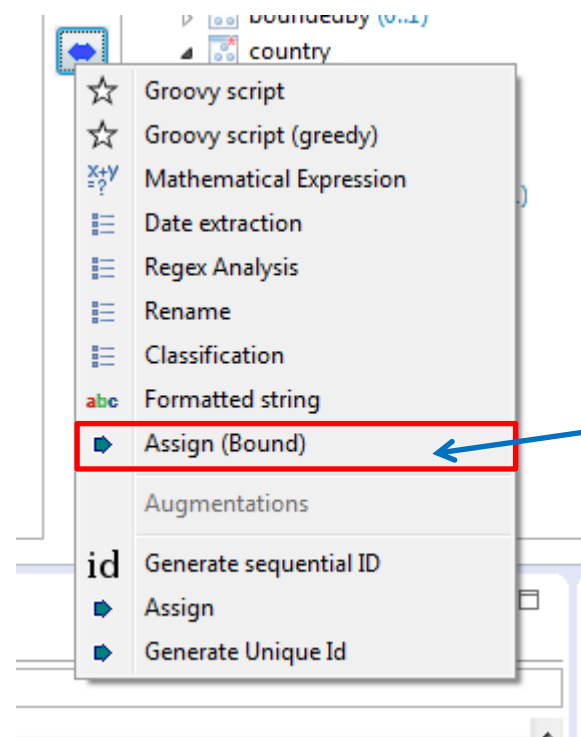
■ Outras funções

Geoserver



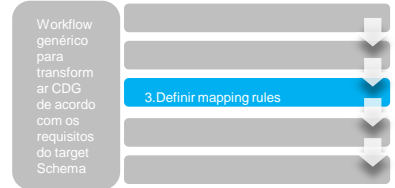
■ *Assign Bound*

■ Atribuí um valor a uma propriedade (***target***) se uma propriedade na fonte estiver presente.





Definir Mapping rules

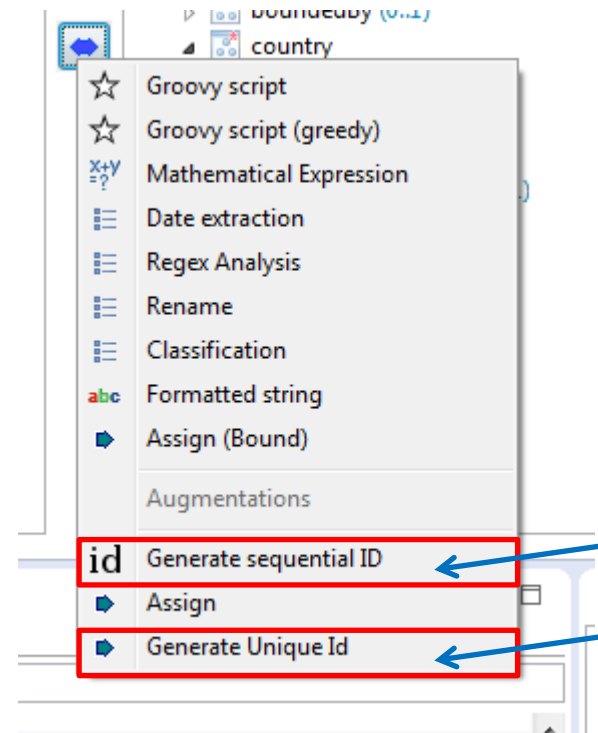


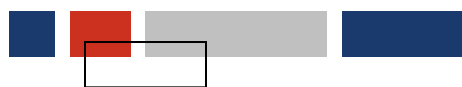
■ Outras funções

Geoserver

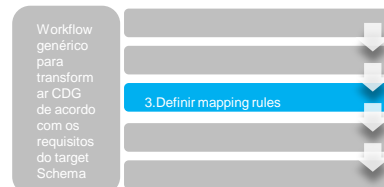
■ **Generate Sequential ID** e **Generate UniqueID**

