

Use of the INSPIRE Reference Validator in 2019 Monitoring: process & lessons learned

Marco Minghini, Lukasz Ziemba, Fabio Vinci, Robert Tomas, Davide Artasensi, Daniele Francioli, Emanuela Epure

INSPIRE Geoportal Workshop - Status & Monitoring 2019 - June 4, 2020





Indicators MDi1.1 & MDi1.2 are calculated using the INSPIRE Validator:

Article 4

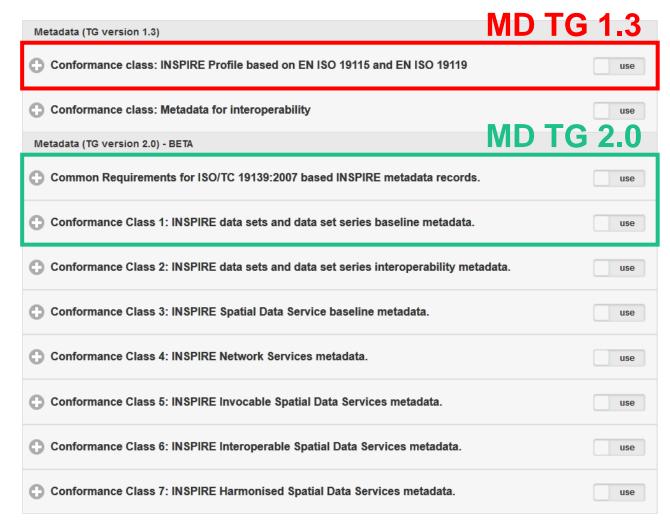
Monitoring of the conformity of metadata with Regulation (EC) No 1205/2008

The following indicators shall be used to measure the percentage of metadata for spatial data sets and spatial data services that are published by Member States through the discovery services referred to in Article 11(1)(a) of Directive 2007/2/EC that are in conformity with Commission Regulation (EC) No 1205/2008⁷ as regards metadata:

- (a) the number of spatial data sets for which metadata are in conformity with Regulation (EC) No 1205/2008 multiplied by a hundred and divided by the number of spatial data sets for which metadata exist as given by indicator "DSi1,1" ("MDi1,1");
- (b) the number of spatial data services for which metadata are in conformity with Regulation (EC) No 1205/2008 multiplied by a hundred and divided by the number of spatial data services for which metadata exist as given by indicator "DSi1,2" ("MDi1,2").

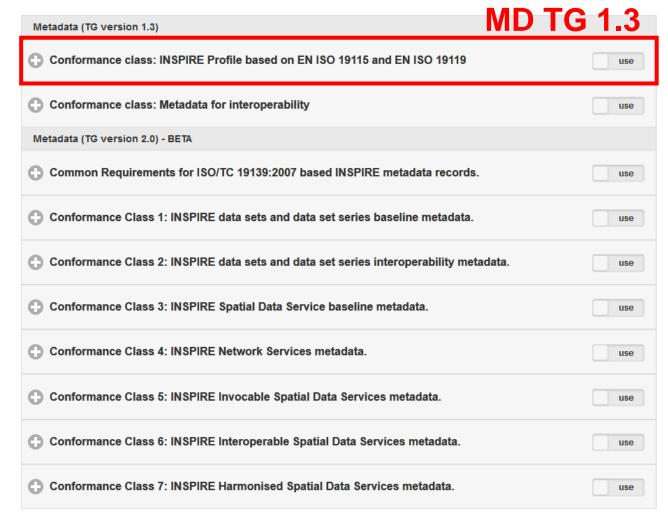


- Indicator MDi1.1:
 - percentage of metadata for spatial data sets conformant with Commission Regulation (EC) No 1205/2008 as regards metadata
 - metadata (TG 1.3 & 2.0)
 filtered as:
 - <gmd:hierarchyLevel>/<gmd:MD_ScopeCode> with codelistvalue dataset or series



