

INSPIRE MIG Action 2017.3

Deliverable 2.2

Analysis of the usability of INSPIRE data

Author: Stefania Morrone (Epsilon Italia)

Reference: D2.2

Title: Analysis of the usability of INSPIRE data

Type of deliverable: Chapter in the 2017.3 Technical Report

Version: 20190924

Table of contents:

Introduction.....	2
1. The testing environment.....	2
1.1. The test data.....	2
1.2. The test suites	3
1.3. The tested tools.....	3
1.4. The study documentation	3
2. Analysis of the collected issues	6
2.1. Issues related to management of INSPIRE data in the GML encoding	7
2.2. Issues related to management of INSPIRE data in the GeoJSON encoding	10
3. Main outcomes.....	11
Annex I.....	12

Introduction

The number of geospatial datasets available through INSPIRE Spatial Data Infrastructures is rapidly increasing and this goes hand in hand with the request for more user-friendly and ready-to-use solutions for efficient consumption of the INSPIRE data. Aiming to keep up with the coming demands and ensure that the INSPIRE infrastructure remains fit for purpose, the INSPIRE Maintenance and Implementation Group (MIG) has put forward Action 2017.2¹ and Action 2017.3² respectively addressing the investigations for alternative INSPIRE data encodings and the improvement of client support for INSPIRE data.

The present document, produced in the context of the Action 2017.3, reports on the study related to the usability of INSPIRE data, either provided in default GML encoding or alternative GeoJSON encoding, in most used libraries, desktop / web clients and ETL tools for data processing.

The findings of the study are being discussed with JRC and MIG representatives as well as with open source community and commercial vendors in order to agree on a list of priority issues and identify a way forward to improve the client support for INSPIRE data.

Deliverable structure:

- **Section 1** describes the testing environment: test data, test suites, list of the tested clients;
- **Section 2** provides an analysis of the test results and the list of major identified issues per client tool;
- **Section 3** presents main conclusions and proposes a prioritisation of the issues and possible next steps.

1. The testing environment

1.1. The test data

To better support the works related to the alternative encodings for INSPIRE data and in agreement with the MIG 2017.2 group, the study focused on test data sets:

- belonging to ‘Addresses’ and ‘Environmental Monitoring Facilities’ data themes, for which specific GeoJSON Encoding Rules for INSPIRE have been developed in the context of the MIG Action 2017.2;
- provided in the default GML encoding or conformant to relevant GeoJSON Encoding Rule for INSPIRE;
- available from the *GitHub repository for MIG Action 2017.2*³ or from the *INSPIRE Geoportal*⁴.

¹

<https://webgate.ec.europa.eu/fpfis/wikis/display/InspireMIG/Action+2017.2+on+alternative+encodings+for+INSPIRE+data>

² <https://ies-svn.jrc.ec.europa.eu/projects/2017-3/wiki>

³ <https://github.com/INSPIRE-MIF/2017.2>

⁴ <https://inspire-geoportal.ec.europa.eu/>

When retained appropriate, so-obtained data sets were modified in order to represent additional complexity e.g. adding a second geometry for the same address object.

When it was not possible to find suitable AD / EF data for specific testing purposes, data from different INSPIRE data themes was considered. This was the case, for example, of the Protected Sites data set (from the INSPIRE Geoportal) used to test client behaviour with mixed geometries and of the AU web service (link available from the GitHub 2017.2 repository) used to test clients' connections to WFS 2.0 serving GeoJSON data.

The details on the data used for each test can be found in the test suites documents described in below Section 1.2.

1.2. The test suites

The identification and the characterization of the tests have built on the technical issues identified during the classification of available information on the usability of INSPIRE data (also part of the activities related to Action 2017.3 and reported in the deliverable D2.1) and on the contractor's experience.

The tests focus on the client capacity to:

- load data (from local files and web services);
- handle big files;
- display the data maintaining the original structure and without loss of information;
- make use of complex properties for styling, processing and filtering;
- display multiple values per property;
- resolve references to other features (in external resources or in the same data set);
- display different geometry types or multiple geometries per feature type.

A '*GML Test Suite*' and a '*GeoJSON Test Suite*' have been created to document the tests used to verify the behaviour of client tools with INSPIRE data, respectively in the GML default encoding and the in GeoJSON alternative encoding.

In the Test Suites, each test is identified by a specific '**Test ID**' and detailed through the '**Test purpose**', the '**Test data**' (also containing a direct link to the data used for testing) and the '**Test expected outcomes**'.

More details and the links to the documents are provided in below Section 1.4.

