Discussion records on fixed issues

Since April 2020 release

2020-23 ISO schemas moved again! (Peter)

Standards.iso.org has done it again, so the schema set advice we give at
https://www.agi.org.uk/40-gemini/1048-uk-gemini-encoding-guidance#2.1 is now longer quite
right.

https://schemas.isotc211.org/schemas/19139/-/gmx/1.0/gmx.xsd works

Amended 2020-09-29, as I needed to also fix the reference to the INSPIRE Validator further
down the same page.

2020-12: Disappearance of (TC211, Mar 2020)

Fixed later in https://www.agi.org.uk/40-gemini/1048-uk-gemini-encoding-guidance

JP: Only the following is working at the moment for validation:
http://inspire.ec.europa.eu/draft-schemas/inspire-md-schemas-temp/apiso-inspire/apiso-
inspire.xsd

Other listed schema have some sort of redirect in place that breaks the validator

April 2020 release

2020-12: Disappearance of (TC211, Mar 2020)

See https://github.com/ISO-TC211/schemas/issues/26 & 24.
Affects GEMINI XML general guidance, an informative comment at 2.1
(https://www.agi.org.uk/40-gemini/1048-uk-gemini-encoding-guidance#2) references the
"old" ISO locations, but we do not recommend importing them directly. OGC & INSPIRE are
working to fix the files that we do recommend.

This from TC211 XMG Convenor: "The only authoritative space is
https://schemas.isotc211.org <https://schemas.isotc211.org/>"

[Done: remove the misleading text; put in relevant text. Also fixed a number of examples on
that page]

all code list references (41 - didn't flag element as changed)
http://standards.iso.org/ittf/PubliclyAvailableStandards/ISO_19139_Schemas/resources/codelist/gmxCodelists.xml should be updated to
https://schemas.isotc211.org/schemas/19139/resources/gmxCodelists.xml

Done in April release, still needs to be done in https://www.agi.org.uk/40-gemini/1046-uk-gemini-encoding-guidance

2020-10: Maintenance information maximum multiplicity (Peter, Mar 2020)

INSPIRE only allows one. Although 19115 allows multiple resourceMaintenance elements, I can't imagine why! GEMINI 2.2 only allowed one "Frequency of update".

Suggestion: make occurrence "single".## 2020-09: How to manage multiple "limitations on public access" (Sean, Feb 2020)

Which is the repeating unit in the XML encoding? Mention in the guidance & give an example.
- *19115 allows an MD_LegalConstraints to contain multiple accessConstraints and otherConstraints*
- *INSPIRE TG Requirement C.17 metadata/2.0/req/common/limitations-on-public-access states that "Limitations on public access shall be described using exactly one gmd:resourceConstraints/gmd:MD_LegalConstraints … not … the same one as used for describing conditions" going on to "at least one instance of gmd:otherConstraints/gmx:Anchor …."*

Item 7 in the "Guidelines" of the "Encoding guidelines" does already say that, so perhaps just an example with two otherConstraints (one INSPIRE, one not) - and change guidelines 5 to show that only one need be INSPIRE.

2020-08: Maintenance information: update scope & maintenance note

GEMINI only mentions these in the encoding guidelines. The main text gives domain as 'complex type' but no details. *Add domain description.*
Also, INSPIRE restricts 19115 by only allowing one update scope & one maintenance note. *Add to domain description (?) and 'corresponding element' text.*

2020-06: Metadata standard name guidance (Sean, 5th Feb 2020)

I (Peter) didn't implement everything we voted on - I did the Metadata version bit, but forgot to change the Metadata standard name guidance & encoding example.

The text content should be the 'preferred label' of the anchor target, so http://vocab.nerc.ac.uk/collection/M25/current/GEMINI/ is "UK GEMINI". Change both the examples & the guidance. Perhaps "AGI GEMINI" if AGI would prefer that.

2020-05: Conformity (Peter Parslow, 30th Jan 2020, UKME Beta testing):

Change history

https://www.agi.org.uk/40-gemini/1062-gemini-datasets-and-data-series#41 change history says "GEMINI 2.2 to 2.3: it is no longer allowed to state that a resource has not been tested against the relevant INSPIRE specification", but that's not strictly true: it's the mechanism (encoding) for saying it that has changed.

