

# Tvorba GML souboru odpovídajícího požadavkům INSPIRE pomocí programu HALE studio

Miroslav Fanta  
CENIA, oddělení geoinformatiky  
poslední aktualizace: 20.10.2020  
miroslav.fanta@cenia.cz

## Příklad zpracování dat ČHMÚ – monitorovací stanice kvality ovzduší

### ***Přípravné kroky:***

Seznámit se s datovou strukturou INSPIRE platnou pro dané téma a datovou sadu – viz dokument „[Jak pracovat s datovými specifikacemi INSPIRE](#)“

Nainstalovat open source software **HALE studio**: <https://www.wetransform.to/products/halestudio/>

Další užitečné odkazy pro práci s HALE studiem:

- uživatelské forum: <https://discuss.wetransform.to/>
- novinky: <https://www.wetransform.to/category/news/>
- webináře: <https://www.wetransform.to/downloads/webinars/>
- nápověda: <http://help.halestudio.org/>
- doporučená prezentace – **Introduction to HALE software**:  
[http://sadl.kuleuven.be/docs/Smespire\\_training\\_presentation\\_adv4\\_INSPIRE%20Data%20specifications%20advanced\\_part6.pdf](http://sadl.kuleuven.be/docs/Smespire_training_presentation_adv4_INSPIRE%20Data%20specifications%20advanced_part6.pdf)

### ***Upozornění na budoucí změnu protokolu http -> https:***

na několika místech tohoto dokumentu se v popisovaném vzorovém projektu v programu **HALE studio** používají **odkazy na stránky INSPIRE** (schémata, codelisty, atd), např:

<http://inspire.ec.europa.eu/schemas/ef/4.0/EnvironmentalMonitoringFacilities.xsd>

<http://inspire.ec.europa.eu/codelist/MediaValue>

atd.

V letech 2020 – 2022 mají být **stránky INSPIRE převáděny z protokolu http na protokol https**. Tuto změnu bude následně třeba zakomponovat do již existujících projektů HALE, aby i nadále fungovaly tak jako dříve.

# Transformace dat do struktury INSPIRE

## Příprava vstupních dat

### Úprava vstupních dat do 2D, pokud obsahují souřadnici Z

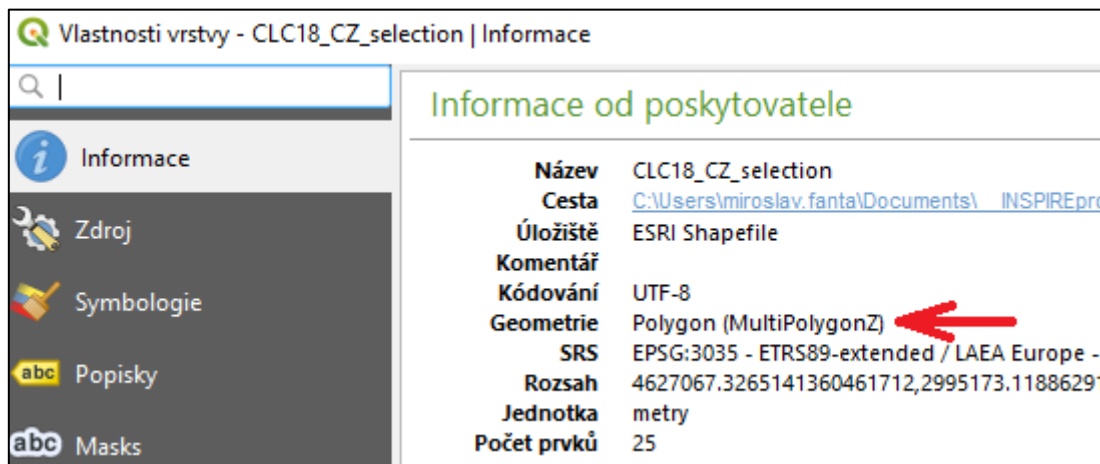
INSPIRE validátor (zjištěno v říjnu 2020 v testu Interoperable data sets in GML: INSPIRE GML application schemas, General requirements) nepracuje správně s daty která mají v GML souboru 3. souřadnici, i když daný prostorový objekt obsahuje specifikaci srsDimension="3". Toto platí i když je 3. souřadnice ve všech případech nulová.

Vstupní data mohou nulovou 3. souřadnici neočekávaně obsahovat, i když nemá žádný praktický význam. Možnost eliminace případné nulové třetí souřadnice v programu HALE studio během transformace dat do struktury INSPIRE se nepodařilo zjistit. Z uvedeného důvodu je třeba vstupní data (shapefile) ještě před načtením do HALE studio zkontrolovat, zda nulovou třetí souřadnici neobsahují a případně ji odstranit.

#### **Postup v QGISu:**

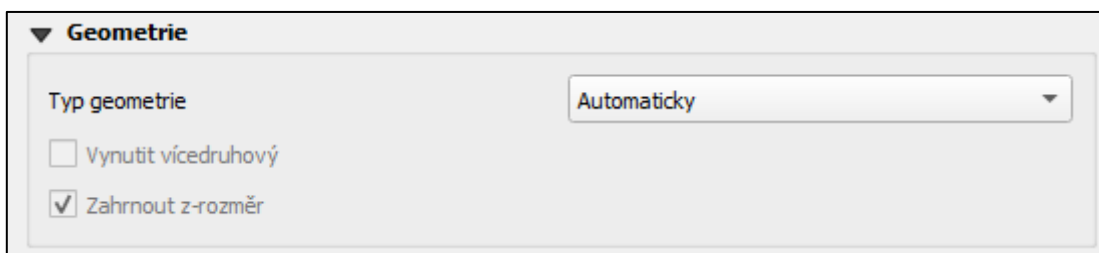
Ověření existence 3. souřadnice:

Načíst danou vektorovou datovou sadu do QGISu - klik pravým tlačítkem na vrstvu – Vlastnosti... - Informace



Odstranění nulové 3. souřadnice:

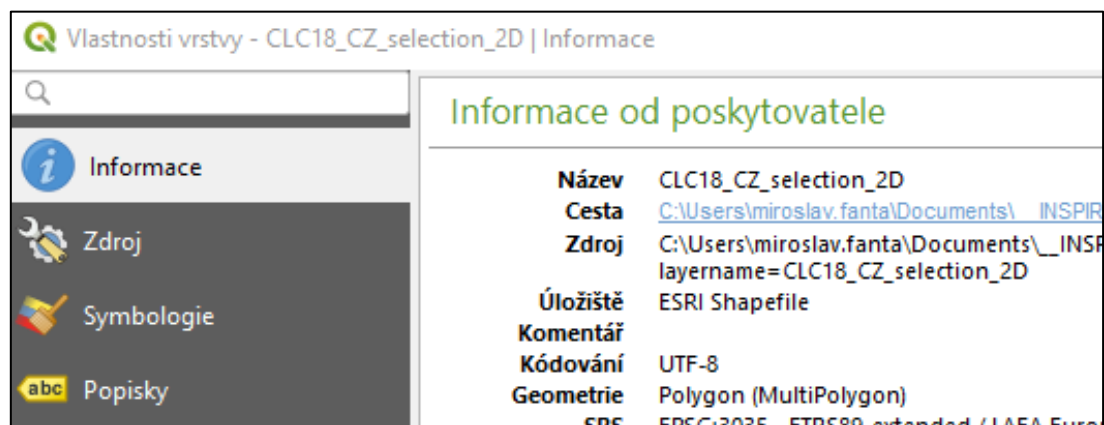
Klik pravým tlačítkem na vrstvu – Export – okno Uložit prvky jako... - rozbalit rámeček Geometrie – v něm bude zaškrtnutý checkbox „Zahrnout Z-rozměr“



Aby bylo možno toto zaškrtnutí zrušit, je třeba nejdříve změnit *Typ geometrie* z *Automaticky* na konkrétní typ (např. Polygon) a poté daný checkbox odškrtnout a dokončit export dat.