1.3. The tested tools

The tests were performed on the following client tools:

QGIS v3.4 - Grass v7.7- ArcGIS Pro v2.3 - ArcMap v10.7 - ArcGIS Online v2018/2019 – OpenLayers v5.3.0 – Leaflet v1.4 – FME 2018 and hale studio v3.4.

1.4. The study documentation

The GitHub platform is the tool designated to document all activities related to the study and to share and discuss the work done with the MIG 2017.2 group, the software providers and all the possible stakeholders.

The public ‘CanIUse INSPIRE’ GitHub repository⁵ makes available:

- the documentation related to the GML ⁶ and the GeoJSON⁷ tests;
- the test data⁸
- the results of the performed tests⁹.

The software providers and the INSPIRE experts are this way allowed to provide feedbacks and contributions - e.g. on possible solutions – in GitHub issue tickets.

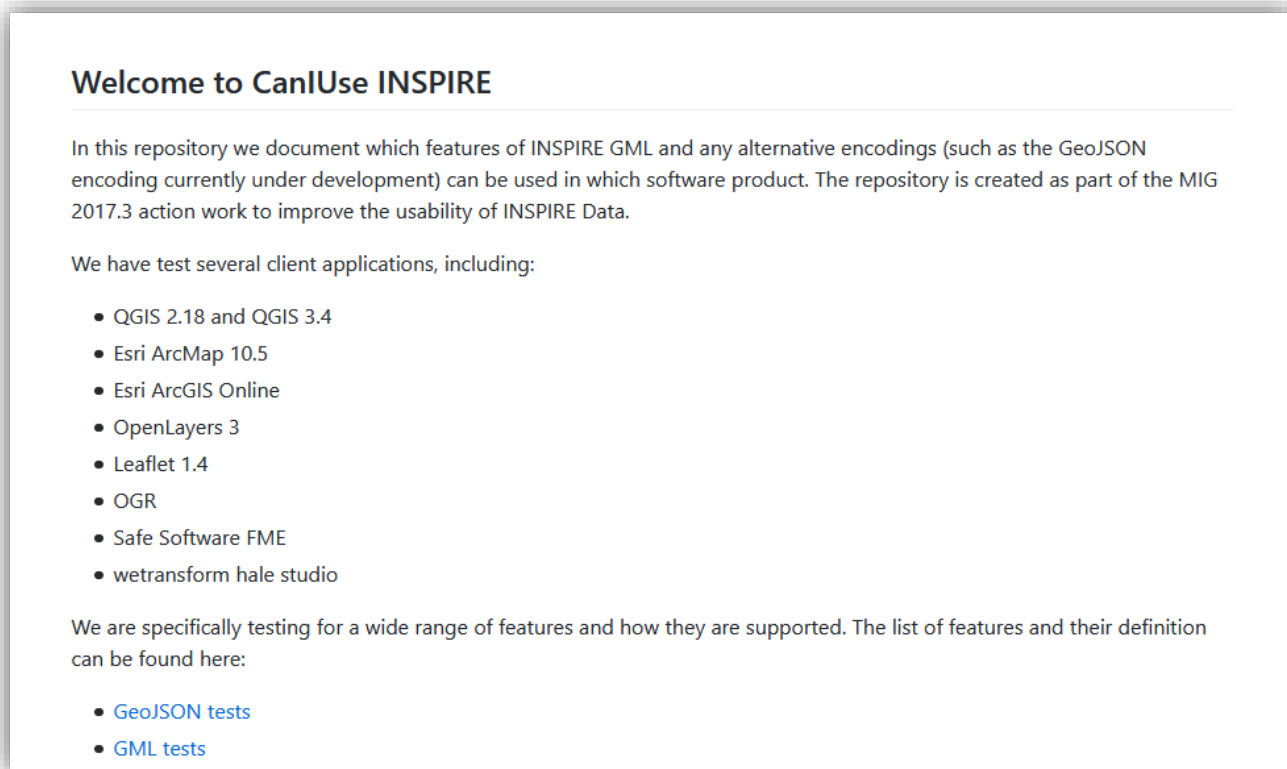


Figure 1 CanIUse INSPIRE welcome page

A ‘GitHub Pages’¹⁰ project site ¹¹ has been connected to the ‘CanIUse INSPIRE’ repository so that all project documentation have been published.

In particular, the test results¹² are made available in the very intuitive form of ‘CanIUse’¹³ tables (see Figure 2 below). The aim is to make straightaway clear which features of INSPIRE GML and GeoJSON encodings can be used in which software product.

In those ‘CanIUse’ tables, the green cells indicate full support for a specific feature, the yellow ones partial support and the red cells mean no support.