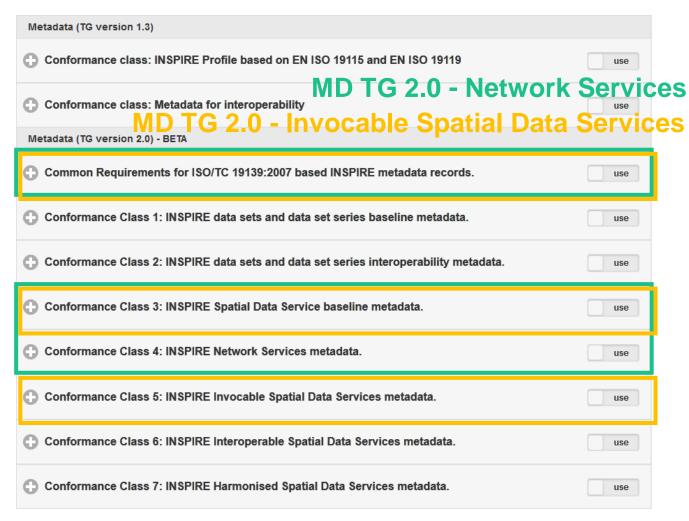


- Indicator MDi1.2:
 - percentage of metadata for spatial data services conformant with Commission Regulation (EC) No 1205/2008 as regards metadata
 - metadata (TG 1.3) filtered as:
 - <gmd:hierarchyLevel>/<gmd
 :MD_ScopeCode> with codelist
 value service

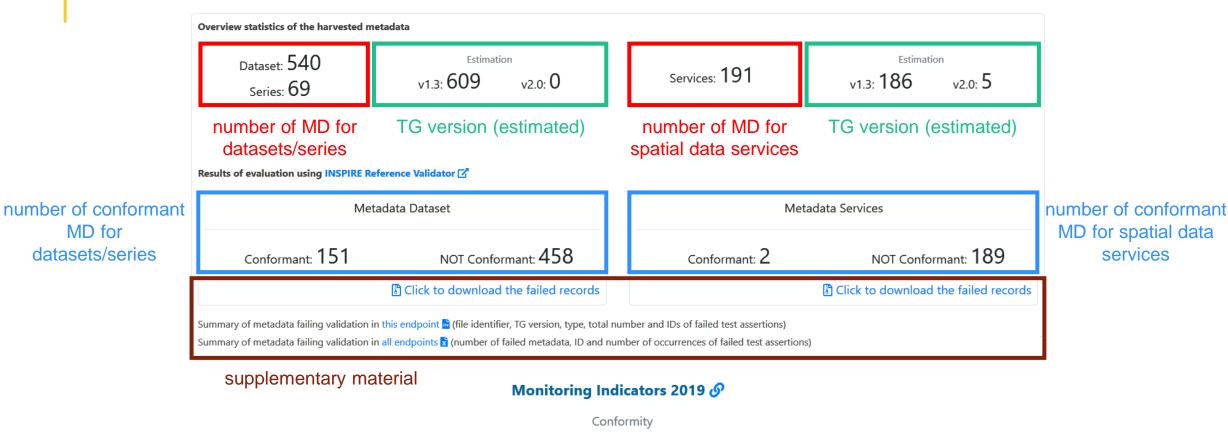




- Indicator MDi1.2:
 - percentage of metadata for spatial data services conformant with Commission Regulation (EC) No 1205/2008 as regards metadata
 - metadata (TG 2.0) filtered as:
 - <gmd:hierarchyLevel>/<gmd:MD_ScopeCode> with codelistvalue service
 - <srv:serviceType>/<gco:Lo
 calName> with value view,
 download, discovery or
 transformation for Network
 Services, other for Invocable
 Spatial Data Services