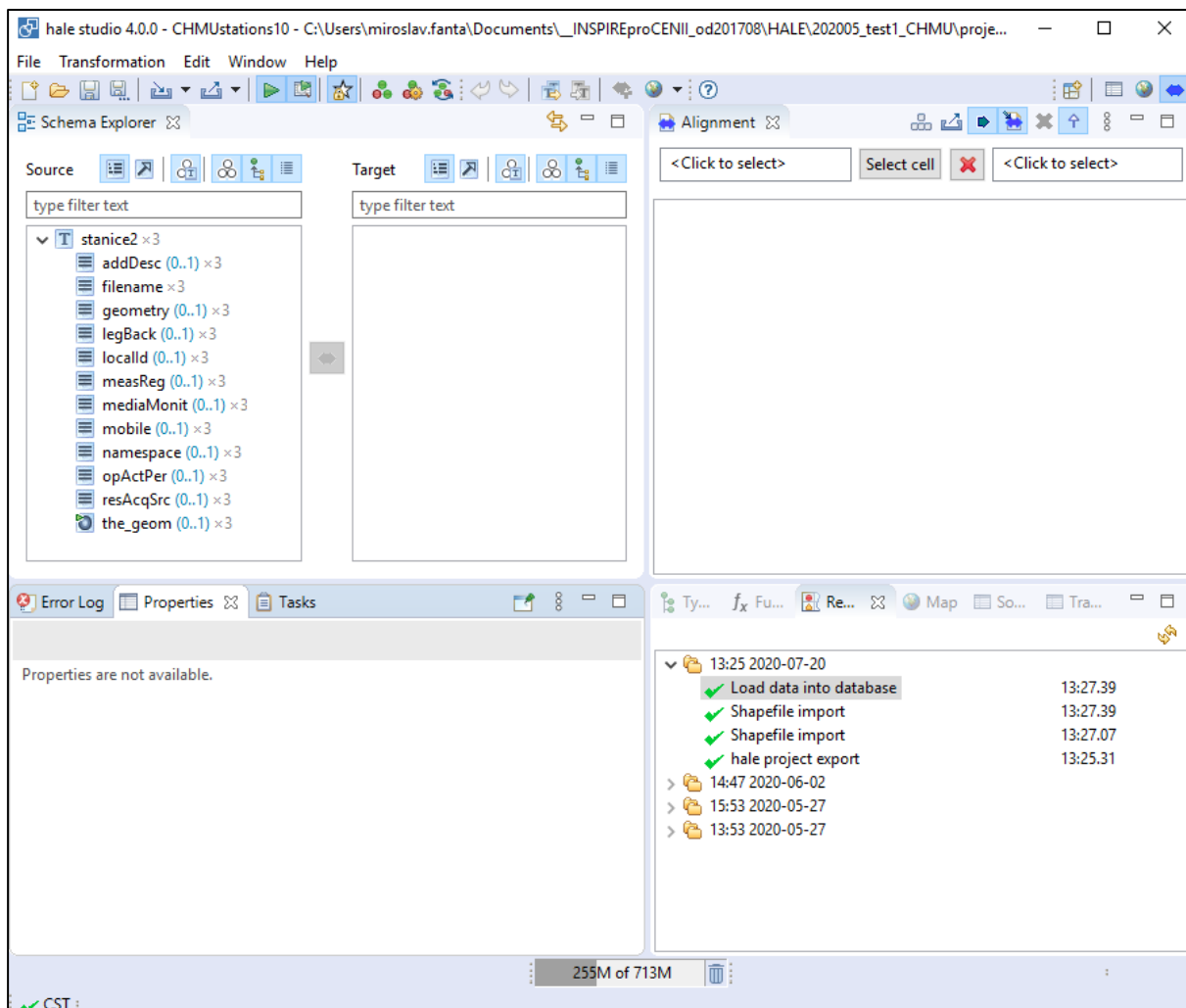
Takto vyexportovaná data již budou mít pouze dvě souřadnice a po transformaci v programu HALE studio bude mít pochopitelně pouze dvě souřadnice i výstupní soubor GML.



# Úvodní operace

## Otevření projektu v programu HALE studio

Výchozí rozložení oken programu HALE studio:



Návrat k výchozímu rozložení oken: Default button vpravo nahoře (oboustranná modrá šipka ).

**Uložení projektu:** File – Save alignment project as – Export format: hale project – zadat jméno projektu a v dalším kroku jméno souboru, ve kterém je projekt uložen.

## Načtení zdrojového schématu a zdrojových dat (source schema, source data), zobrazení source dat

source schema:

File – Import – Source schema – okno Import location – karta From file - vybrat shp soubor (další karty v okně Import location jsou From URL / From preset / From WFS / From database (JDBC))

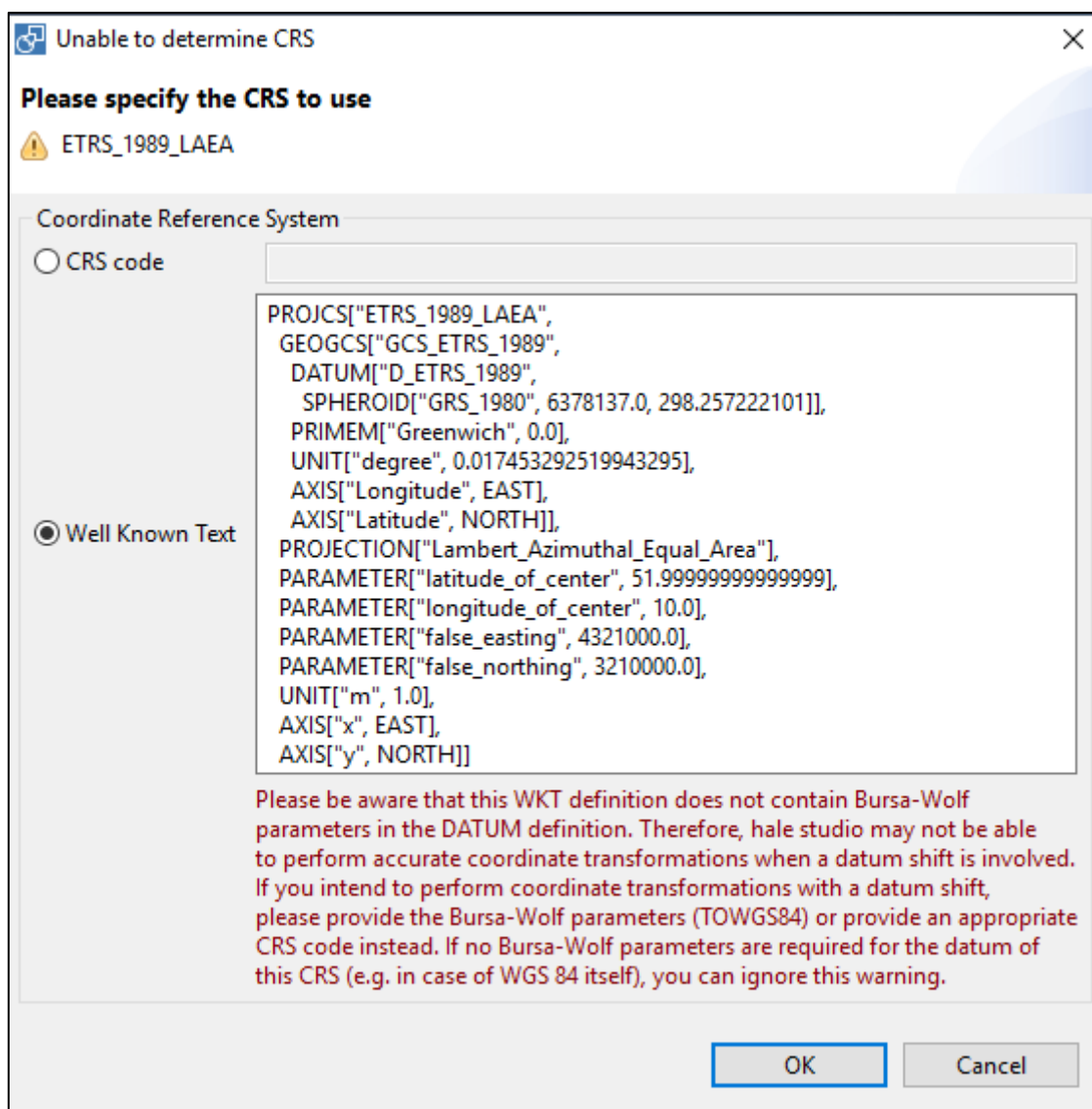
source data:

File – Import – Source data – okno Import location - karta From file - vybrat shp soubor (tentýž jako výše)

(další karty v okně Import location jsou From URL / From WFS / From database (JDBC))