Text changed to "changed the way to encode that a resource has not been tested against the relevant INSPIRE specification"

Fix change history & "no" to "not" in comment in example

INSPIRE conformity statement mandatory

The Schematron always checks for at least one INSPIRE conformity statement. This is in line with the overall aim that a GEMINI record is an INSPIRE record. Therefore, GEMINI should state that 'at least one conformity statement shall be to an INSPIRE specification, even if simply to say that the data set is not conformant or not tested'

2020-04: Resource identifier examples, reported by Jo Cook 31st May

"According to

https://www.agi.org.uk/agi-groups/standards-committee/uk-gemini/40-gemini/1062-gemini-datasets-and-data-series#36

it's at least strongly recommended that this uses the
gmd:MD_Identifier element rather than gmd:RS_Identifier, however I
notice that the two dataset examples in the schema download use
gmd:RS_Identifier or a mix of the two."

OS Boundary Line example file fixed on GitHub December 2019

Peter: need to fix BGS example (c) has both, using RS_Identifier to encode the DOI. ~~According to https://www.agi.org.uk/agi-groups/standards-committee/uk-gemini/40-gemini/1062-gemini-datasets-and-data-series#36, the DOI should be encoded in otherCitationDetails:~~

~~<gmd:otherCitationDetails>~~
~~<gco:CharacterString>doi:10.5285/1e7d5e08-9e24-471b-ae37-49b477f695e3</gco:CharacterString>~~
~~</gmd:otherCitationDetails>~~

Fixed in GitHub

Metadata working group to confirm this

Remove discussion of DOI from alternative title (guidance, encoding guidance & example). Add as a note/comment in Additional information that the other resource could be referenced by its DOI. Fixed

The Resource identifier encoding should be MD_Identifier, with an Anchor - URL to the landing page & text element being the DOI. Fixed

2020-11: Invalid example file (Jo, Mar 2020)

See also 2020-04
The example file at https://github.com/AGIGemini/Schematron/blob/master/samples/BGSds-example1c.xml contains an invalid Data quality element, at lines 890-900. It needs to contain a result element. Remove lines 890-900. The report between lines 859 and 870 doesn't look that helpful either!
Fixed in GitHub
Jo adds: "a record with comments for the placeholder elements would be hugely useful". Good idea - needs work

2020-07: The series XML example is invalid (Jo/James, 11 Feb 2020)

We don't have a valid GEMINI 2.3 "series" file. The guidance links to the GEMINI 2.2 one, which isn't valid.
Fix: we need to create a valid series metadata instance.

Suggestion: link from the guidance to the examples in GitHub. Fixed on website

January 2020 release

Live Feb 5th 2020
- Keyword for INSPIRE theme: encoding example uses the wrong URL. Also, could describe it better in the guidance.
- (Jo Cook) Typo in UseConstraints encoding "<gmd:otherContraints>" *actually occurred four times*

- HierarchyLevelName (Jo Cook 17th July 2019) To match the INSPIRE TG, it's not just "an appropriate text string", but specifically "service", "series", or "dataset" (in which case it's not required). Changed guidance & examples; encoding seems OK.
- Conformity (INSPIRE): minor fix to guidance & new 'known error' (reported by Phil Trembath, Jan 2020)
- Improve Limitations on public access: list/explain the INSPIRE reasons, and hyperlink to their register (idea from Peter, Dec 2019)
- (re)introduce Metadata standard name & version, as optional, but with clear guidance what to put in? (Suggested by Jo Cook, February 2019) - yes; much email correspondence, resolved by January 2020

June 2019: Schematron on GitHub

Suggested by Jo Cook, Feb 2019; implemented by June 2019:
https://github.com/AGIGemini/Schematron

March 2019

1. Schema location - changed https://www.agi.org.uk/1048-uk-gemini-encoding-guidance#2.1 to describe a choice of XSDs

Live on 2nd April 2019:

2. *Dataset* reference date: removed mention of the maximum occurence for service records, when viewing this element for dataset records.
3. Corrected the INSPIRE "corresponding element" entry for Conformity, Abstract, Topic category, Keyword, Spatial reference system, Data format, File Identifier, Hierarchy level name, and Quality scope.
4. Corrected the "corresponding element" entries for service elements to reference ISO 19119 for Extent, Vertical extent information, Spatial data service type, and Coupled resource.
5. Corrected the ISO 19139 "corresponding element" entries for service elements Extent, Vertical extent information, and Spatial resolution
6. Corrected the ISO 19139 "corresponding element" entries for Title, Alternative title, and Abstract.