Monitoring of the conformity of metadata 27.5% MDi1 1 1.1% MDi12

indicator values



MD for

datasets/series

Filter for the estimation of MD TG version

- The filter is based on the only element that is different in the XML encoding between TG 1.3 and 2.0, i.e. <gmd:useLimitation> (required for the MD element "Conditions applying to access and use" in TG 1.3).
- But sometimes the <gmd:useLimitation> is ALSO used in MD TG 2.0!

```
<gmd:resourceConstraints>
                                                                                                                                                 limitations on public access
                 <gmd:MD LegalConstraints>
                   <gmd:accessConstraints>
                       <gmd:MD RestrictionCode codeList="http://standards.iso.org/iso/19139/resources/gmxCodelists.xml#MD RestrictionCode" codeListValue="otherRestrictions"/>
                    </gmd:accessConstraints>
                    <gmd:otherConstraints xsi:type="gmd:PT FreeText PropertyType">
                      <gmx:Anchor xlink:href="http://inspire.ec.europa.eu/metadata-codelist/LimitationsOnPublicAccess/noLimitations">Pas de restriction d'accès public selon INSPIRE/gmx:Anchor>
                     </gmd:otherConstraints>
                 </gmd:MD LegalConstraints>
              (gmd:resourceConstraints>
                                                                                                                                conditions applying to access and use
                 <gmd:MD LegalConstraints>
                    <gmd:useLimitation xsi:type="gmd:PT FreeText PropertyType">
                       <gco:CharacterString>Pas de restriction d'accès public selon INSPIRE</gco:CharacterString>
                   </gmd:useLimitation>
                    <gmd:useConstraints>
                       <gmd:MD RestrictionCode codeList="http://standards.iso.org/iso/19139/resources/gmxCodelists.xml#MD RestrictionCode" codeListValue="otherRestrictions"/>
                    </gmd:useConstraints>
                    <gmd:otherConstraints xsi:type="gmd:PT FreeText PropertyType">
                      <gco:CharacterString>Les données peuvent être (ré)utilisées librement, sous réserve de mentionner la source. Les conditions d'accès et d'utilisation sont décrites dans la licence
standardisée CC-By (Creative Commons-Attribution) consultable sur https://creativecommons.org/licenses/by/4.0/legalcode</gco:CharacterString>
                   </gmd:otherConstraints>
                 </gmd:MD LegalConstraints>
              </gmd:resourceConstraints>
```



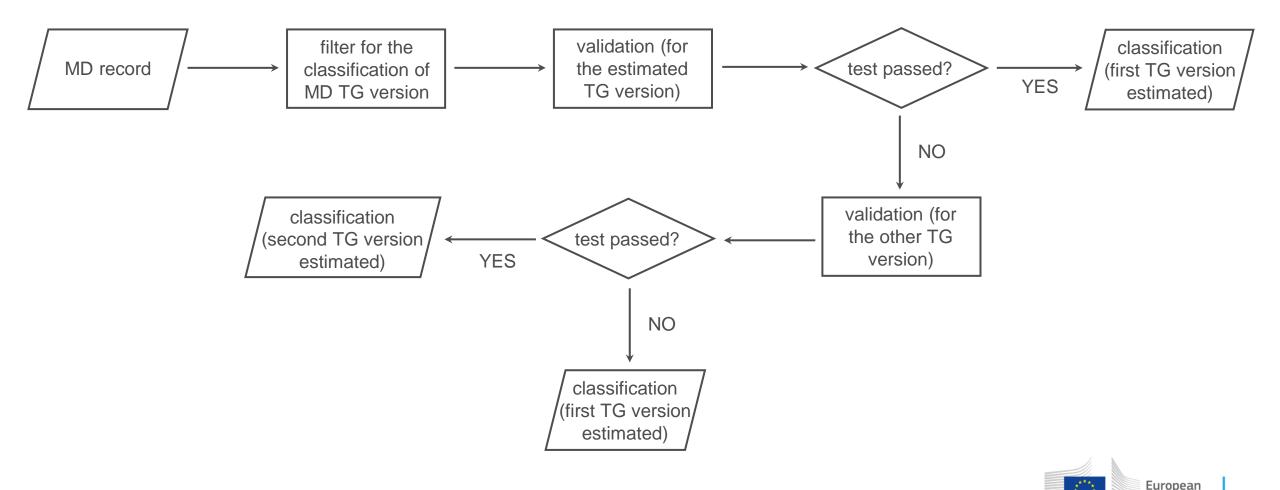
Workflow for estimation of MD TG version