### **Upozornění:**

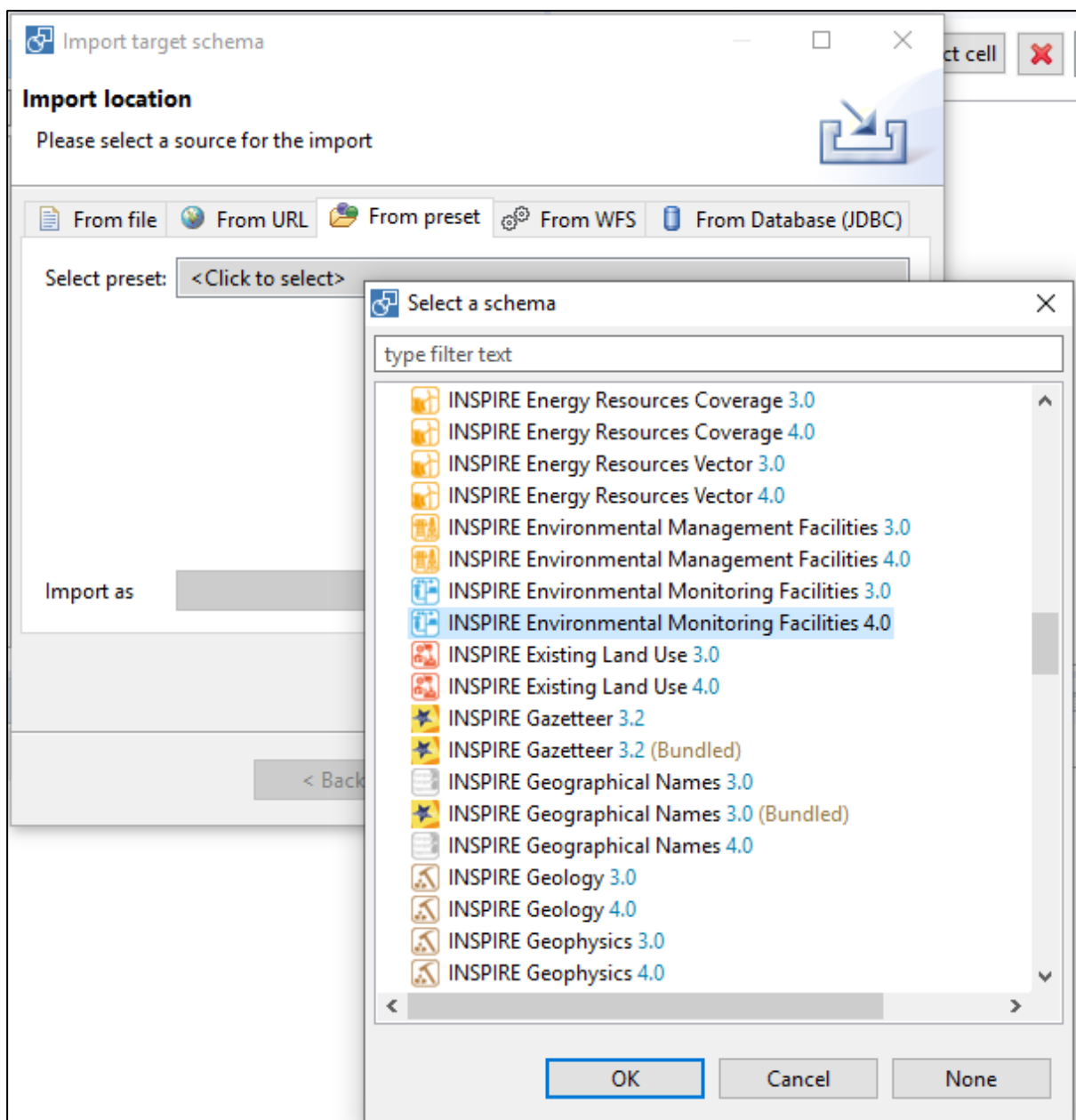
v posledním kroku importu zdrojových dat při zadávání souřadnicového systému ponechat to, co se nabídne v rámečku Well Known Text – následně ověřit v mapovém okně, že zdrojová data jsou zobrazena správně (při zadání na pohled téhož CRS, který má daný shapefile nadefinován ve svém .prj souboru, pomocí příslušného EPSG nedaly správný výsledek – data se v map. okně zobrazila jinde (v případě monitorovacích stanic ČHMÚ, která jsou dle .prj souboru v projekci „ETRS\_1989\_LAEA“ se při zadání odpovídajícího EPSG:3035 data zobrazila jižně od Islandu a otočená o 90° proti směru hod. ručiček; pokud se zvolila možnost ponechání údajů uvedených automaticky v rámečku Well Known Text, data se zobrazila správně v Praze); asi to bude případ od případu – parametry dané projekce v definici EPSG a v prj souboru se mohou lišit; při zadávání CRS pomocí EPSG kódu zapsat včetně prefixu „EPSG:“, tj. např. „EPSG:3035“, „EPSG:4258“ atd.



Po načtení source dat se v okně Source objeví u „root“ položky (Stanice2) i u všech jejích atributů počet načtených prvků (např „x 3“ pokud se načtou 3 prvky)

## Načtení cílového (target) schématu

File – Import – Target schema – okno Import location - karta From preset -> INSPIRE environmental monitoring facilities 4.0

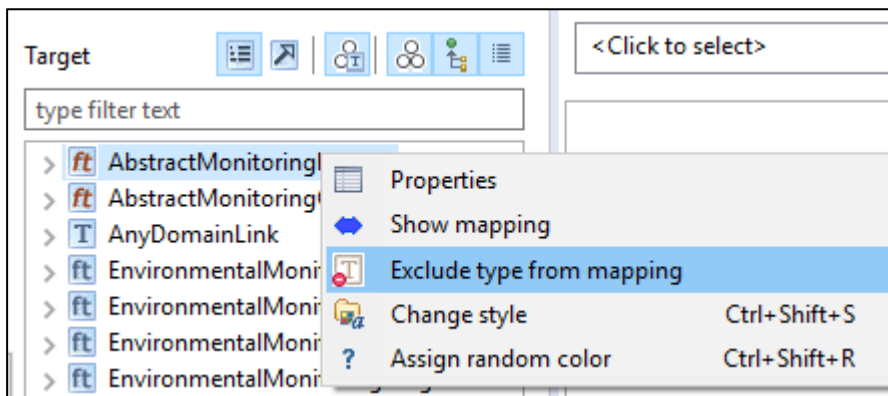


nebo xsd schema stáhnout zde:

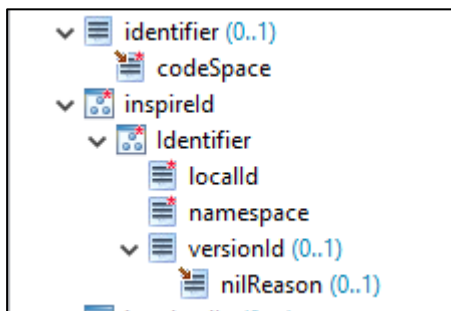
<http://inspire.ec.europa.eu/schemas/ef/4.0/EnvironmentalMonitoringFacilities.xsd>

a poté v okně Import location použít namísto karty From preset kartu From file

Objeví-li se po natažení target schématu v daném okně více položek, které v projektu nebudou potřeba, odstraní se kliknutím pravým tlačítkem na položku a výběrem „Exclude type from mapping“



Ikona „ft“ u target dat = feature type  
 Červenou hvězdičkou označena povinná pole  
 Upozornění: tato hvězdička se může objevit až v určité úrovni vnoření!!



## Načtení INSPIRE codelistů

Zjistit v daném schématu (v datových specifikacích, UML diagramu, Mapping table), zda daný typ objektu (feature type) používá v povinných položkách codelisty - u environmental monitoring facilities jsou dva takovéto parametry povinné:

non-voidable parametr:

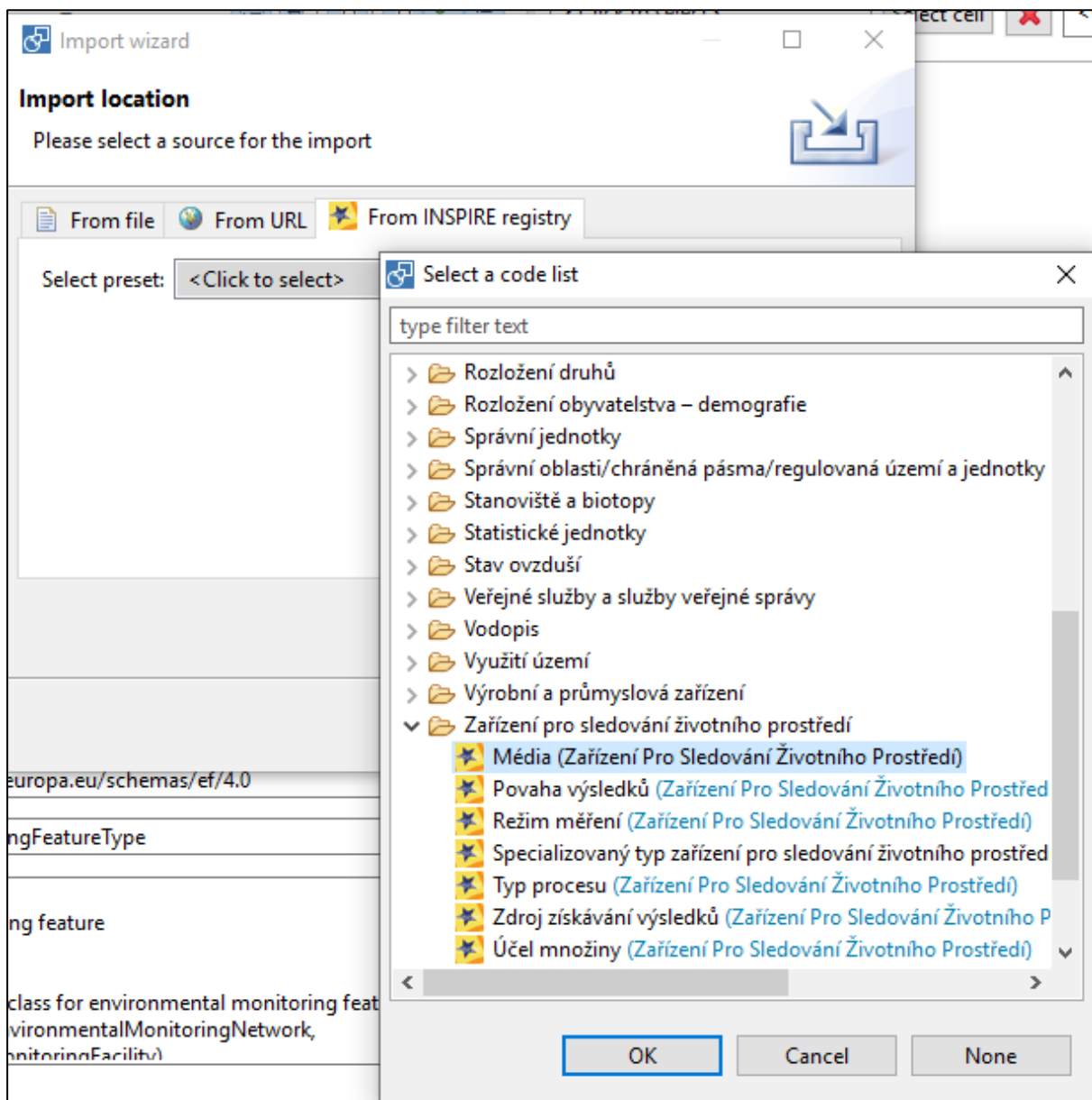
mediaMonitored (multiplicity 1...\*) - codelist: <http://inspire.ec.europa.eu/codelist/MediaValue>

voidable parametr:

measurementRegime (multiplicity 1) – codelist:  
<http://inspire.ec.europa.eu/codelist/MeasurementRegimeValue>

