Structuring Semantic Information with Respect to Conventional Metadata (Syntactic) Structures

Metadata Interest Group (MIG) - RDA Plenary 21
Monday, 23 October 2023
Breakout 1
Metadata Required for FAIR Digital Objects

Rainer Stotzka and many others:

RDA IG FAIR DO Fabric, Helmholtz Metadata Collaboration Platform, FDO Forum, NFDI4Ing, NFDI-MatWerk, and more

DFG project numbers 442146713 & 460247524
FAIR Digital Object:
Representation of research data and other digital assets that contains all information required for FAIRness

- Data content
- Referenced and identified by a persistent ID
- Has properties described by metadata, e.g. type

- General concept
- Long-lasting & sustainable
- Technology agnostic
- Interoperability and harmonization
- Bridges between research data repositories, disciplines, etc.

- Seamless inclusion of legacy data
- Data has a type
- Operations can be connected to the types

How to implement?
The representation contains at minimum an **Identifier** and an **Information Record** with the following properties:

- **Identifier**: *unambiguously assigned*, e.g. a PID
- **Information Record**: well and *unambiguously defined structure* containing a *minimum of information* entries or links pointing to the information:
  - identifier,
  - definition of the structure of the information record, e.g. profile,
  - type of the data,
  - link to the data, e.g. a PURL
RDA Draft Kernel Information Profile (KIP) defines **15 basic attributes**, mostly administrative information.

- **Extension** of Draft KIP by contextual and relational attributes agreed on between representatives from all research fields.
- **Goal**: Increase immediate (scientific) benefit of using FAIR DOs.
- **Compatible** to RDA Recommendations.
- **Basis for all FAIR DOs** created within the Helmholtz Association.
- **Extensible** by additional attributes if required.
- **Guidance document available**, publication soon.

<table>
<thead>
<tr>
<th>Additional Helmholtz KIP Attributes</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>digitalObjectLocation-AccessProtocol</td>
<td>Access information for digitalObjectLocation, e.g., protocol, protocol version, and client</td>
</tr>
<tr>
<td>underEmbargoUntil</td>
<td>Access restrictions probably apply before</td>
</tr>
<tr>
<td>license</td>
<td>Extracted from digitalObjectPolicy</td>
</tr>
<tr>
<td>checksum</td>
<td>Renamed from 'etag' to be more specific</td>
</tr>
<tr>
<td>signature</td>
<td>Cryptographic signature of PID record</td>
</tr>
<tr>
<td>topic</td>
<td>Topic term from vocabulary for additional context</td>
</tr>
<tr>
<td>locationPreview</td>
<td>Optional preview for digitalObjectLocation</td>
</tr>
<tr>
<td>contact</td>
<td>Contact information, e.g., ORCiD or ROR</td>
</tr>
<tr>
<td>hasMetadata</td>
<td>PID pointing to a related FDO containing metadata</td>
</tr>
<tr>
<td>isMetadataFor</td>
<td>Inversion for hasMetadata</td>
</tr>
<tr>
<td>wasGeneratedBy</td>
<td>W3C PROV-DM element to refer to tool/agent used for generating the digital object</td>
</tr>
<tr>
<td>provenanceGraph</td>
<td>Optional PID of full provenance graph</td>
</tr>
</tbody>
</table>