Note: in order to visualise the results in the ‘CanIUse’ table form, the test results in the GitHub repository are provided as json files.

⁵ <https://github.com/INSPIRE-MIF/caniuse/tree/master>

⁶ <https://inspire-mif.github.io/caniuse/docs/gml.html>

⁷ <https://inspire-mif.github.io/caniuse/docs/geoJSON.html>

⁸ <https://github.com/INSPIRE-MIF/caniuse/tree/master/testcases>

⁹ <https://github.com/INSPIRE-MIF/caniuse/tree/master/results>

¹⁰ <https://help.github.com/en/articles/about-github-pages>

¹¹ <https://inspire-mif.github.io/caniuse/>

¹² <https://inspire-mif.github.io/caniuse/generator/out.html>

¹³ https://caniuse.com/#info_about

Search for features..

1. Geography Markup Language (GML)

This section contains information on usability tests with INSPIRE GML data. The source data and overview of the individual test suits is available on [GitHub](#).

gml_file_load

QGIS	ArcGIS Online	hale studio	OpenLayers	FME Desktop	Leaflet
3.4.4-Madeira	December 2018 update	3.4.1	5.3.0	2018.1	1.4.0

gml_file_display

QGIS	ArcGIS Online	hale studio	OpenLayers	FME Desktop	Leaflet
3.4.4-Madeira	December 2018 update	3.4.1	5.3.0	2018.1	1.4.0

gml_WFS2_load

QGIS	ArcGIS Online	hale studio	OpenLayers	FME Desktop	Leaflet
3.4.4-Madeira	December 2018 update	3.4.1	5.3.0	2018.1	1.4.0

gml_WFS2_display

Figure 2 Canluse INSPIRE - results

2. Analysis of the collected issues

The test results for each client software have been thoroughly analysed and the issues found have been reported in the 'Client SW Issues' Google spreadsheet¹⁴, shared with the MIG group and the software providers.

This living excel file constitutes a common reference for discussions, prioritisation of issues and agreement on next steps.

Two different sheets report the issues experienced respectively with the GML and GeoJSON encoding and provide the current possible workarounds /known possible solutions (if any) to overcome them. Moreover, all the issues have been assigned an 'INSPIRE priority' (High, Medium, Low) that indicates the level of importance their fixing has for the INSPIRE community.

The software providers are asked to contribute to this dynamic document with their feedbacks regarding the importance that the resolution of the issue has for them, the difficulty of the relevant fixing and an estimation of the needed workload.

Those pieces of information will allow the identification of the most significant unresolved issues that have the higher possibility to be solved as they are also significant from the SW point of view / are easy to solve.

Taking into consideration the discussions in the face-to-face meeting in ISPRA on July 8th and 9th 2019, the excel file also proposes possible next steps that the software providers are asked to comment (e.g. are these steps indeed feasible in the short/medium run?) and update when they have any news.

The excel structure is described in Annex I.

Client Tool	Failing Test ID	Current possible solutions	Notes	INSPIRE Priority (H,M,L)	SW Product Priority (H,M,L)	Difficulty of Solution (H,M,L)	Next steps SW Product
QGIS v3.4	gml_complex_property_load	Set 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' OR use the GML Application Schema Toolbox	Useful instructions on how it is currently possible to set 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' can be found at https://github.com/INSPIRE-MIF/canluse/blob/master/docs/gml.md NOTE: only the GML Application Schema Toolbox is able to render the original INSPIRE structure, using the 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' the data structure is flattened.	H			Evaluate whether it is safe to overwrite the default behavior of QGIS and set 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' as default see https://github.com/qgis/QGIS/issues/29641
QGIS v3.4	gml_complex_property_display	Set 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' OR use the GML Application Schema Toolbox	Same as above	H			Same as above
QGIS v3.4	gml_complex_property_style	Set 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES'	Same as above	M			Same as above
QGIS v3.4	gml_complex_property_processing	Set 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES'	Same as above	M			Same as above
QGIS v3.4	gml_WFS2_load	Use the GML Application Schema Toolbox	Connection to WFS 2.0 is available also without using the GML Application schema plugin, but in that case, it is not possible to download INSPIRE complex features	H			Investigate the feasibility of proposing the GML Application Schema Toolbox moving to QGIS Core via a QEP
QGIS v3.4	gml_WFS2_display	Use the GML Application Schema Toolbox	Even using the plugin, however, the 'Filter by extent' is not working - for more details see issue https://github.com/INSPIRE-MIF/canluse/issues/15 and https://issues.qgis.org/issues/21576	M			Same as above

Figure 3 Client SW Issues Google spreadsheet

¹⁴https://docs.google.com/spreadsheets/d/1nHW2QCTDyClf2_5fdtHiwwAMuFVwyQt/edit?pli=1#gid=22364322

Major issues with the consumption of INSPIRE data.