Postup načtení codelistů:

File – Import – Code list – okno Import location - karta From Inspire registry – vybrat příslušné téma INSPIRE a v něm správný codelist a načíst příslušný codelist (další karty okna Import location jsou From file a From URL)



## Přidání dynamických proměnných do projektu

Výhoda použití proměnných – hodnota se nastaví, případně změní jen jednou, na jednom místě, přičemž je použita na mnoha místech

Window – Settings – okno Preferences – Project – Variables – zelené tlačítko „+“ vpravo nahoře – zadat jméno proměnné a její hodnotu

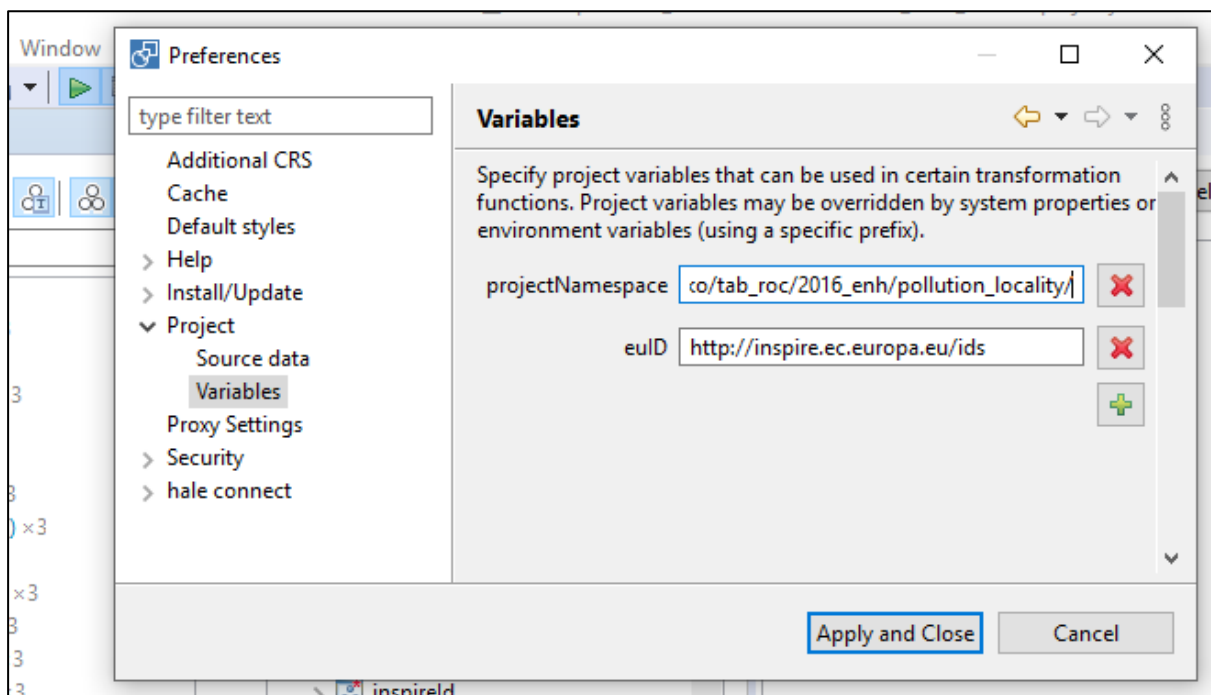
**projectNamespace**, vložena hodnota fiktivního URL:

[http://portal.chmi.cz/files/portal/docs/uoco/isko/tab\\_roc/2016\\_enh/pollution\\_locality/](http://portal.chmi.cz/files/portal/docs/uoco/isko/tab_roc/2016_enh/pollution_locality/)

(!!!!!! pozn: nezapomenout na lomítko na konci URL)

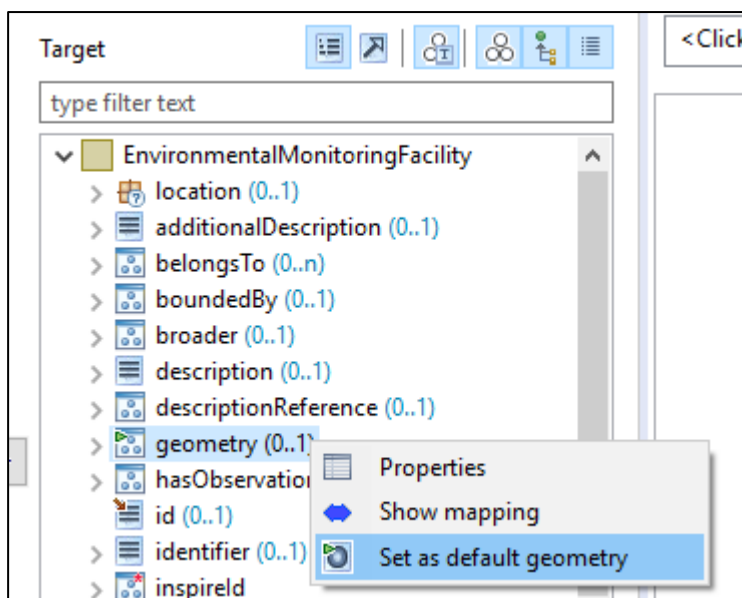
**euID**, potřebné pro vytvoření INSPIRE identifikátoru: <http://inspire.ec.europa.eu/ids>

(euID codelist, to je vhodné pro každý INSPIRE projekt)

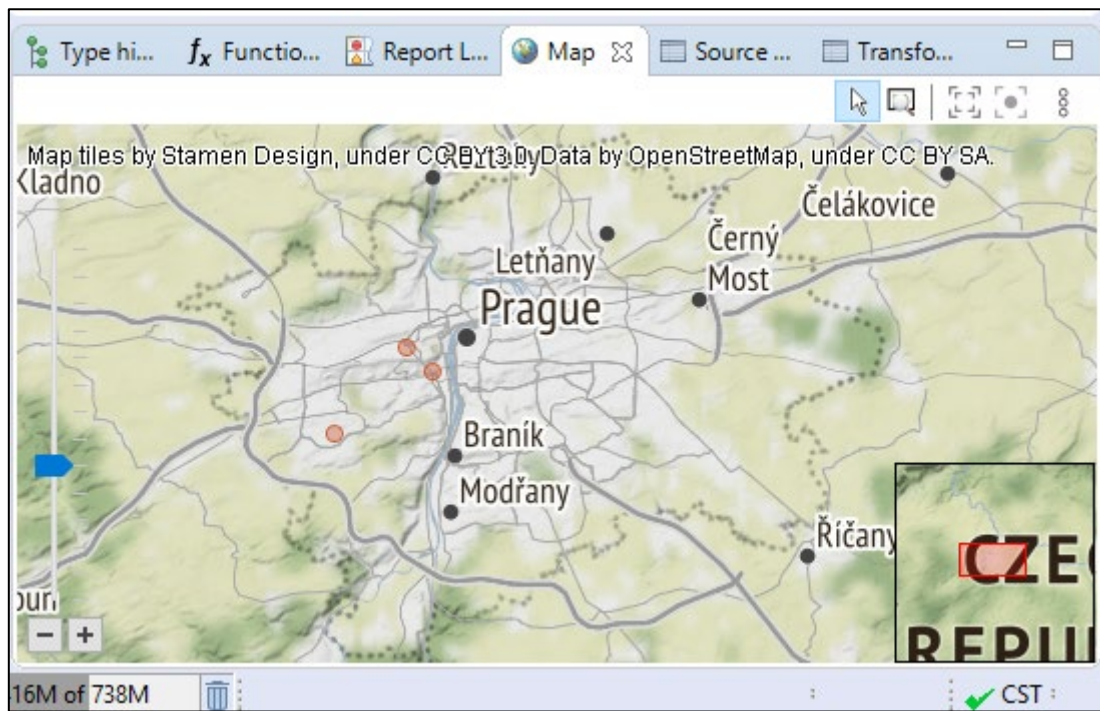


## Zobrazení zdrojových a cílových dat v mapovém okně

Pro zobrazení dat v mapovém okně musí být daný parametr (!!!! **Source: the\_geom, Target: geometry**) nastaven jako výchozí geometrie (klik pravým na parametr – Set as default geometry). U takto nastavených parametrů se v levém horním rohu jejich ikony objeví zelená šipka mířící vpravo

