Documenting XML Standards with the TDWG SDS

An assessment from the PlinianCore and ABCD Working Groups

Should XML Standards be documented using the SDS

YES

- Advantages
    - Provides a lot of the documentation needs: names, definitions, examples,
    - Common Style and Structure of the documentation as the other TDWG Standards

- What is missing?
    - Structure and Context

Adding Structure to the SDS

Term-Structure
- Link to xsd-schema file (or RelaxNG or DTD)*
- XPath

SDS {

Term
- TermIRI*
- Label*
- Definition*
- modified
- TermVersionIRI
- Example
- Usage
- …

XPath: XPath location with the XML Document that is described by the XML-Schema

* normative

How to handle element reuse?

```
<root>
    <section1>
       ...
       <reference>
            <referenceText></referenceText>
            <referenceURI></referenceURI>
       </reference>
    </section1>
    <section2>
       ...
       <reference>
            <referenceText></referenceText>
            <referenceURI></referenceURI>
       </reference>
    </section2>
</root>
```

Option 1:

Reference
Definition: A reference to a source
XPath(s): /root/section1/reference
              /root/section2/reference

Option 2:

Section1-Reference
Definition: The reference to the source of section 1.
XPath: /root/section1/reference

Section2-Reference
Definition: The reference to the source of section 2.
XPath: /root/section2/reference

Open Questions/Potential Issues

Assumptions about the XML Standards

- not mixed content

      <section>
            Text within an element that also contains other elements
            <reference>...</reference>
      </section>

- no loops

      <section>
            <data>
                  <section>...</section>
            </data>
            <reference>...</reference>
      </section>

Special Cases (that are already covered by our approach)

- Attributes
    - /root/element/subelement@attribute
- Ordering
    - so far not important
    - could be documented as:1.3.3.2 (*)
- Cardinality of elements
    - defined in Schema
    - could be documented as (*)
        - required/optional
        - repeatable
        - 1/0-1/0-n/1-n

*: tied to the XPath location

Other thoughts

- So far we have only looked at the path from an XML Standard towards SDS compliance, not from having an SDS compliant standard and wanting to add a structural layer on top of it.
- We need to document the steps and tools that we used to go from an xsd-Schema to the flat documentation (not part of today's discussion)
- The same approach should also work with JSON Schema, with JSON-Path instead of XPath.
- We only use a subsection of the XPath specification.

How to manage user experience if there are too many xpaths

4.1 Metadata

Basic information to identify the record; basic information to identify the record; and information about the Version, Revision, the language and target audiences of the Taxon Record.

Term Name plic:Audience

Term IRI	http://rs.tdwg.org/plic/terms/Audience
Modified	2022-06-24
Term version IRI	http://rs.tdwg.org/plic/terms/version/Audience-2022-06-24
Label	Audience
Definition	Repetitive concept where the target audiences are written one by one.
Notes	Predetermined list: 01. Expert (High). 02. Knowledgeable (Medium). 03. General (Low).
	/AncillaryData/audience /AnnualCycleAtomized/AncillaryData/audience /AnnualCycles/AncillaryData/audience /AnnualCycles/AnnualCycleAtomized/AncillaryData/audience /Audience /Behavior/AncillaryData/audience /Behavior/BehaviorAtomized/AncillaryData/audience /BehaviorAtomized/AncillaryData/audience /CommonNames/AncillaryData/audience /CommonNames/CommonNamesAtomized/AncillaryData/audience /CommonNames/CommonNamesAtomized/UsedIn/DistributionScope/AncillaryData/audience /CommonNamesAtomized/AncillaryData/audience /CommonNamesAtomized/UsedIn/DistributionScope/AncillaryData/audience /Dataset/AncillaryData/audience /Dataset/TaxonRecord/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/DirectThreats/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/DirectThreats/DirectThreatAtomized/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/Legislation/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/Legislation/LegislationAtomized/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/Legislation/LegislationAtomized/AppliesTo/DistributionScope /AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/PopulationBiology/AncillaryData/audience /Dataset/TaxonRecord/DemographyAndThreat/PopulationBiology/PopulationBiologyAtomized/AbundanceData /AncillaryData/audience

Term Name plic:Revision

Term IRI	http://rs.tdwg.org/plic/terms/Revision
Modified	2022-06-24
Term version IRI	http://rs.tdwg.org/plic/terms/version/Revision-2022-06-24
Label	Revision
Definition	Creators, revision status and dates of the record.
Equivalent XPath	/Dataset/TaxonRecord/RecordMetadata/Revision /RecordMetadata/Revision /Revision /TaxonRecord/RecordMetadata/Revision
Type	Class