■ Cria um identificador sequencial numérico





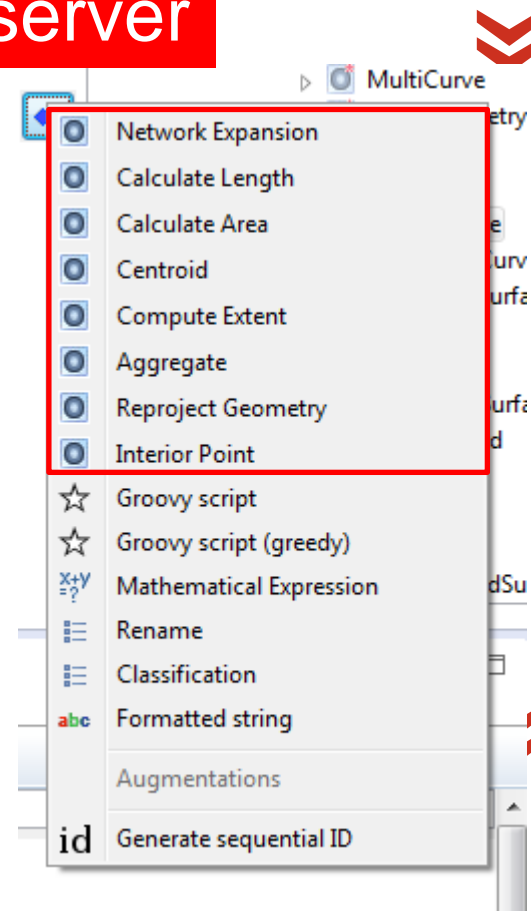
Definir Mapping rules



■ Funções de Geometria

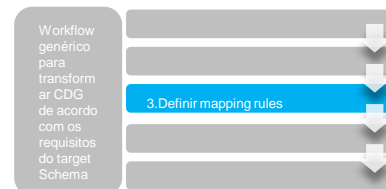
- Ficam ativas se clicar na geometria do **source schema**
- Funções de geometria simples na maioria.
- **Network expansion** permite converter para um diferente tipo de geometria (linhas -> Pontos)
- **Aggregate** - Junta geometrias

Geoserver





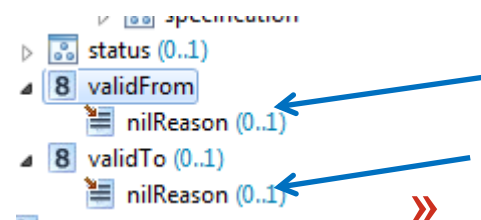
Definir Mapping rules



- Algumas propriedades no **target Schema**, especialmente as obrigatórias podem receber o valor de **void**

- As propriedades **voidable**, podem ter uma subpropriedade **nilReason**

Para caracterizar a **nilReason** podemos utilizar uma **codelist (void reason value)** está no folder **Others**.