Note: these are all minor corrections of the way the element is referenced, rather than substantive changes to which element we consider equivalent. The largest impact is to note that INSPIRE only allows one Quality scope element - i.e. the metadata has to describe the whole dataset.

Reported; resolved in March 2019 as not errors

1. INSPIRE "Conditions for access and use" / GEMINI Use constraints: INSPIRE requires (harvest checker) this to be ISO useLimitation

Suggested this be changed at https://www.agi.org.uk/40-gemini/1048-uk-gemini-encoding-guidance and in the encoding guidance & examples at https://www.agi.org.uk/agi-groups/standards-committee/uk-gemini/40-gemini/1062-gemini-datasets-and-data-series#26

2. (Sean) "Equivalent metadata element" for services Coupled Resource: change from "9 operatesOn" to something like SV_ServiceIdentification:operatesOn. Similarly, for Spatial data service type, change from "1 serviceType";

October 2018

1. Resource identifier (reported by Peter)

Typos in comment:
- "it is no longer required to use RS_Identifier if a value for if a code space value is provided. " delete either "a value for if" or "for if a" & 2nd "value".
- "MD_Iₙdentifier"

Fixed in 20181031 on Peter's laptop; uploaded 2nd November

2. Additional information source (reported by Sean in September)

For background, see meeting notes at October 24th. Note this element only applies to dataset & series metadata, not service metadata.
Draft revision text:

1.   Resource locator (element 19):
    1.1.   Amend Guidance bullet point 4, which currently reads "If no online access for the dataset or dataset series is available, but there is a publicly available online resource providing additional information about **the described dataset or dataset series**, the URL pointing to this resource shall be given instead." to say "If no online access for the dataset or dataset series is available, but there is a publicly available online resource providing additional information about **how to access the dataset or dataset series being described**, the URL pointing to this ~~resource~~ shall be given instead."
    1.2.   Amend Example 2 to one which fits this less ambiguously - the current BGS worldStatistics page is more general, although it does contain the download link. Perhaps https://os.uk/opendatadownload/products.html - but I'd be happier with a BGS (or other non-OS) example

*Also: tweaked the "Purpose and meaning" from "if no direct access is avaialble, to an online resource providing more information about the dataset or dataset series", to "if no direct access is availalble, to an online resource providing more information about **accessing** the dataset or dataset series" (& equivalent for service), and removed third bullet "This may be different from where the data may be ordered online, which should be included in the web address of the distributor." Also improved the text introducing the examples (which was common across dataset & service, and therefore somewhat misleading).*

2.   Additional information source
    2.1.   Rename it to "additional information", or even "supplemental information"
    2.2.   Amend "Purpose and meaning" from "The purpose of this element is to record references to relevant information held externally, for example a reference (e.g. a URL) to background information." to "The purpose of this element is to record ~~references to~~ relevant information **that does not clearly belong in another element.** ~~held externally, for example a~~ This may be a reference (~~e.g. a~~ URL) ~~to background information~~."
    2.3.   Replace Guidance with:
        ●   Information about access to the resource should be in 'Resource locator'
        ●   Information about fees should be in 'Limitations on public access'
        ●   Other relevant information should be in this element, or provided by a link which is given in this element
    2.4.   Add a second encoding example using a gmx:Anchor:

```
<gmd:supplementalInformation>
    <gmx:Anchor
xlink:href='http://www.bgs.ac.uk/mineralsuk/statistics/worldStatisti
cs.html/'>Background information</gmx:Anchor>
<gmd:supplementalInformation>
```

    2.5.   Add as INSPIRE corresponding element "supplemental information".

(Note: this is only mentioned in the INSPIRE Elevation spec and therefore at C.7.3 of the current INSPIRE Metadata TG "This metadata element is recommended to concisely describe general properties of the spatial data set")

Fixed in 20181031 on Peter's laptop; uploaded 2nd November

September 2018

All fixed in the XML 2018-09-27 which was uploaded on 28th September, and documented in the change log.

1. Use constraint (reported by Sean Gaffney)

Typo in encoding - has gmd:LegalConstraints, which should be gmd:MD_LegalConstraints - in Examples one, two, and three. Also change it in the encoding rules.
Also, change the service Example One to be just like the Dataset one, with the exception of srv:SV_ServiceIdentification

2. Irish language (reported by Jo Cook)

The official name of the Irish language is "Irish", not "Gaelic (Irish)" - change in the Guidance for elements 3 and 33.