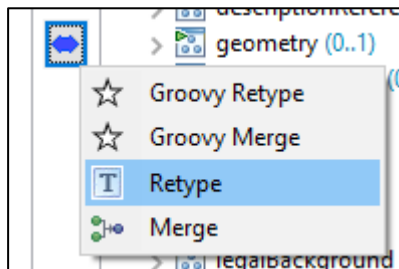


Mapové okno aplikace HALE studio se zobrazením zdrojových dat (vzorová data ČHMÚ – 3 stanice monitorující kvalitu vzduchu v Praze – označeny červenými kolečky):

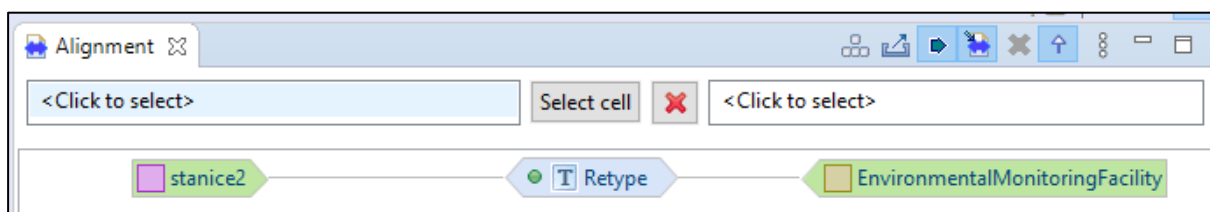


## Postup definice schématu transformace (alignment, mapování)

Klik na „root“ položky (feature type) v oknech Source i Target (v daném případě „Stanice2“ a „EnvironmentalMonitoringFacility“). Poté se stane funkční ikona pro výběr **mapovací funkce** (oboustranná šipka mezi okny Source a Target), klik na ní levým tlačítkem, z otevřeného menu zvolit mapovací funkci „**Retype**“.



Po výběru funkce **Retype** se do okna **Alignment** přidá základní mapovací vazba mezi Source a Target datasety (Stanice2 a EnvironmentalMonitoringFacility).



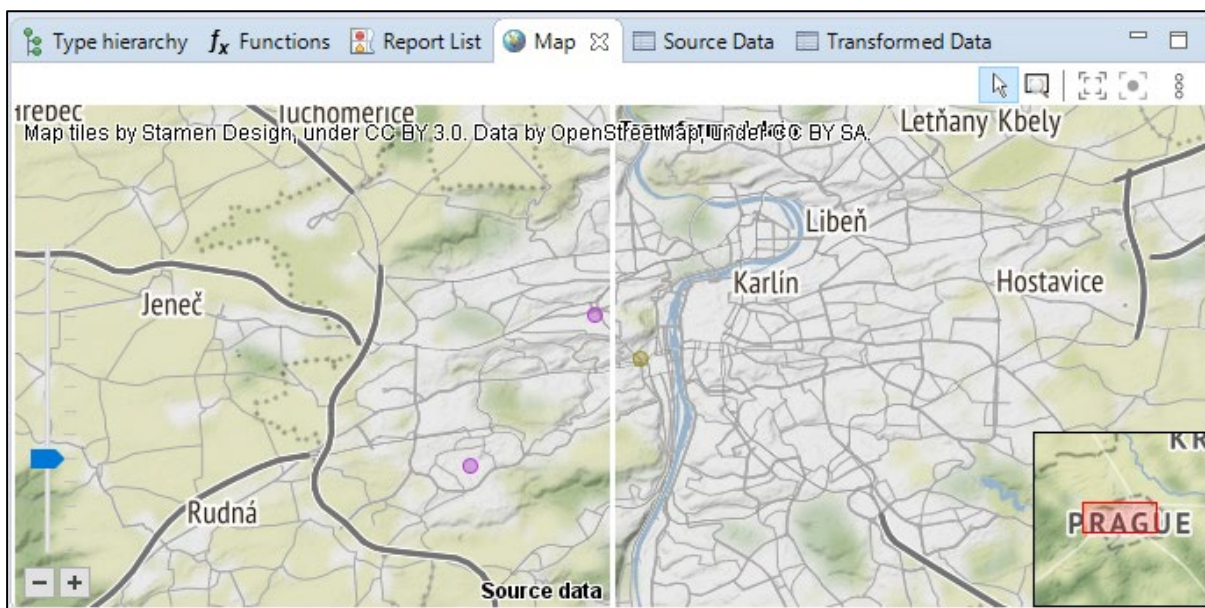
**Retype** kopíruje prvky (feature) z jednoho feature typu do druhého. Po provedení Retype se v target okně objeví u „root“ položky počet překopírovaných prvků (analogicky jako u Source, tj. „x 3“, prozatím nikoliv ovšem u atributů).

## Mapování geometrie

Source – klik na položku **the\_geom**, Target – klik na položku **geometry** (v základní úrovni, nerozklikávat) – klik na tlačítko mapovací funkce - vybrat **Rename** (zkopíruje geometrii) – v target okně se u položky geometry objeví počet prvků („x 3“) a v okně Alignment se objeví mapování geometrie z položky the\_geom přes mapovací funkci Rename na položku geometry  
Též by asi bylo možno použít mapovací funkci **Reproject geometry**

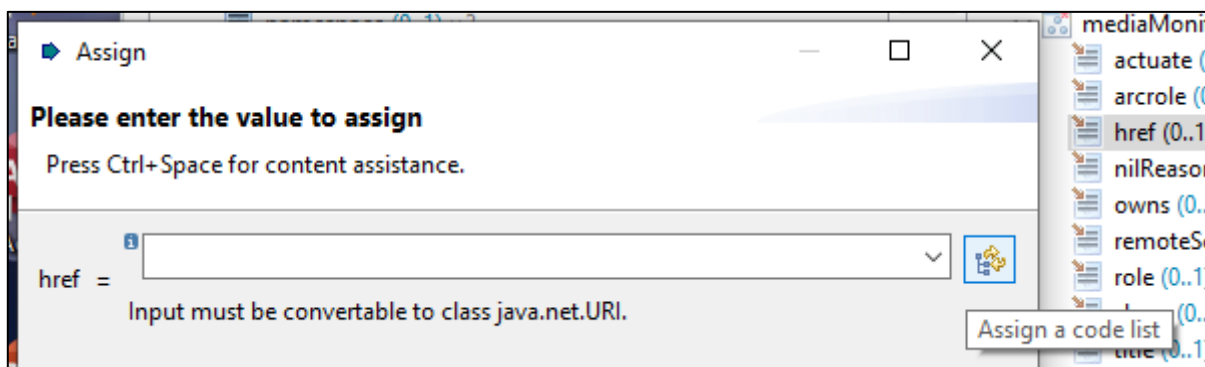
Jsou-li již načtena **source data** a je-li položka **Target: geometry** nastavena jako default geometry (viz výše „**Zobrazení source a target dat v mapovém okně**“), zobrazí se **target data v mapovém okně** ihned po nastavení mapování geometrie, není třeba provádět vlastní transformaci a export dat do výstupního souboru.

Dělené mapové okno – vlevo Source, vpravo Target; mapu lze pod dělicí linií posunovat na obě strany. V daném příkladě je nejvýhodnější monitorovací stanice zobrazena v části Target. Výběr způsobu zobrazení v mapovém okně se provádí tlačítkem „View menu“ v PH rohu mapového okna.

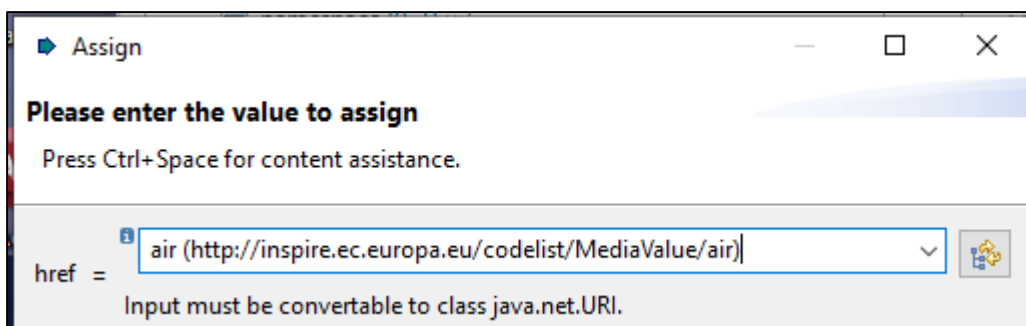


## Přiřazení hodnoty z codelistu

Pokud mají všechny prvky prostorových dat v daném parametru stejnou hodnotu, není třeba žádný vstup ze source dat. Rozkliknout daný atribut v okně Target (**mediaMonitored**) – klik na **href** – klik na ikonu mapovací funkce – zvolit mapovací funkci **Assign** – otevře se **okno Assign** – Next – vpravo od rámečku „href =“ je funkce „Assign a code list“