Definir Mapping rules

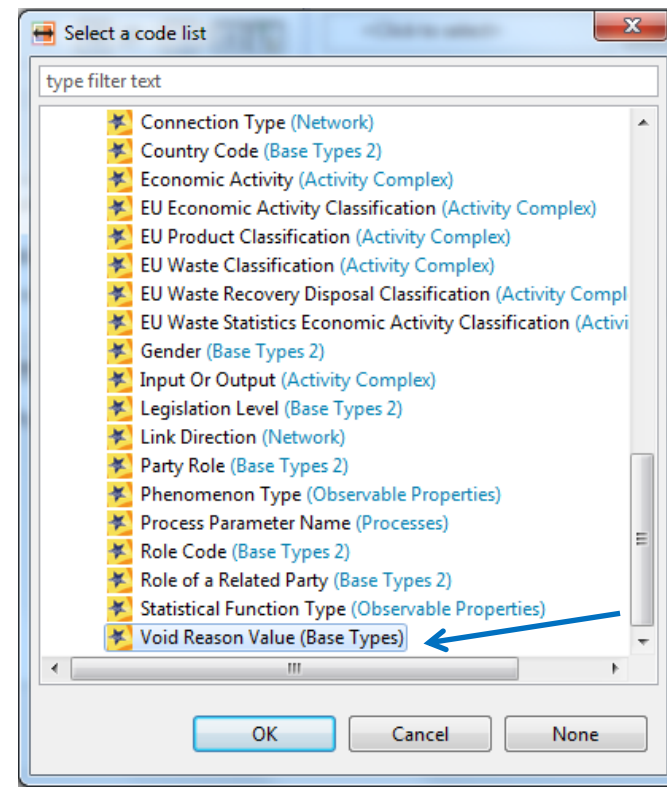
▪ ***Void reason value***

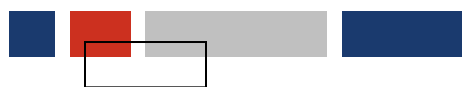
▪ Pode conter 3 valores:

▪ ***Unknown***

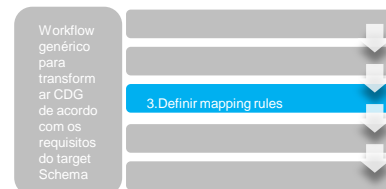
▪ The correct value for the specific spatial object is not known to, and not computable by, the data provider. However, a correct value may exist.

▪ Example When the elevation of the water body above the sea level of a certain lake has not been measured, then the reason for a void value of this property would be 'Unknown'.





Definir Mapping rules



- ***Void reason value***

- ***Unpopulated***

The characteristic is not part of the dataset maintained by the data provider. However, the characteristic may exist in the real world.

EXAMPLE When the "elevation of the water body above the sea level" has not been included in a dataset containing lake spatial objects, then the reason for a void value of this property would be 'Unpopulated'.





- ***Withheld***

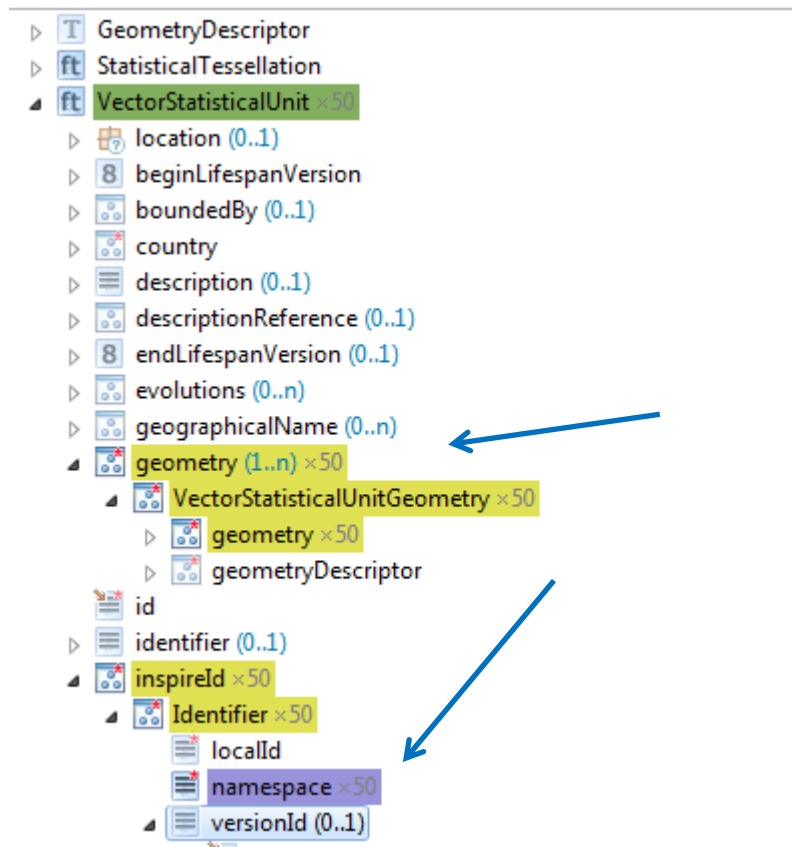
The characteristic may exist, but is confidential and not divulged by the data provider.

