Title	Alternative encodings for INSPIRE data			
ID	MIWP-2017.2			
Status	☐ Proposed	☒ Endorsed	☒ In Progress	☐ Completed
Date of last update	2017-12-15			
Issue	The current data specification TGs define (complex) xml schemas based on GML as the default encoding for all INSPIRE spatial data themes. Many existing (web and desktop) applications and tools have difficulties in consuming and/or fully making use of data shared according to these schemas.			
	The INSPIRE xml schemas are complex, because they are generated automatically from the conceptual UML model (according to the normative UML-to-GML encoding rules described in the GML standard and INSPIRE Technical Guidelines *D2.7 Encoding Guidelines*) and therefore reflect all the complex structures present in the conceptual model. In contrast, most existing clients, including the popular GDAL/OGR open source library (that is underlying most OS and proprietary client solutions) consumes and writes flat data structures, where e.g. each attribute can only have at most one value and attributes can have only simple types (e.g. integer, string, boolean) and not complex ones. This means that, while INSPIRE data encoded according to the current schemas can be downloaded and viewed, simple use (visualisation, simple joins, visual overlays, spatial search, …) is difficult in standard GIS clients.			
	One way to address this gap is to create alternative encodings for basic data exchange and direct visualisation in standard GI tools[1].			
	Proposals already exist for alternative encodings, mainly for simplified XML schemas[2], but also for RDF vocabularies[3] or ESRI geodatabase[4] structures. Alternative encodings could also be based on other standards such as JSON or GeoPackage.			
	According to Art. 7 of the IRs on data interoperability, alternative encodings can be used as long as an encoding rule is publicly available that specifies schema conversion rules for all spatial object types and all attributes and association roles and the output data structure used. Such an encoding rule would need to include cross-cutting aspects (e.g. how to flatten recurring complex structures such as geographical names) as well as theme-specific aspects.			
Proposed action	The action shall define alternative encoding rules (mainly for the purpose of viewing/analysis in mainstream GIS systems) for a number of selected application schemas and a template and procedure for proposing and endorsing additional encoding rules in the future.			
Link to REFIT evaluation	Specific proposed action to "assist the Member States in applying and implementing the INSPIRE Directive (simplification of use), e.g. by the use of common tools, and promote			

---

[1] Another one is to encourage better support for consuming INSPIRE data by vendors – this solution will be investigated in action 2017.3.

[2] This approach is often also referred to as "flattening" of the existing xml schemas and has been applied e.g. in the ELF project or the example presented by DK in the 2nd 2016.1 meeting) and is already implemented in tools (e.g. ShapeChange).

[3] Draft guidelines have been developed in the ARE3NA ISA action – see https://github.com/inspire-eu-rdf/inspire-rdf-guidelines/blob/master/README.md

[4] http://server.arcgis.com/en/inspire/latest/get-started/geodatabase-template.htm

	priority setting together with the Member States " (page 12 of COM(2016)478).
Links & dependencies	Dependencies: • 2016.4: Discussions in the Thematic Clusters on use cases and requirements for simplification of xml schemas Links: • 2016.3: Validation and conformity testing – additional encoding rules will require additional abstract and executable test suites to be developed and potentially an extension of the INSPIRE test framework. • 2017.3: The action should be carried out in close collaboration with the action to encourage better support for consuming INSPIRE data by vendors
Organisational set-up	The work will be carried out by a temporary MIG sub-group, coordinated by JRC and supported by a contractor (e.g. for tasks 1c and 2). The temporary sub-group will carry out is work in several meetings during 2018 and via electronic exchange. The first meeting will be scheduled in Feb/March 2018.
Lead	JRC
Scope	This action will only address complexities that is introduced by the encoding. Discussions about changes in the conceptual models are out of scope.

Tasks	1) Develop concrete proposals for **alternative encodings**
	    a. In collaboration with thematic communities (through the Thematic Clusters platform and MIWP-14 sub-group), collect proposals for **alternative encodings**. These can be based on existing examples and/or on specific use cases and requirements. The proposals can be cross-cutting (i.e. cover all INSPIRE themes) or specific for one or several related themes.
	    b. In agreement with the MIG-T, prioritise the collected examples and select a small number of proposals, for which alternative encodings will be developed by the action.
	    c. For the selected examples, develop encoding rules that explain how (and/or under which conditions) the proposed encoding meets the requirements of the IRs[5].
	2) Define a **template and procedure for** proposing and endorsing **additional encoding rules**
	    a. Based on the work on (1a), elaborate a template for proposals for additional encoding rules. The template should cover the actual encoding rule (including possible approaches for explicitly documenting mappings in UML), but also target use cases, expected benefits, known limitations and tools for conformity testing & validation.
	    b. Develop a procedure for how additional encoding rules can be proposed using the template and checked/endorsed by the MIG, and how these are referred to from the existing data specification TGs.
	    c. Develop a proposal for the update or maintenance of endorsed encoding

---

[5] The action should also consider the potential loss of information for certain encodings and discuss whether such "lossy" encodings are meeting the IR requirements.

	rules.
	d. Create a repository of addition encoding rules that have been endorsed by the MIG.
	e. Develop a proposal for documenting the used encoding rules in data set/service metadata.
	3) Give recommendations on an update of D2.7 Encoding Guidelines (for future work)
Outcomes	• Repository of additional encodings • Template for proposing additional encoding rules • Procedure for endorsing and maintaining additional encoding rules • Proposal for documenting the used encoding rules in data set/service metadata
Proposed Impact	☐ Technical Adjustment / Bug Fixing ☐ Technical Improvement / Development ☒ Practical Support for Implementing Process ☒ Cost Reducing Effect for Implementing Process ☐ Direct Support on Policy-Making / - Activities
Timeline	Date of kick-off: January 2018
	Proposed Date of Completion: 31/12/2018
Required human resources and expertise	The members of the temporary sub-group should have expertise in one or several of the following areas: • Data modelling & specification of encoding rules • Implementation of data transformation processes • Implementing Rules and Technical Guidelines for data interoperability
Required financial resources	• Meeting reimbursement • Expert contract(s) for supporting tasks 1c and 2

Risk factors	Overall risk level of the action	Risk factors to be considered
	☐ High	☐ Missing Resources
	☒ Medium	☒ High Complexity
	☐ Low	☐ Interdependencies with other Actions
		Others: The risk of decreasing technical interoperability needs to be mitigated by ensuring proper documentation of additional encodings (in documentation and metadata)

Possible funding	• DG ENV funding (through Administrative Arrangement) • MS funding / in-kind contributions