- Final approach:
 - 1. classification (1.3 vs. 2.0) based on the same <gmd:useLimitation> filter
 - 2. validation against the corresponding Conformance Class(es)
 - 3. if the test is passed, the MD record is classified as initially determined
 - 4. if the test is NOT passed, the test against the Conformance Class(es) of the other TG is run
 - 5. if this second test is passed, the MD record is classified as compiled according to the TG tested later
 - 6. if this second test is NOT passed, the MD record is classified back as initially determined by the filter



Workflow for estimation of MD TG version

Final approach:



Workflow for MD validation

- Final approach:
 - we cannot be 100% sure that the final classification (1.3 or 2.0)
 corresponds to the actual 'intention' of the metadata creator
 - it is only an estimation!
 - BUT this does NOT impact on the value of the indicators:
 - if a MD record is conformant against TG 1.3 OR 2.0, this is detected!

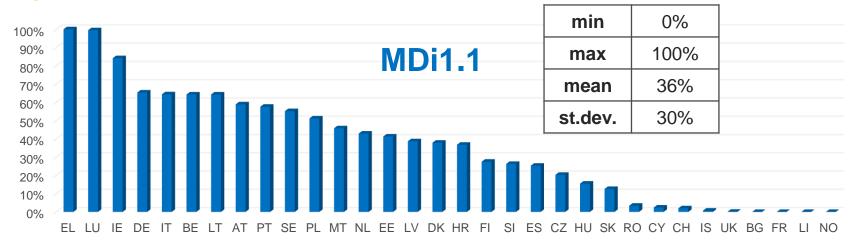


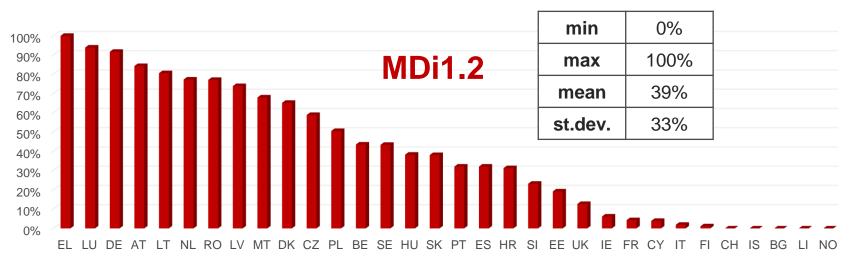
Global results

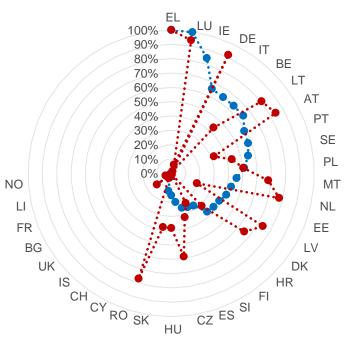
- Total number of metadata: 257,524
 - data set metadata: 159,003
 - service metadata: 98,521
- Total portion of conformant metadata: 36.1%
 - total portion of conformant data set metadata: 35.2%
 - total portion of conformant service metadata: 38.8%