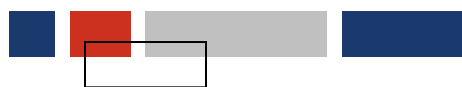
Definir Mapping rules

▪ Sistema de cores dos elementos

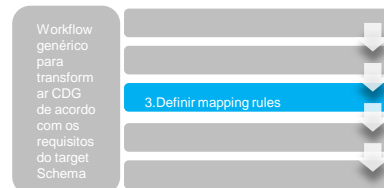
- Para identificar os elementos do **schema** já mapeados o **Hale** utiliza um sistema de cores que nos permitem ter uma noção do trabalho já realizado

	city	Not mapped
	street	Mapped explicitly with a relation
	address	Mapped implicitly due to the mapping of a sub-property
	type	Value assignment independent of the source schema



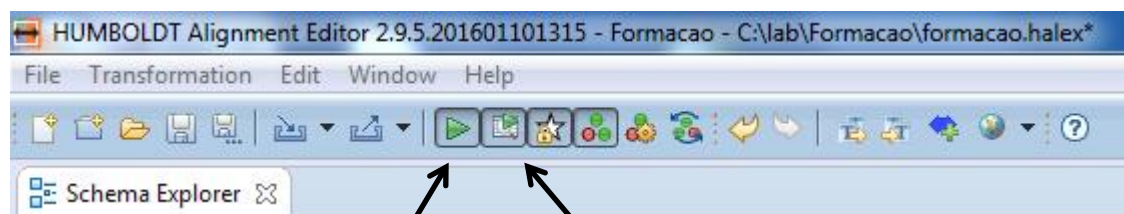


Definir Mapping rules



▪ Sistema de validação

- Verifica se existem inconsistências no mapeamento nomeadamente atributos obrigatórios no **target schema** não mapeado, ou restrições ignoradas.



Live transformation button

Permite que as instancias carregadas no projeto sejam transformadas sempre que haja um mapeamento.

Validate transformed instances automatically

Executa uma validação das instâncias carregadas sempre que haja um mapeamento.

Definir Mapping rules

▪ Sistema de validação, com erros

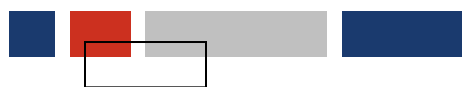
The screenshot displays a software interface with two main panes. The left pane, titled 'Error Log' and 'Properties', shows a tree view of warnings under the heading 'Instance validation'. It lists 'VectorStatisticalUnit (200 warnings)' and 'country (50 warnings)', with a 'Cardinality (showing 5 of 50)' warning for the 'country' entry. The right pane, titled 'Type hierarchy', 'Functions', 'Report List', 'Transformed Data', and 'Mapping', shows a list of validation tasks with status icons (warning, success) and timestamps. The tasks include 'Instance validation', 'Instance transformation', and 'INSPIRE code list import'. Arrows point from the text labels below to specific elements in the interface.

Total e lista de erros

Instance validation

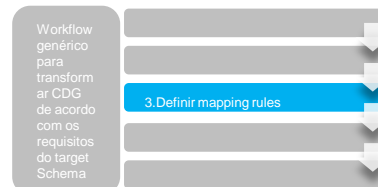
Símbolo *Warning*

Significa que existem erros



Definir Mapping rules

▪ Sistema de validação, sem erros



The screenshot displays a software interface with two main panels. The left panel, titled 'Instance validation', shows a 'Report' section with the following details: Success: true, Summary: Finished successfully, Time: Mon Apr 11 14:51:04 BST 2016, and Duration: 31 milliseconds. The right panel, titled 'Type hierarchy', shows a list of tasks for the date 14:47 2016-04-11. The first task, 'Instance validation', is highlighted with a blue box and a green checkmark. Below it, a list of other tasks is shown, each preceded by a green checkmark. A red double arrow points from the 'Instance validation' task in the right panel to the 'Instance validation' report in the left panel. Another red double arrow points from the bottom of the right panel towards the bottom right of the slide.