The two tables below provide an overview of major issues experienced with the consumption of INSPIRE data, respectively in GML and GeoJSON encoding, grouped by client software tool and affected feature.

Detailed description of the issues can be found in the above mentioned 'Client SW Issues' excel file.

2.1. Issues related to management of INSPIRE data in the GML encoding

Client	Issues
<p>QGIS v3.4</p>	<p><u>Management of Complex Properties:</u></p> <ul style="list-style-type: none"> • Visualisation of complex properties is possible only using the 'GML Application Schema toolbox' plugin OR setting the 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' (https://github.com/qgis/QGIS/issues/29641) • When using the 'GML Application Schema toolbox' plugin, styling and processing are not supported • When setting the 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' the sub-elements are all visible but the data structure is flattened. <p><u>Management of 'xlink' references:</u></p> <ul style="list-style-type: none"> • Visualisation of 'xlink' attributes is possible only using the 'GML Application Schema toolbox' plugin OR setting the 'GML_ATTRIBUTES_TO_OGR_FIELDS=YES' • It is possible to resolve only the external links - the internal links (#) are NOT resolved. It is worth noting that using the GMLAS plugin (external) links can be resolved as 'embedded' description or 'new layer'. <p><u>Management of Multiple occurrences:</u></p> <ul style="list-style-type: none"> • Using GML_ATTRIBUTES_TO_OGR_FIELDS option, multiple occurrences of a property (e.g. the different <ad:components> related to a specific address) are displayed as elements of an array and there are limitations in the number of elements that can be displayed. <p><u>Management of Geometries:</u></p>

	<ul style="list-style-type: none"> • Multiple occurrences of the same geometry property (e.g. <ad:position>) are displayed only using the specific <i>GML Application Schema toolbox</i> plugin options '<i>Load in relational mode (GMLAS)</i>' or '<i>Load from GMLAS database</i>'. • Mixed geometries (e.g. polygons and points in same dataset) are not displayed when using the '<i>GML Application Schema Toolbox</i>' plugin • 3D geometries display correctly only if both '<i>srsName</i>' and '<i>srsDimension</i>' attributes are set. <p><u>Download of INSPIRE complex features from WFS 2.0</u></p> <ul style="list-style-type: none"> • needs the '<i>GML Application Schema toolbox</i>' plugin • 'Filter by extent' option loads no data. The following error is returned '<Crs' object has no attribute 'split'>' - see https://issues.qgis.org/issues/21575 <p><u>Partial support for feature editing/creation</u></p> <ul style="list-style-type: none"> • Not possible to operate directly on GML dataset
<p>ArcGIS Pro v2.3 - ArcMap v10.7</p>	<p><u>Management of Complex Properties</u></p> <ul style="list-style-type: none"> • Visualisation of INSPIRE GML data (either from files or WFS 2.0) needs Data Interoperability Extension (DIE) • Displayed structure is flat: the complexity of the original data model is not rendered and this makes displayed attributes difficult to understand <p><u>Management of 'xlink' references:</u></p> <ul style="list-style-type: none"> • It is possible to resolve only the external links - the internal links (#) are NOT resolved. It is worth noting that using the GMLAS plugin (external) links can be resolved as 'embedded' description or 'new layer'. <p><u>Partial support for feature editing/creation</u></p> <ul style="list-style-type: none"> • Not possible to operate directly on GML dataset <p><u>Tests to be completed by ESRI</u></p>

<p><i>ArcGIS Online</i></p>	<p>GML encoding not supported</p>
<p><i>Leaflet</i></p>	<p>GML encoding not supported</p>
<p><i>OpenLayers</i></p>	<p><u>Management of Complex Properties</u> Not supported. However, interesting step forward in upcoming v5.3.x https://github.com/openlayers/openlayers/pull/8519 .</p>
<p><i>hale Studio</i></p>	<p><u>Management of Geometries:</u></p> <ul style="list-style-type: none"> • Nested geometries, such in the case of the AD theme, are not displayed on the map in the “Map perspective” <p><u>Management of ‘xlink’ references:</u></p> <ul style="list-style-type: none"> • ‘xlink’ references cannot be resolved <p>Styling and processing of the properties</p> <ul style="list-style-type: none"> • not supported <p><u>Partial support for feature editing/creation</u></p> <ul style="list-style-type: none"> • Not possible to operate directly on source GML dataset (ETL needed)
<p><i>FME</i></p>	<p><u>Management of ‘xlink’ references:</u></p> <ul style="list-style-type: none"> • ‘xlink’ references cannot be resolved <p><u>Management of Complex Properties</u></p> <ul style="list-style-type: none"> • Displayed structure is flat: the complexity of the original data model is not rendered and this makes displayed attributes difficult to understand <p><u>Partial support for feature editing/creation</u></p> <ul style="list-style-type: none"> • Not possible to operate directly on source GML dataset (ETL needed)