Country-specific results







•••••• MDi1.1 ••••• MDi1.2

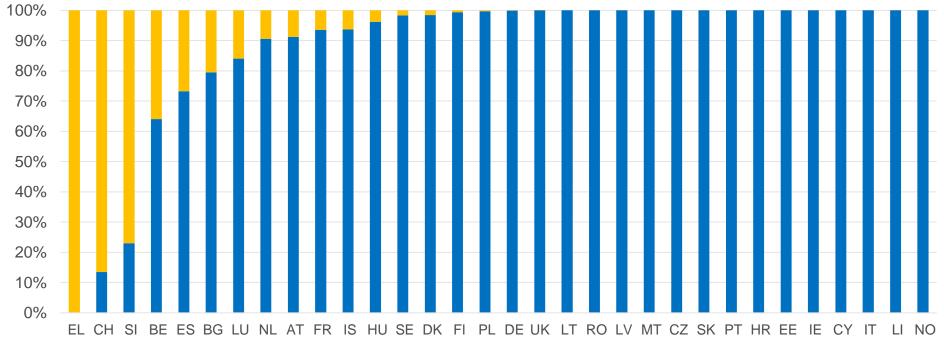


Estimation of MD TG version

Global results:

	data set MD	service MD	Total
TG 1.3	96%	99%	97%
TG 2.0	4%	1%	3%

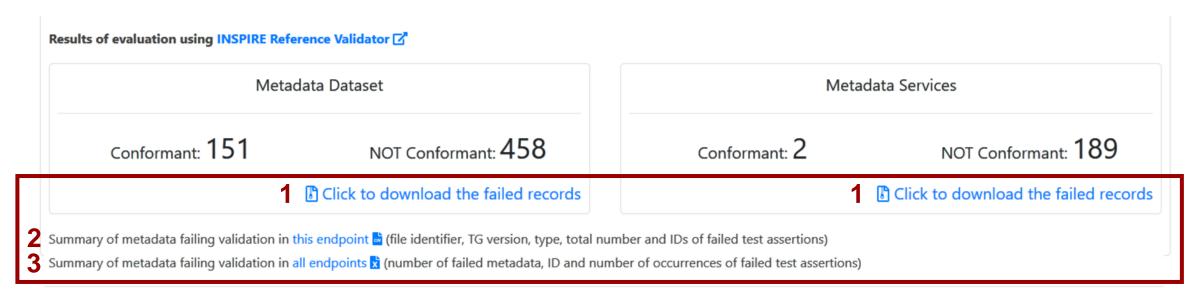
Country-specific results:





Supplementary material

- Offered to help data providers understand validation errors & improve MD:
 - 1. test reports for failed data set & service metadata (HTML, JSON)
 - 2. summary of errors for all metadata failing validation in the selected endpoint (CSV)
 - 3. summary of errors for all metadata failing validation in all endpoints (XLSX)



58% of the survey respondents rated this material very useful



1. Download the test reports for the failed MD records:

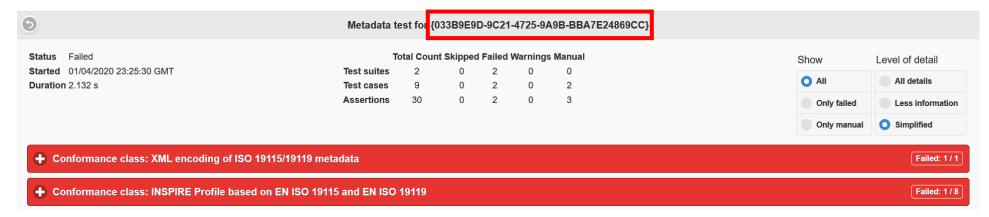


the zipped folder includes reports in HTML and JSON:

Name	Date modified	Туре	Size
e 1-100_1.1.html	02/04/2020 08:26	HTML File	190 KB
	02/04/2020 08:26	JSON File	212 KB
e 1-100_1.html	02/04/2020 08:26	HTML File	198 KB
1-100_1.json	02/04/2020 08:26	JSON File	217 KB
e 1-100_2.1.html	02/04/2020 08:26	HTML File	190 KB
1-100_2.1.json	02/04/2020 08:26	JSON File	212 KB
e 1-100_2.html	02/04/2020 08:26	HTML File	198 KB
	02/04/2020 08:26	JSON File	217 KB
e 1-100_5.1.html	02/04/2020 08:26	HTML File	190 KB
1-100_5.1.json	02/04/2020 08:26	JSON File	212 KB
● 1-100_5.html	02/04/2020 08:26	HTML File	198 KB
№ 1-100_5.json	02/04/2020 08:26	JSON File	217 KB
e 1-100_6.1.html	02/04/2020 08:26	HTML File	190 KB
	02/04/2020 08:26	JSON File	212 KB
e 1-100_6.html	02/04/2020 08:26	HTML File	197 KB
1-100_6.json	02/04/2020 08:26	JSON File	217 KB



2. Open the test report (e.g. in the HTML version) and analyse the errors:



3. Find the corresponding metadata in the CSV file (using the file name of the test report or the metadata fileIdentifier) to know the estimated TG version, the type, total number of errors and the tests failing validation:

file_id	md_id	version	type	error_count	errors
201-266_63	{FD6E1A28-3A7C-43D9-B8D4-839AD3A83308}	1.3	dataset	1	c.3
1-100_47	{262C85FA-2822-482F-882B-721DB7D754EA}	1.3	dataset	2	a.5 c.4
1-100_5	{033B9E9D-9C21-4725-9A9B-BBA7E24869CC}	1.3	dataset	2	xml.a.1 f.4
1-100_63	{3B2F157E-3DE5-4CFC-AB36-192E97156D68}	1.3	service	1	g.1
1-100_74	{4F7FCCD2-B060-4399-A5C4-57181EE24D51}	1.3	dataset	3	a.5 c.4 f.4



4. Find the corresponding MD record (using the value of </gmd:fileIdentifier>):

5. Fix the MD record based on the error message in the test report; reading the ATS (and optionally the ETS) can also help find the error.



 The XLSX file provides a summary of errors for all metadata failing validation in all endpoints & includes the title and ATS link of each test:

md_failed	number of metadata records that failed validation (includes md_unknown_type)								
md_unknown_type	number of metadata records with unknown or missing type (in dataset only)								
Test assertion ID	Title of the test assertion	ATS link							
xml.a.1	md-xml.a.1: Validate XML documents	http://inspire.ec.europa.eu/id/ats/metadata/1.3/xml-encoding/schema-validation							
a.1	md-iso.a.1: Title	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/title							
a.2	md-iso.a.2: Abstract	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/abstract							
a.3	md-iso.a.3: Access and use conditions	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-access-use							
a.4	md-iso.a.4: Public access	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-public-access							
a.5	md-iso.a.5: Specification	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-specification							
a.6	md-iso.a.6: Language	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/language							
a.7	md-iso.a.7: Metadata contact	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/md-contact							
a.8	md-iso.a.8: Metadata contact role	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/md-contact-role							
a.9	md-iso.a.9: Resource creation date	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/resource-creation-date							
a.10	md-iso.a.10: Responsible party contact info	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/responsible-party-contact-info							
a.11	md-iso.a.11: Responsible party role	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/responsible-party-role							
b.1	md-iso.b.1: Hierarchy	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/hierarchy							
c.1	md-iso.c.1: Dataset identification	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-identification							
c.2	md-iso.c.2: Dataset language	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-language							
c.3	md-iso.c.3: Dataset linkage	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-linkage							
c.4	md-iso.c.4: Dataset conformity	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-conformity							
c.5	md-iso.c.5: Dataset topic	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/ds-topic							
c.6	md-iso.c.6: Dataset geographic Bounding box	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/geographic-bounding-box							
c.7	md-iso.c.7: Dataset lineage	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/lineage							
d.1	md-iso.d.1: Service type	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/srv-type							
d.2	md-iso.d.2: Service linkage	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/srv-linkage							
d.3	md-iso.d.3: Coupled resource	http://inspire.ec.europa.eu/id/ats/metadata/1.3/iso-19115-19119/coupled-resource							
readme 1.3	dataset 1.3 service 1.3 readme 2.0	dataset 2.0 service +							