Report	Success:	Summary:	Time:	Duration
	true	Finished successfully	Mon Apr 11 14:51:04 BST 2016	31 milliseconds

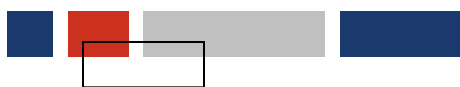
- 14:47 2016-04-11
 - ✓ Instance validation
 - ✓ Instance transformation
 - ✓ HALE XML project export
 - ✓ Instance validation
 - ✓ Instance transformation
 - ✓ Instance validation
 - ✓ Instance transformation
 - ✓ Load data into database
 - ✓ Shapefile import
 - ✓ Shapefile import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import
 - ✓ INSPIRE code list import

Lista de erros deixa de existir

Instance validation

Símbolo = visto (*checked*)





Exportar dados transformados



■ Exportar dados quando:

- não existem erros reportados no ***Instance validation***
- não existe mais nenhum mapeamento para realizar
- não existir ***instance sampling***



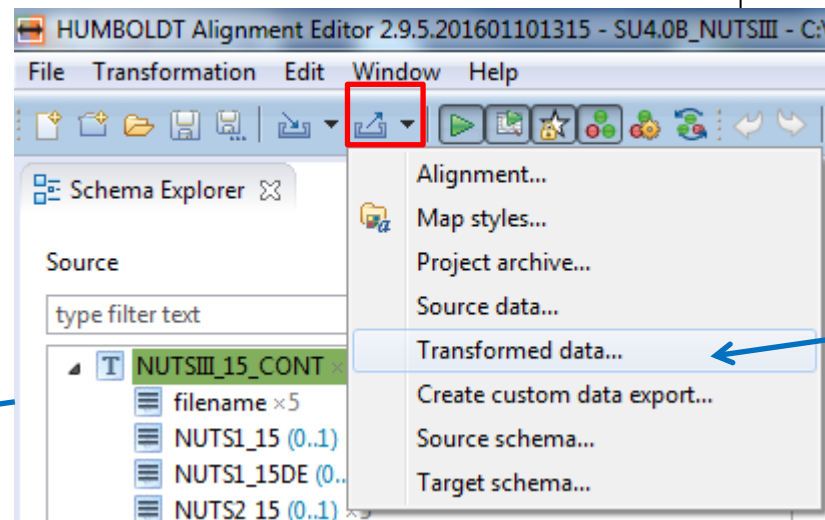
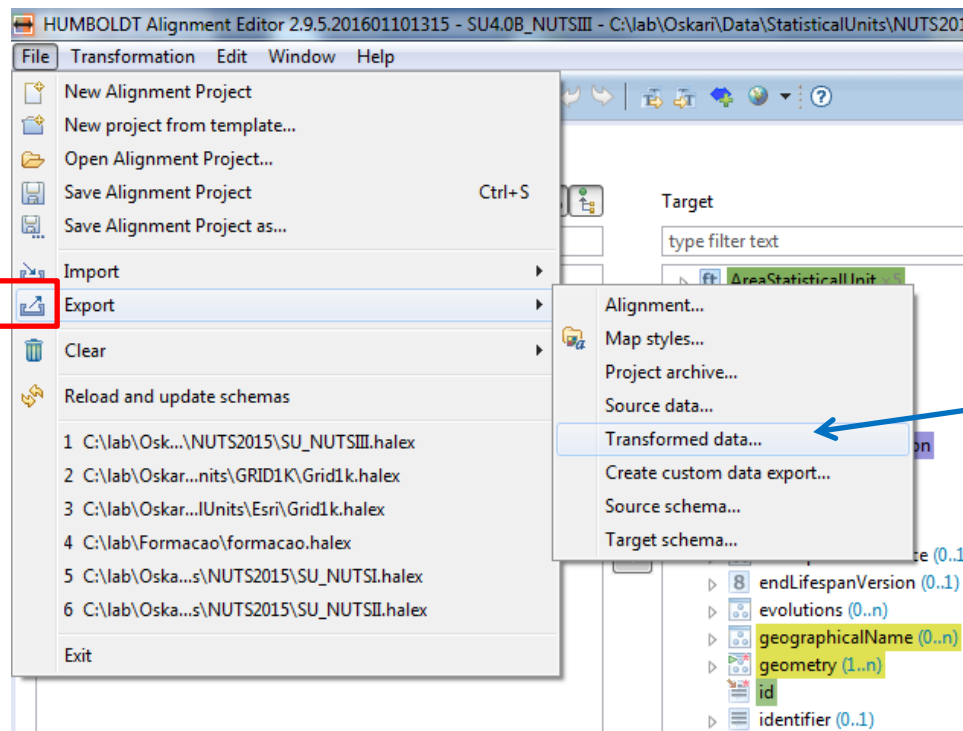
Exportar dados transformados