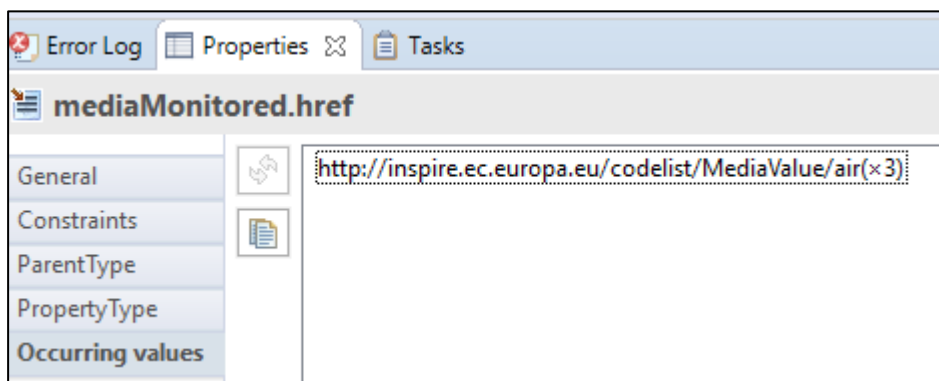
– okno Assign a code list – vybrat code list z nabídky dříve natažených – OK – zpět okno Assign – rozbalovací šipka vpravo u rámečku „href =“ – nabídka hodnot – vybrat vhodnou („air (http://inspire.ec.europa.eu/codelist/MediaValue/air)“) -



– operaci dokončit tlačítkem Finish

-> u položky mediaMonitored/href se objeví počet prvků („x 3“)

**Ověření přiřazené hodnoty:** klik pravým na mediaMonitored/href – vybrat **Properties** – vlevo dole rámeček Properties – klik na Occuring values v menu po levé straně – klik na funkci obnovení hodnot (dvě šipky proti směru hodin ručiček , mezi menu po levé straně a rámečkem pro výpis hodnot) – v rámečku se objeví přiřazená hodnota včetně počtu prvků („x 3“)

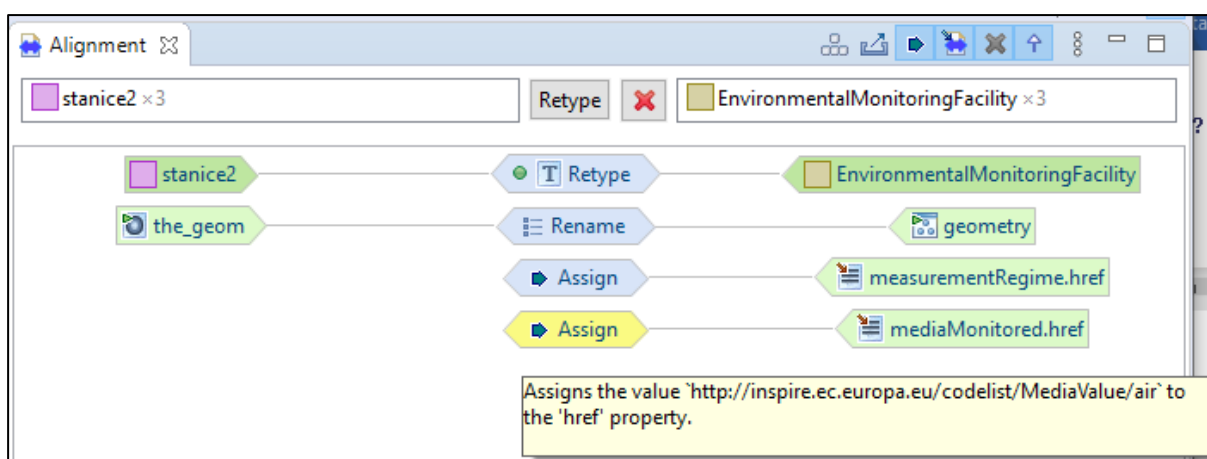


Analogicky:

**measurementRegime/href** – Assign -

<https://inspire.ec.europa.eu/codelist/MeasurementRegimeValue/periodicDataCollection>

Okno Alignment po namapování geometrie a přiřazení hodnot z codelistů. Tooltip v rámečku se žlutým pozadím se objeví po kliknutí na ikonu mapovací funkce Assign:



## Přiřazení hodnoty identifikátoru

V target struktuře jsou následující položky

**id** (0..1), bez další vnořené úrovně; jedná se o tzv. gml\_id

**identifier** (0..1)

identifier/codeSpace, povinná položka

**inspireId**, bez násobnosti, tj. násobnost 1, povinná položka

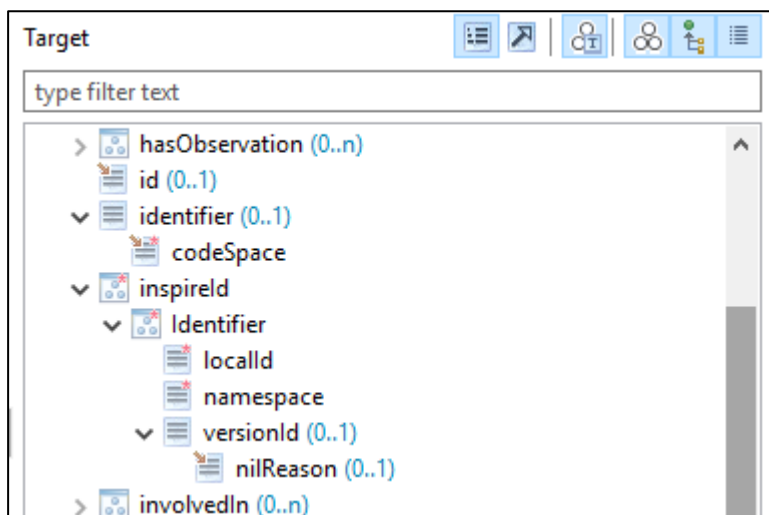
inspireId/Identifier, povinná položka

inspireId/Identifier/localId, povinná položka

/namespace, povinná položka

/versionId (0..1)

/nilReason(0..1)



Podrobně je rozlišení a význam jednotlivých identifikátorů popsán zde:

<https://www.wetransform.to/news/2018/02/12/best-practices-for-inspire-ids/>

včetně několika příkladů a vhodných pravidel. Z této stránky doporučujeme především následující část:

**gml:id**: This is the **mandatory** XML element ID, and it is encoded as an attribute of the element. It is used to uniquely identify that element in the current document, and serves to identify the target object of an Xlink. It has to match a defined pattern, e.g. it must start with a letter or underscore. It is first and foremost a technical identifier, though it should be stable over time (e.g. over multiple transformation runs) and should thus be grounded in a property of the source feature. Only if it is stable over time, Xlink references across documents can actually work. The **gml:id** is used by the WFS standard query **GetFeatureById**.

**inspireId**: This is a specific, often **mandatory**, complex property of INSPIRE objects, which consists of three sub-properties - **localId**, **namespace**, and **version**. The INSPIRE ID should be stable, and is usually used to clearly identify the object in its specific domain. Often, existing keys are re-used 1:1 as the **localId**.

**gml:identifier**: This is the **optional** external element ID, i.e. it should include a namespace to make it globally unique, not just in the current document. It is a standard property of all GML objects, it is encoded as an element, and is of the type **gml:CodeWithAuthorityType**. This is also a technical identifier which should be stable over time. INSPIRE recommends to use the **namespace** and **localId** from the **inspireId** to build the **identifier**, and INSPIRE **identifiers** use this codespace: <http://inspire.ec.europa.eu/ids>

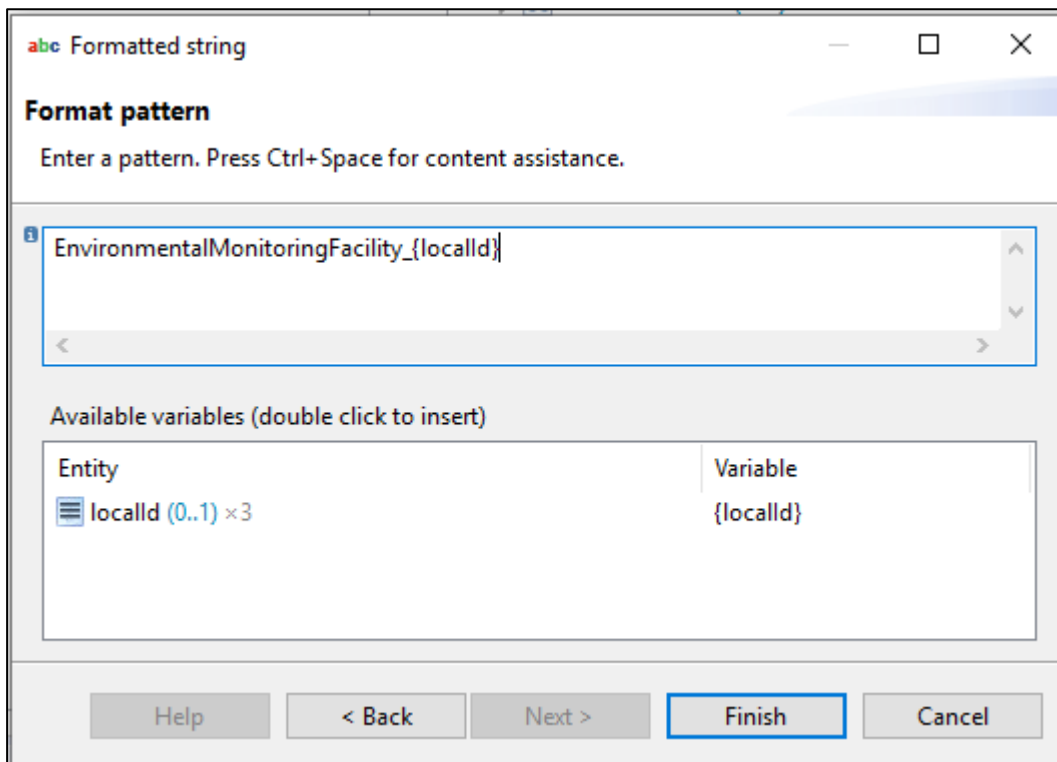
*Pozn:*

*Kromě těchto identifikátorů mohou v některých případech INSPIRE data obsahovat ještě tzv.*