 The XLSX file provides a summary of errors for all metadata failing validation in all endpoints & includes the title and ATS link of each test:

Country	Endpoint ID	md_failed	xml.a.1	a.1	a.2	a.3	a.4	a.5	a.6	a.7	a.8	a.9	a.10	a.11	b.1	d.1	d.2	d.3	e.1	f.1	f.2	f.3	f.4	g.1	h.1
DE	INSPIRE-4fed3eb0-06fa-11ea-8480-525400695e9c	2871	10	0	0	6	20	0	35	146	4	0	8	0	0	7	1115	2081	0	0	41	69	7	0	2
DK	INSPIRE-6e8353b4-de80-11e7-a188-52540023a883	47	14	0	0	1	7	0	0	0	2	0	0	0	0	7	8	10	0	0	3	4	0	0	0
EE	INSPIRE-608e479c-616e-11e2-b563-52540004b857	65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53	63	0	0	0	0	0	0	0
ES	INSPIRE-c6f329a0-4c3d-11e7-9e8f-52540023a883	156	0	0	0	0	0	0	0	0	7	0	0	0	0	0	29	150	0	0	0	0	2	0	0
FI	INSPIRE-f670705f-f4e9-11e6-81e4-52540023a883	184	47	1	1	3	1	1	180	12	45	1	16	1	1	14	55	137	1	1	16	5	9	1	1
FR	INSPIRE-5145fa60-0067-11e5-9ea6-52540004b857	49747	601	0	455	148	147	0	944	379	678	8	419	5	0	276	3673	10719	3	0	48063	33	285	0	4
HR	INSPIRE-697db035-9af0-11e3-8508-52540004b857	92	47	0	0	0	1	0	0	0	3	0	0	0	0	0	18	35	0	0	3	0	0	0	0
HU	INSPIRE-412cdb45-0b05-11e7-9a72-52540023a883	34	8	0	0	0	0	0	0	0	15	0	0	1	0	0	3	11	0	0	0	0	4	0	0
IE	INSPIRE-67c9c760-1be3-11e3-851a-52540004b857	15	4	0	0	0	0	0	0	0	0	0	0	0	0	0	14	1	0	0	10	0	0	0	0
IS	INSPIRE-7ba666ea-05bd-11e7-9a72-52540023a883	14	5	0	0	0	3	0	6	2	4	0	1	0	0	3	9	3	1	0	3	0	0	0	0
IT	INSPIRE-c22038a7-4e03-11e8-a459-52540023a883	2586	847	0	0	0	0	0	0	1	1	0	3	0	0	0	331	126	0	0	0	2586	9	0	0
LI	INSPIRE-86b8eec9-581d-11e4-b478-52540004b857	3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	3	2	0	0	0	0	0	0	0
LT	INSPIRE-106902a4-2bd0-11e9-a83c-52540023a883	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0
LU	INSPIRE-93ee1068-1dc3-11e7-a02d-52540023a883	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	0	0	0	0
LV	INSPIRE-10108b88-d195-11e5-91ce-52540023a883	37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	4	0	0	0	20	0	3	0
MT	INSPIRE-80e86358-9378-11e5-a300-a0369f4c5bc0	56	30	0	0	0	0	0	0	0	0	0	0	0	0	0	7	32	0	0	0	0	0	0	0
NL	INSPIRE-8c93a17a-05f4-11e1-b7de-52540004b857	67	2	1	1	1	1	1	1	4	5	1	6	1	1	1	26	42	3	1	4	2	1	1	1
NO	INSPIRE-ccf3ad04-9003-11e3-aef9-52540004b857	245	41	0	0	17	1	0	244	0	18	0	50	0	0	0	17	32	0	0	3	0	0	0	0
PL	INSPIRE-d81e48c4-b4cf-11e3-a455-52540004b857	1878	16	0	0	0	2	0	0	13	0	0	127	0	0	32	59	1741	1	0	23	5	1	0	12
PT	INSPIRE-d60bf7f3-ea96-11e4-a2c7-52540004b857	421	12	0	0	0	0	0	0	0	0	0	3	0	0	0	189	312	0	0	0	45	1	0	0
RO	INSPIRE-7edbed58-ddbc-11e4-b469-52540004b857	11	0	0	0	0	0	0	0	0	4	0	0	0	0	0	3	3	0	0	0	1	0	0	0
SE	INSPIRE-adae82ae-5364-11e8-bd03-52540023a883	118	8	0	0	24	7	0	0	2	2	0	0	2	0	0	36	63	0	0	6	3	3	0	0
SI	INSPIRE-6f0cd439-226d-11e6-9ff2-52540023a883	7	5	0	2	2	3	0	1	0	0	0	0	0	0	0	3	0	0	0	0	0	1	0	0
SK	INSPIRE-da77b119-9d6e-11e7-b5a7-52540023a883	130	2	0	0	1	1	0	0	1	0	0	0	0	0	2	87	81	0	0	1	10	0	0	0
UK	INSPIRE-f89f4772-05f5-11e1-b7de-52540004b857	552	6	0	0	2	51	0	0	1	0	0	3	0	0	0	254	352	0	0	146	7	12	1	2
Total		59622	1751	2	459	207	246	2	1412	565	817	10	636	10	2	347	6116	16182	9	2	48334	2798	352	6	26
>	readme 1.3 dataset 1.3 service 1.3 readme	2.0 datase	t 2.0 se	rvice	e 2.0		+				:	4													F