Workflow genérico para transformar ar CDG de acordo com os requisitos do target Schema

4. Exportar dados transformados

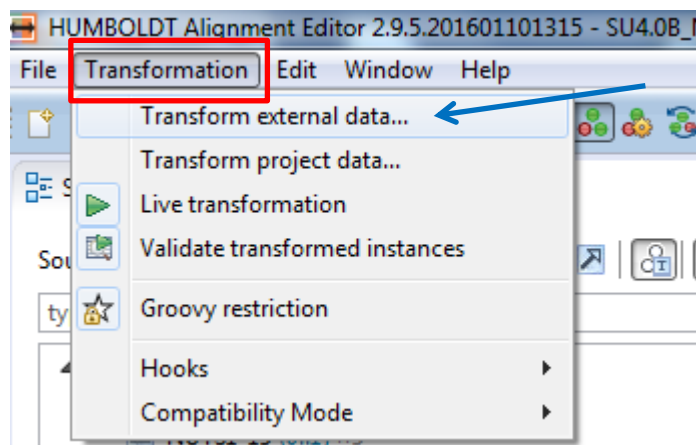
■ Existem duas formas para transformar dados

■ 1. Transformar as instâncias carregadas no projeto



Existem duas formas para transformar dados

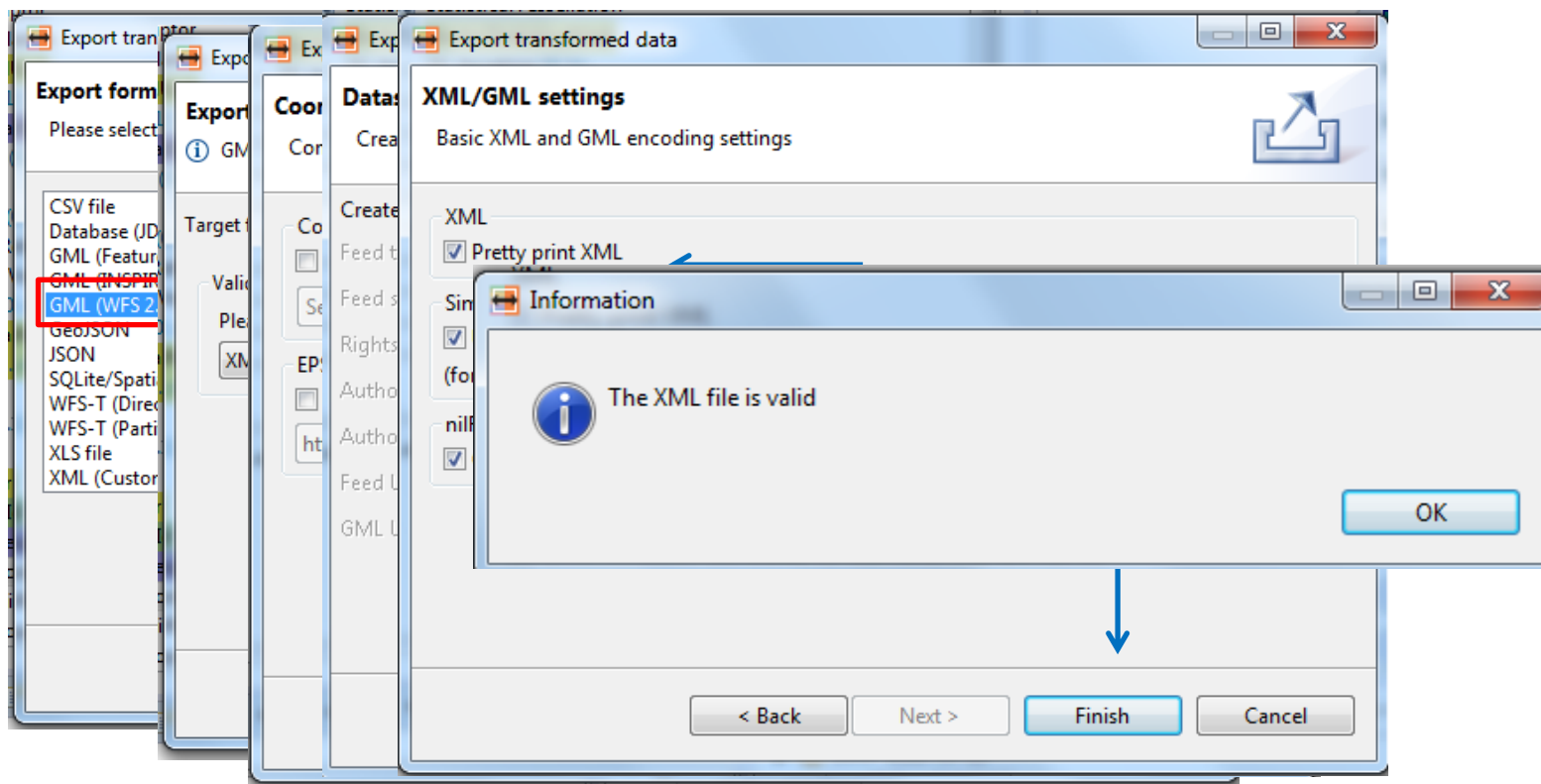
2. Transformar dados externos



- Não se recomenda este método para dados com muitas instâncias: no caso da BGRI (~200 000 registos) não chega a terminar.
- Utiliza-se para partilhar o projeto Hale com dados de Portugal continental e R.A.'s

Exportar dados transformados

■ Tipo de formato e outras opções



Validação dados transformados

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

5. Validar dados transformados

- **eENVplus Validation Service**
- http://cloud.epsilon-italia.it/eenvplus_new/
- Só 4 temas validam com **Schematron**

The screenshot displays the eENVplus Validation Service interface. At the top, the OGC logo and the service name 'eENVplus Validation Service' are visible. Below this, the text 'TEAM Engine v4' is present. The main section is titled 'Results for session s0007' and shows the 'Test Suite: GML 3.2 (ISO 19136:2007) Conformance Test Suite'. A status bar indicates 'Test tns:Main (View Details): Passed'. A 'Summary of results' table shows the following counts: Best Practice (0), Passed (1), Continue (0), Not Tested (0), Warning (0), Skipped (0), Failed (0), and Failed (Inherited) (0). Below the table, there is a link to 'See the detailed test report.' and buttons for 'Execute this session again', 'Delete this session', 'Download log Files', and 'Create execution log report file'. A 'Sessions list' link is also present. At the bottom, a note states 'If you have any questions or suggestions, feel free to contact the eENVplus team.' and the version 'TEAM Engine 4.0.5' is listed.

OGC[®]
Making location count.

eENVplus Validation Service

TEAM Engine v4

Results for session s0007

Test Suite: GML 3.2 (ISO 19136:2007) Conformance Test Suite

✓ Test tns:Main (View Details): Passed

Summary of results

Best Practice	Passed	Continue	Not Tested	Warning	Skipped	Failed	Failed (Inherited)
0	1	0	0	0	0	0	0

See the [detailed test report.](#)

[Execute this session again](#) [Delete this session](#) [Download log Files](#) [Create execution log report file](#)

[Sessions list](#)

If you have any questions or suggestions, feel free to contact the [eENVplus team.](#)

TEAM Engine 4.0.5

Validação dados transformados

Workflow genérico para transformar CDG de acordo com os requisitos do target Schema

5. Validar dados transformados

■ **Oxygen**

■ **Software com custos**

NUTSIII.gml [C:\tmp\NUTSIII.gml] - <Oxygen/> XML Editor (Evaluation use only)

File Edit Find Project Options Tools Document Window Help



XPath 2.0 Execute XPath on 'Current File'