3. Spatial reference system (reported by Peter Parslow)

GEMINI states that Spatial reference system has no INSPIRE equivalent (at least, in the data view). But see 3.2.1.1 in the INSPIRE guidance: metadata/2.0/req/isdss/crs. It's just that it doesn't have a "name"!
Similarly, 'Quality scope' is required by INSPIRE metadata/2.0/req/datasets-and-series/one-data-quality-element

4. Use constraints (reported by Peter Parslow)

We state that the equivalent ISO 19115 element is useLimitation. The examples are correct - it is useConstraints.
Similarly, the 'equivalent element' for 'limitations on public access' should actually mention both the ISO elements used (access constraints and other constraints)

August 2018

1. Parent identifier (reported by Sean Gaffney)

Fix published 31st August.
Guidance: "only to be used if the dataset is part of a series"; "only to be used if the service is designed to be part of a set of services" Peter to put into content.

(Peter) It looks like I failed to do that properly, and everyone failed to spot it. The text at https://www.agi.org.uk/gemini/40-gemini/1062#49 is the same as https://www.agi.org.uk/gemini/40-gemini/1063#49.

This is a "bug" in the XML that is common across the two (datasets & services).

2. Schema validation (reported by James Passmore)

Fixed online (29th August)

https://www.agi.org.uk/gemini/40-gemini/1053-common-metadata-errors-uk-location-discovery-metadata-service:

Under section 1.5 on Validation it says:

"First, a metadata set is validated against the ISO 19139 schemas. UK Location / data.gov.uk use the XSD files provided by EDEN; see UK GEMINI Encoding Guidance for details."

This should be reworded, I feel, because EDEN no longer supply any schema. I've been validating using http://inspire.ec.europa.eu/draft-schemas/inspire-md-schemas/apiso-inspire/apiso-inspire.xsd. I guess the existing validation system uses copies of the EDEN schema, but into the future it will need to switch to something else.

The next two bullet points in that section go on to describe the process that exists for GEMINI 2.1/2 validation, but I'm not sure if the same system will apply for GEMINI 2.3, or will the existing GEMINI 2.3 schematron be split up for the data.gov validator?

For example looking at the following copy of the ISO 19139 Table A.1 schematron (https://raw.githubusercontent.com/metadata101/iso19139.gemini22/3.4.x/src/main/plugin/iso19139.gemini22/schematron/schematron-rules-ISOTS19139A1Const.sch) I see there are tests in there covered by the GEMINI 2.3 schematron.

So perhaps the line "Validating a UK Location metadata record is a three stage process:" could be changed to "Validating a UK Location GEMINI 2.1 metadata record is a three stage process:"

OR maybe it should be 1.5 Validation / 1.5.1 GEMINI 2.1 and 1.5.2 GEMINI 2.3

Then in GEMINI 2.3 section you could have:

"There is no currently no UK Location validation tool for GEMINI 2.3". To validate a metadata record in your own software you should

Validate using the ISO 19139 (APISO) schemas. (http://inspire.ec.europa.eu/draft-schemas/inspire-md-schemas/apiso-inspire/apiso-inspire.xsd)

Validate for additional requirements using the GEMINI 2.3 Schematron Schema. (https://www.agi.org.uk/about/resources/category/125-schematron)

It doesn't matter which order validation is done in, but to be valid records must be validate against both the APSIO schema and the GEMINI 2.3 Schematron.

In addition a second (supplemental) schematron is provided to test for recommendations in your metadata structure."

3. Spatial data service type (reported by Sean)

Fix published 31st August.
In the GEMINI 2.3 service metadata profile, we have spatial data service type
https://www.agi.org.uk/gemini/40-gemini/1063-gemini-services#37

 This has a requirement level of conditional. However, the conditionality is that it is required if the data resource is a service. Given we're dealing with service metadata here, does this then not become a mandatory requirement instead?

4. Minimum length of Abstract (reported by Sean)

Fix published 31st August.
Add line to guidance for Abstract that the Schematron will complain if it's < 100 characters.

5. INSPIRE Default CRSs (reported by Peter)

Fix published 31st August.
At https://www.agi.org.uk/gemini/40-gemini/1062-gemini-datasets-and-data-series#17, we say these are listed in Annex D.4 of the Metadata Regulation; they're not - it's in the Metadata TG