European Commission

Bulk validation of metadata

- The whole processing (estimation of TG version, validation, output storage, extraction of failed tests, calculation of indicators, etc.) was automated using Pentaho Data Integration (PDI) platform.
- The PDI scripts will be published under an open source license:
 - very soon
 - under the INSPIRE Validation organization in GitHub (https://github.com/inspire-eu-validation)
- The scripts will allow to:
 - replicate the validation process performed by the JRC
 - customize the process according to your needs (e.g. validate using your own instance of the Validator, validate only the MD of one single data provider, etc.)
- An updated script based on the tests performed in the next 2020 Monitoring round will be shared soon.



M&R 2019 conclusions & lessons learnt

- MD conformity was heterogeneous across MS but overall quite low.
- Little use of the INSPIRE Validator before the Monitoring deadline in December 2019.
- There was a high percentage of MD TG 1.3 although the deadline to switch to MD TG 2.0 expired in Dec 2019.
- This was our first-ever look at all MD records from MS, which allowed to:
 - discover & fix bugs in the tests performed by the Validator
 - improve the Validator in terms of stability & performance
 - identify inconsistencies with the Geoportal and related issues
- We recommend each MS to setup its own instance of the INSPIRE Validator to prepare to the next Monitoring round!



(Some) news on the Validator

- A release plan of the INSPIRE Validator was published to explain the release planning process in an open & transparent way, in preparation of the 2020 M&R process (https://github.com/inspire-eu-validation/community/tree/master/release%20strategy):
 - next releases are expected in mid-June and mid-September
 - from the June release, tests will NOT be made stricter
- A description of the helpdesk management workflow was published to facilitate user participation & interaction with the helpdesk (https://github.com/inspireeu-validation/community/tree/master/helpdesk%20management).
- A new INSPIRE UI of the Validator is under development.
- Don't miss the webinar on the INSPIRE Reference Validator (for users & developers) on June 9 at 9:00 CET!



Thank you!





© European Union 2020

European Commission

Keep in touch



EU Science Hub: ec.europa.eu/jrc



@EU_ScienceHub



EU Science Hub – Joint Research Centre



EU Science, Research and Innovation



Eu Science Hub