Project

sample.xpr

- css
- debugger
- epub
- fo
- import
- json
- jsp

The Master Files support is disabled

Enable [Read more](#)

Project Open/Find Resource

Outline

Element name filter

- wfs:FeatureCollection <http://inspire.ec.europa.eu/schemas/au/4.0/>
- wfs:member
- wfs:member
- wfs:member
- wfs:member
- wfs:member

• NUTSIII.gml

wfs:FeatureCollection

```
1 <?xml version="1.0" ?>
2 <wfs:FeatureCollection xmlns:au="http://inspire.ec.europa.eu/schemas/au/4.0"
3   xmlns:gts="http://www.isotc211.org/2005/gts"
4   xmlns:gco="http://www.isotc211.org/2005/gco"
5   xmlns:ad="http://inspire.ec.europa.eu/schemas/ad/4.0"
6   xmlns:hfp="http://www.w3.org/2001/XMLSchema-hasFacetAndProperty"
7   xmlns:gn="http://inspire.ec.europa.eu/schemas/gn/4.0"
8   xmlns:bu-base="http://inspire.ec.europa.eu/schemas/bu-base/4.0"
9   xmlns:base="http://inspire.ec.europa.eu/schemas/base/3.3"
10  xmlns:cp="http://inspire.ec.europa.eu/schemas/cp/4.0"
11  xmlns:su-core="http://inspire.ec.europa.eu/schemas/su-core/4.0"
12  xmlns:xlink="http://www.w3.org/1999/xlink"
13  xmlns:net="http://inspire.ec.europa.eu/schemas/net/4.0"
14  xmlns:tn="http://inspire.ec.europa.eu/schemas/tn/4.0"
15  xmlns:ns1="http://www.w3.org/2001/XMLSchema-instance"
16  xmlns:base2="http://inspire.ec.europa.eu/schemas/base2/4.0"
17  xmlns:gml="http://www.opengis.net/gml/3.2"
18  xmlns:su-vector="http://inspire.ec.europa.eu/schemas/su-vector/4.0"
19  xmlns:gss="http://www.isotc211.org/2005/gss"
20  xmlns:sc="http://www.interact.org/2005/sc"
21  xmlns:gssr="http://www.isotc211.org/2005/gssr"
22  xmlns:gmd="http://www.isotc211.org/2005/gmd"
23  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
24  xmlns:wfs="http://www.opengis.net/wfs/2.0"
25  xsi:schemaLocation="http://inspire.ec.europa.eu/schemas/au/4.0/StatisticalUnitVector.xsd
26  http://www.opengis.net/wfs/2.0 http://schemas.opengis.net/wfs/2.0/wfs.xsd"
27  numberMatched="5" numberReturned="5" timeStamp="2016-04-11T16:10:09.578+01:00">
28   <wfs:member>
29     <su-vector:AreaStatisticalUnit gml:id="Id1">
30       <gml:name>NUTSIII 2015</gml:name>
31       <su-vector:inspireId>
```

Validate with

URL: <file:///C:/lab/Oskari/Data/SLD/StatisticalUnitVector.xsd>

Schema type: XML Schema ☒ Embedded schematron rules

Public ID:

Schematron phase:

OK

Cancel

E [Xerces] cvc-elt.1.a: Cannot find the declaration of element 'wfs:FeatureCollection'.

Attributes

wfs:FeatureCollection [http://www.opengis.net/wfs/2.0]

Attribute	Value
numberMatched	5
numberReturned	5
timeStamp	2016-04-11T16:10:09.578+01:00
xmlns:ad	http://inspire.ec.europa.eu/schemas/ad/4.0
xmlns:au	http://inspire.ec.europa.eu/schemas/au/4.0
xmlns:base	http://inspire.ec.europa.eu/schemas/base/3.3
xmlns:base2	http://inspire.ec.europa.eu/schemas/base2/4.0
xmlns:bu-base	http://inspire.ec.europa.eu/schemas/bu-base/4.0
xmlns:cp	http://inspire.ec.europa.eu/schemas/cp/4.0

ios - NUTSIII.gml

Type