*ThematicIdentifier ( thematicId), který většinou vzniká modifikací localId a je vyžadován pro reporting k určitým konkrétním účelům (např. prostorové objekty reportované do Water Framework Directive).*

Přiřazení hodnoty **localId** (resp. **loc\_code** v originálním zdrojovém shapefile *III\_7\_CHMU\_imisni\_merici\_stanice\_epsg3035*) k parametru **id** (= **gml\_id**):

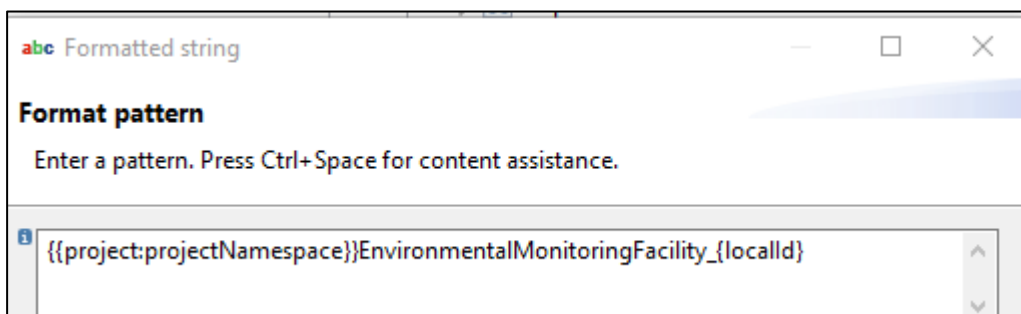
Source: **localId** (**loc\_code**) – target: **id** – mapovací funkce: **Formatted string** – Next – okno Format pattern – do rámečku vyplnit (best practice): feature type name hodnota **localId** (**loc\_code**), tj „**EnvironmentalMonitoringFacility\_{localId}**“, kde {**localId**} (resp. {**loc\_code**}) je “obyčejná” (tj. nikoliv dynamická) proměnná.



V target okně se u položky id objeví počet prvků („x 3“), v Alignment okně se přidá příslušná položka mapování.

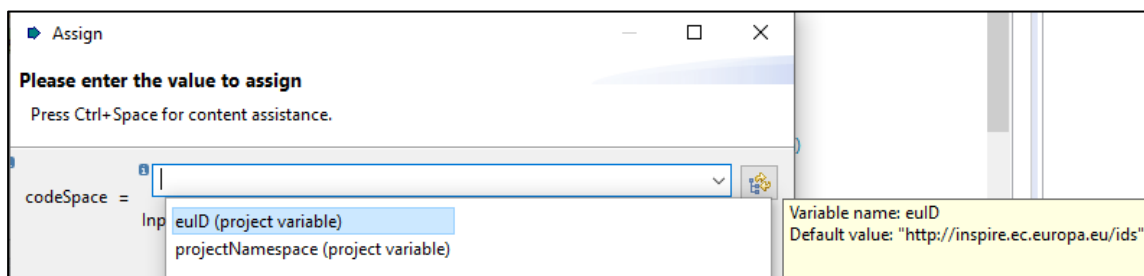
Přiřazení hodnoty **localId** (*loc\_code*) k parametru **identifier** (jedná se o parametr identifier který má vnořenou položku codeSpace, nikoliv o parametr Identifier který je sám vnořený pod parametrem inspireId):

Source: **localId** (*loc\_code*) – target: **identifier** - mapovací funkce: **Formatted string** – Next - okno Format pattern – kliknout do rámečku pro zadání hodnoty do rámečku - následně stisknout **Ctrl + mezerník** – objeví se nabídka využitelných **proměnných** (*localId* (resp. *loc\_code*) a obě na začátku definované dynamické proměnné) – dvojklikem vybrat proměnnou **projectNamespace** – přidáme feature type name a doplníme *localId* (*loc\_code*) -> výsledný řetězec: „**{{project:projectNamespace}}**EnvironmentalMonitoringFacility\_{localId}“ (resp. ...\_{loc\_code})

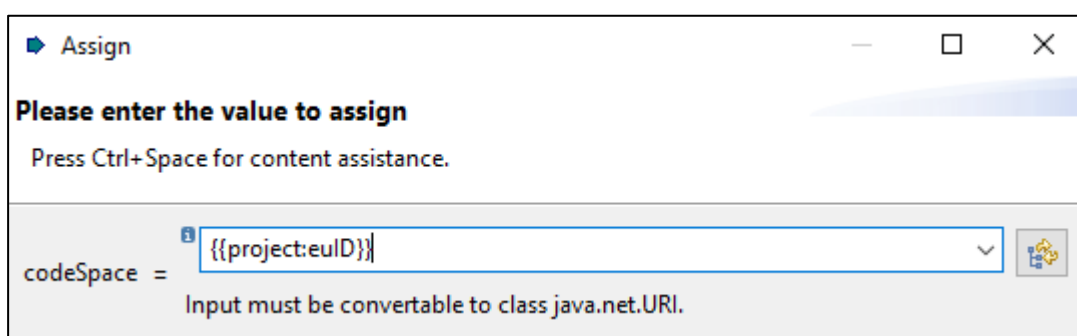


Přiřazení hodnoty k parametru **identifier/codeSpace**:  
využití dynamické proměnné **euID** nadefinované dříve

Target: **identifier/codeSpace** - mapovací funkce: **Assign** – Next – kliknout do rámečku pro zadání hodnoty „codespace =“, následně stisknout **Ctrl + mezerník** – objeví se nabídka dříve definovaných dynamických **proměnných**



– dvojklikem vybrat **euID** (v rámečku pro zadání hodnoty se objeví „**{{project:euID}}**“) – Finish  
 (=> *dynamické proměnné se zapisují ve výrazech v dvojitých složených závorkách, „obyčejné proměnné“ představující hodnotu atributů se zapisují v jednoduchých složených závorkách*)



V target okně se u identifier/codeSpace objeví počet prvků („x 3“), v Alignment okně se přidá příslušná položka mapování

Přiřazení hodnoty **inspireId** – skládá se z **localId** a **namespace** (a z nepovinné položky **versionId**):

- **localId**

Source: **localId (loc\_code)** - Target: **inspireId/Identifier/localId** - mapovací funkce: **Formatted string** – Next - okno Format pattern - do rámečku vyplnit „**EnvironmentalMonitoringFacility\_{localId}**“ (resp. *...\_{loc\_code}*)

- **namespace**

Target: **inspireId/Identifier/namespace** - mapovací funkce: **Assign** – Next – kliknout do rámečku pro zadání hodnoty „namespace =“, následně stisknout **Ctrl + mezerník** – objeví se nabídka dříve definovaných dynamických **proměnných** - dvojklikem vybrat **projectNamespace** (v rámečku pro zadání hodnoty se objeví „**{{project:projectNamespace}}**“) – Finish