2.2. Issues related to management of INSPIRE data in the GeoJSON encoding

Client	Issues
<p>QGIS v3.4</p>	<p><u>Management of array objects</u></p> <ul style="list-style-type: none"> • There is a limitation in the number of displayed elements • Styling and processing of single elements is not supported <p><u>Management of Geometries</u></p> <ul style="list-style-type: none"> • Geometry Collections not handled (nothing displayed on the map) <p><u>Management of multiple feature types (in the same file):</u> If the GeoJSON file contains spatial objects from different feature types (e.g. environmental monitoring facility and its observation results), the properties from the different features are merged i.e. all features are displayed with the same structure, resulting from a merge of the properties of all the single features.</p>
<p>ArcMap v10.7</p>	<p><u>Management of multiple feature types (in the same file):</u> If the GeoJSON file contains spatial objects from different feature types (e.g. environmental monitoring facility and its observation results), the properties from the different features are merged i.e. all features are displayed with the same structure, resulting from a merge of the properties of all the single features.</p> <p>Tests to be completed by ESRI</p>
<p>ArcGIS Pro v2.3</p>	<p><u>Management of multiple feature types (in the same file):</u> If the GeoJSON file contains spatial objects from different feature types (e.g. environmental monitoring facility and its observation results), the properties from the different features are merged i.e. all features are displayed with the same structure, resulting from a merge of the properties of all the single features.</p>

	Tests to be completed by ESRI
<i>ArcGIS Online 2019</i>	Limitations on the size of the files that can be loaded
<i>hale Studio</i>	<ul style="list-style-type: none"> • GeoJSON data import not supported • GeoJSON output partially supported
<i>OpenLayers</i>	<p><u>Management of multiple feature types (in the same file):</u></p> <ul style="list-style-type: none"> • If the GeoJSON file contains spatial objects from different feature types having overlapping geometries (e.g. environmental monitoring facility and its observation results), OpenLayers can show properties from only one feature. Custom logic could line up the properties of the different features in the info popup.
<i>Leaflet</i>	<p><u>Management of multiple feature types (in the same file):</u></p> <ul style="list-style-type: none"> • If the GeoJSON file contains spatial objects from different feature types having overlapping geometries (e.g. environmental monitoring facility and its observation results), Leaflet can show properties from only one feature. Custom logic could line up the properties of the different features in the info popup

3. Main outcomes

The results of the study confirm that all the most common client applications and tools have difficulties in consuming and/or fully making use of data shared according to INSPIRE xml schemas (and therefore reflecting all the complex structures present in the relevant conceptual model). Indeed, one of the main issues related to INSPIRE data in GML default encoding regards the

management of complex properties: even when the property displays without loss of information, the client operates a flattening of the original structure that affects the readability of the data. Another common issue affecting all the examined clients is the inability to resolve ('xlink') references: while external (https) references generally show up in a browser page, it is not possible to resolve internal links i.e. references to other features in the same data sets (e.g. the 'components' of the address in the AD data).

Likewise, even when the data is in GeoJSON encoding, the management of multiple feature types in the same file and the references between different features are challenging for all the tested client SW tools, as demonstrated by the test results for the 'EF' datasets conformant to the GeoJSON Encoding Rule for INSPIRE Environmental Monitoring Facilities.

The performed tests show that the desktop clients provide better GML data handling whilst GML encoding has poor support by JS tools. Almost all the client tools provide good support for the GeoJSON encoding.

Annex I

Find below a description of the fields in the Client SW tool excel

<i>Excel field</i>	<i>Description</i>
Client Tool	Name and version of Client SW under test
Failing Test ID	'Test ID' uniquely identifying the test in the relevant Test Suite
Current possible solutions	known workaround / possible tips to overcome the issue
Notes	Notes for a better delineation of the issue
INSPIRE Priority (H,M,L)	How important is the resolution of the issue for INSPIRE community? (High, Medium, Low importance)
SW Product Priority (H,M,L)	How important is the resolution of the issue for the SW provider? (High, Medium, Low importance)
Difficulty of Solution (H,M,L)	How difficult is the resolution of the issue ? (High, Medium, Low difficulty)
Next steps SW Product	what could be next steps to solve the problem? How much effort is required?