## Přiřazení nil values parametrům, které jsou voidable a přitom mají násobnost 1 nebo 1..n

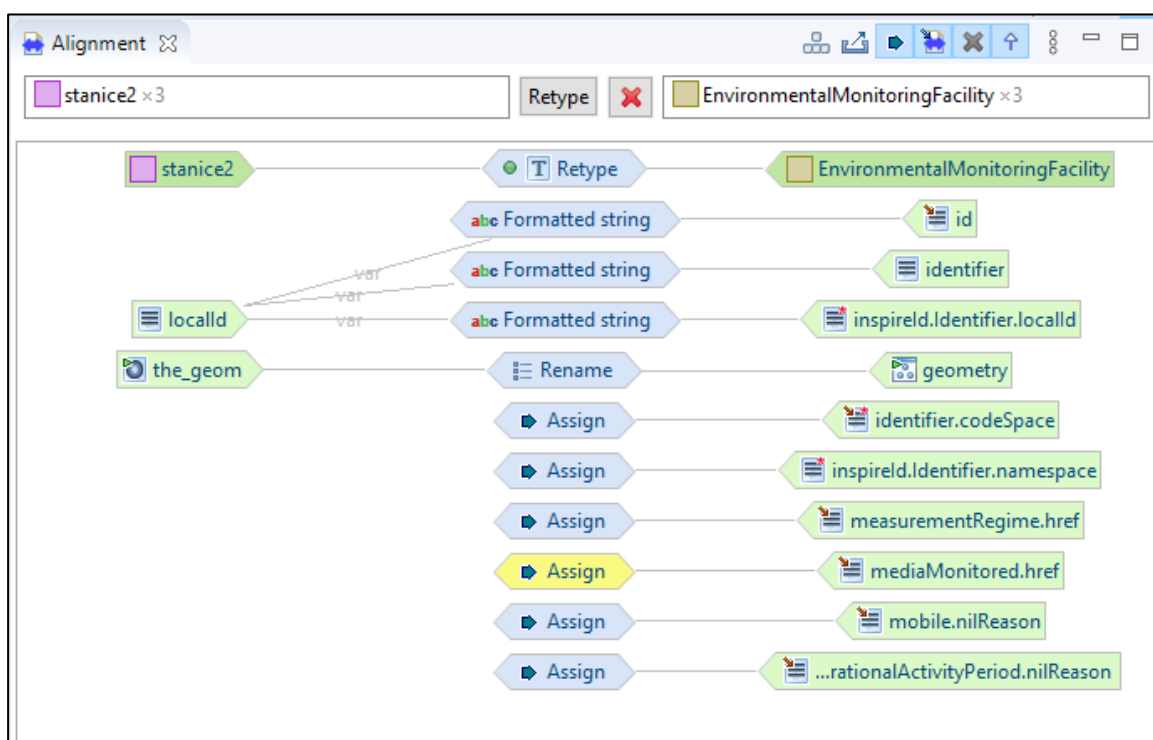
Target: **operationalActivityPeriod/nilReason** - mapovací funkce: **Assign** – Next – do rámečku pro zadání hodnoty vyplnit „**other:unpopulated**“ – Finish  
dtto pro **mobile/nilReason**

??? v budoucnu ověřit zdali by se spíše neměly používat hodnoty z číselníku

<https://inspire.ec.europa.eu/codelist/VoidReasonValue>

zde konkrétně <https://inspire.ec.europa.eu/codelist/VoidReasonValue/Unpopulated>

Okno Alignment po dokončení definice mapování:



## Transformace dat a tvorba výstupního souboru ve formátu GML

File – Export – Transformed data –

**okno Export format** – vybrat „GML (INSPIRE SpatialDataSet) [deprecated by INSPIRE]“ – Next

**okno Export destination** – Browse – zadat jméno souboru (např. stanice2\_GML1) - Next

*Lze doplnit Add validator – vybrat „XML validator (Java XML API)“*

- Next

**okno SpatialDataSet configuration** (v ukázce toto okno není):

vyplnit: localId: stanice2\_GML1

namespace:

[http://portal.chmi.cz/files/portal/docs/uoco/isko/tab\\_roc/2016\\_enh/pollution\\_locality/](http://portal.chmi.cz/files/portal/docs/uoco/isko/tab_roc/2016_enh/pollution_locality/)

(= fiktivní URL); **nezapomenout na koncové lomítko !!!**

Pozn:

*Program HALE studio umožní vložit do rámečku pro namespace i specifikaci proměnné (např. {{project:projectNamespace}}), ale v tomto případě je nutno dané URL přímo vypsat. Pokud se použije proměnná, do výsledného GML souboru se vloží její specifikace a nikoliv obsah.*

- Next

**okno Coordinate reference system:**

Rámeček Convert to CRS:

zaškrtnout “Convert all geometries to the given target CRS:“

Select...

**okno Please specify the CRS to use**

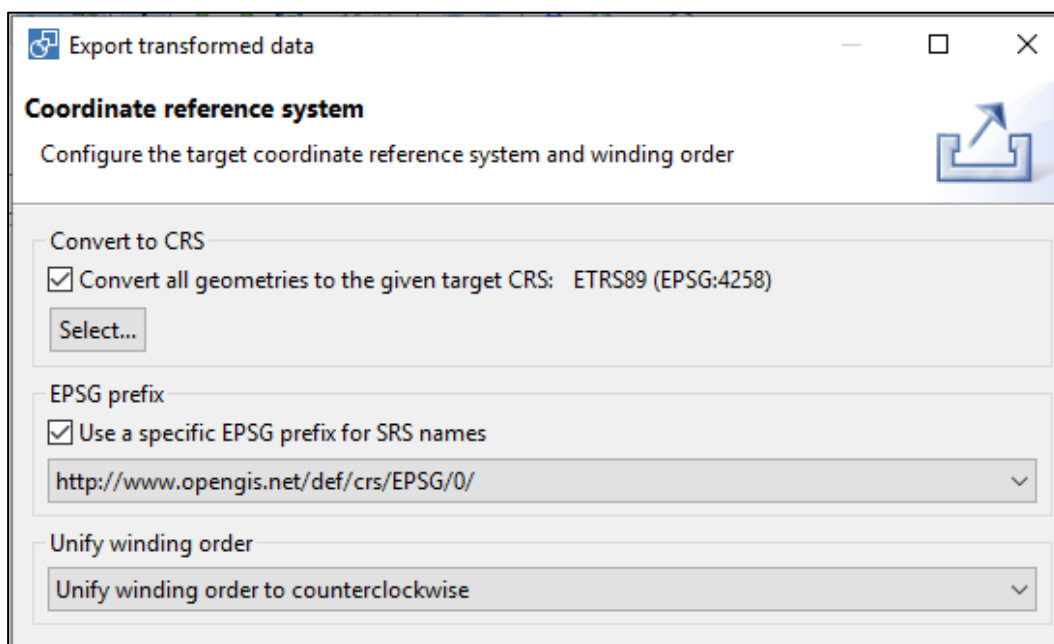
Zaškrtnout CRS code – v rámečku ponechat EPSG:4258

Rámeček EPSG prefix:

zaškrtnout Use a specific EPSG prefix for SRS names

v okénku níže zvolit <http://www.opengis.net/def/crs/EPSSG/0/>

*Pozn: v případě zpracování polygonových dat je třeba v tomto okně nastavit v rámečku Unify winding order variantu Unify winding order to clockwise.*



- Next

**okno Dataset feed creation:**

nevyplňovat nic

- Next

**okno XML/GML settings:**

zaškrtnout „Pretty print XML“

- Next

**okno XML/GML settings:**

nevyplňovat nic

Finish

Na kartě **Report list** (PD kvadrant okna aplikace HALE studio) zkontrolovat, zdali všechny kroky proběhly úspěšně (hale project export, GML file export, Project validator, XML validation)

Uložit projekt: File – Save Alignment Project (doporučeno ukládat průběžně)

Ověřit existenci výstupního GML souboru v zadaném adresáři

## Kontrola výstupního GML souboru

Výstupní GML soubor zkontrolovat např. v programu Notepad++ a v QGISu (se zobrazením GML souborů v ArcMapu jsem měl problémy)

Následně výstupní GML soubor ověřit v **INSPIRE validátoru** ( <https://inspire.ec.europa.eu/validator/> ) – dokud pro dané téma INSPIRE nebudou k dispozici kompletní testy, použít alespoň skupinu testů pro obecné části dat „**Interoperable data sets in GML (Guidelines for the Encoding of Spatial data version 3.3)**“

Tato skupina obsahuje 4 testy:

- Conformance class: Data consistency, General requirements
- Conformance class: INSPIRE GML application schemasm General requirements
- Conformance class: Information accessibility, General requirements
- Conformance class: reference systems, General requirements

Výsledek 1. a 2. testu může obsahovat výzvu k „ruční“ kontrole dat (oranžové podbarvení):

ETF Test Report

inspire.ec.europa.eu/validator/v2/TestRuns/EIDb49edd1b-7075-40ec-948c-43ed480b00e8.html

Test run on 15:21 - 21.07.2020 with test suite Conformance class:  
Data consistency, General requirements

Status Passed, manual checks required

	Total	Count	Skipped	Failed	Warnings	Manual
Test suites	2	0	0	0	0	1
Test cases	3	0	0	0	0	1
Assertions	7	0	0	0	0	2

Started 21/07/2020 13:21:55 GMT  
Duration 0.001 s

Show:  All  Only failed  Only manual

Level of detail:  All details  Less information  Simplified

- + Conformance class: INSPIRE GML encoding 1
- + Conformance class: Data consistency, General requirements 2

Report generated by ETF

3. a 4. test obvykle končí bez jakékoliv připomínky:

ETF Test Report

inspire.ec.europa.eu/validator/v2/TestRuns/EID68a4c969-c5d5-468e-89be-95854af46d46.html

Test run on 15:30 - 21.07.2020 with test suite Conformance class:  
Information accessibility, General requirements

Status Passed

	Total	Count	Skipped	Failed	Warnings	Manual
Test suites	2	0	0	0	0	0
Test cases	2	0	0	0	0	0
Assertions	4	0	0	0	0	0

Started 21/07/2020 13:31:07 GMT  
Duration 0.489 s

Show:  All  Only failed  Only manual

Level of detail:  All details  Less information  Simplified

- + Conformance class: INSPIRE GML encoding 1
- + Conformance class: Information accessibility, General requirements 1

Report generated by ETF