

FAIR Metadata Concepts in DataCite Metadata Schema

Ted Habermann, Metadata Game Changers

Findable Text	
Documentation concepts used to support many types of data discovery, i.e. text, spatial/temporal, author, publisher, etc.	
Concept	Path(s)
Abstract	data[*].attributes.descriptions[?(@.descriptionType="Abstract")].description
Award Title	data[*].attributes[*].fundingReferences[*].awardTitle
Date Created	data[*].attributes.dates[?(@.dateType="Created")].date
Keyword	data[*].attributes.subjects[*].subject
Keyword Vocabulary	data[*].attributes.subjects[*].subjectScheme
Project Funder	data[*].attributes.fundingReferences[*].funderName
Resource Author	data[*].attributes.creators[*].name
Resource Author Affiliation	data[*].attributes.creators[*].affiliation
Resource Identifier	data[*].attributes.doi data[*].attributes.identifiers[*].identifier
Resource Publication Date	data[*].attributes.publicationYear
Resource Publisher	data[*].attributes.publisher
Resource Title	data[*].attributes.titles[*].title
Resource Type General	data[*].attributes.types.resourceTypeGeneral
Funder Project Identifier	data[*].attributes.fundingReferences[*].awardNumber
Spatial Extent	data[*].attributes.geoLocations[*].geoLocationBox data[*].attributes.geoLocations[*].geoLocationPoint

	data[*].attributes.geoLocations[*].geoLocationPolygon data[*].attributes.geoLocations[*].geoLocationPlace
Temporal Extent	data[*].attributes.dates[?(@.dateType="Collected")].date

Findable Identifiers

Documentation concepts that provide extra information, i.e. identifiers and references, about discovery concepts.

Concept	Path(s)
Award Number	data[*].attributes[*].fundingReferences[*].award-Number
Award URI	data[*].attributes[*].fundingReferences[*].awardUri
Date Submitted	data[*].attributes.dates[?(@.dateType="Submitted")].date
Funder Identifier	data[*].attributes[*].fundingReferences[*].funderIdentifier
Funder Identifier Type	data[*].attributes[*].fundingReferences[*].funderIdentifierType
Keyword Value URI	data[*].attributes.subjects[*].valueURI
Keyword Vocabulary URI	data[*].attributes.subjects[*].schemeUri
Resource Author Affiliation Identifier	data[*].attributes.creators[*].nameIdentifiers[*].nameIdentifier
Resource Author Affiliation Identifier Scheme URI	data[*].attributes.creators[*].nameIdentifiers[*].schemeUri
Resource Author Affiliation Identifier Type	data[*].attributes.creators[*].nameIdentifiers[*].nameIdentifierScheme
Resource Author Identifier	data[*].attributes.creators[*].nameIdentifiers[*].nameIdentifier
Resource Author Identifier Type	data[*].attributes.creators[*].nameIdentifiers[*].nameIdentifierScheme
Resource Author Type	data[*].attributes.creators[*].nameType

Resource Identifier Type	data[*].attributes.identifiers[*].identifierType
Resource Type	data[*].attributes.types.resourceType

AIR Connections

Documentation concepts for datasets interoperability and connections for resource documentation, understanding, and trust.

Concept	Path(s)
CitedBy	data[*].attributes.relatedIdentifiers[?(@.relationType='IsCitedBy')].relatedIdentifier
Date Available	data[*].attributes.dates[?(@.dateType="Available")]
DescribedBy	data[*].attributes.relatedIdentifiers[?(@.relationType='IsDescribedBy')].relatedIdentifier
Distribution Contact	data[*].attributes.contributors[?(@.contributorType='Distributor')].name
DocumentedBy	data[*].attributes.relatedIdentifiers[?(@.relationType='IsDocumentedBy')].relatedIdentifier
HasMetadata	data[*].attributes.relatedIdentifiers[?(@.relationType='HasMetadata')].relatedIdentifier
Methods	data[*].attributes.descriptions[?(@.descriptionType="Methods")].description
ReferencedBy	data[*].attributes.relatedIdentifiers[?(@.relationType='IsReferencedBy')].relatedIdentifier
Resource Contact	data[*].attributes.contributors[?(@.contributorType='ContactPerson')].name
Resource Format	data[*].attributes.formats
Resource Size	data[*].attributes.sizes
Resource URL	data[*].attributes.url
ReviewedBy	data[*].attributes.relatedIdentifiers[?(@.relationType='IsReviewedBy')].relatedIdentifier

Rights	data[*].attributes.rightsList[*].rights
Rights URI	data[*].attributes.rightsList[*].rightsUri
RightsHolder	data[*].attributes.contributors[?(@.contributorType='RightsHolder')].name
SourceOf	data[*].attributes.relatedIdentifiers[?(@.relationType='IsSourceOf')].relatedIdentifier
SupplementTo	data[*].attributes.relatedIdentifiers[?(@.relationType='IsSupplementTo')].relatedIdentifier
TechnicalInfo	data[*].attributes.descriptions[?(@.descriptionType='TechnicalInfo')].description

AIR Contacts

Documentation concepts for contacts that can answer questions not addressed in the metadata or other documentation.

Concept	Path(s)
Distribution Contact Identifier	data[*].attributes.contributors[?(@.contributorType='Distributor')].nameIdentifiers[*].nameIdentifier
Distribution Contact Identifier Scheme	data[*].attributes.contributors[?(@.contributorType='Distributor')].nameIdentifiers[*].nameIdentifierScheme
Distribution Contact Identifier Scheme URI	data[*].attributes.contributors[?(@.contributorType='Distributor')].nameIdentifiers[*].schemeUri
Resource Contact Identifier	data[*].attributes.contributors[?(@.contributorType='ContactPerson')].nameIdentifiers[*].nameIdentifier
Resource Contact Identifier Scheme	data[*].attributes.contributors[?(@.contributorType='ContactPerson')].nameIdentifiers[*].nameIdentifierScheme
Resource Contact Identifier Scheme URI	data[*].attributes.contributors[?(@.contributorType='ContactPerson')].nameIdentifiers[*].schemeUri
Rights Holder Identifier	data[*].attributes.contributors[?(@.contributorType='RightsHolder')].nameIdentifiers[*].nameIdentifier
Rights Holder Identifier Scheme	data[*].attributes.contributors[?(@.contributorType='RightsHolder')].nameIdentifiers[*].nameIdentifierScheme

4

erin@metadatagamechangers.com
<https://orcid.org/0000-0001-9998-0114>

ted@metadatagamechangers.com
<https://orcid.org/0000-0003-3585-6733>



Rights Holder Identifier Scheme URI	data[*].attributes.contributors[?(@.contributorType='RightsHolder')].nameIdentifiers[*].schemeUri
Rights URI	data[*].attributes.rightsList[*].rightsUri

Acknowledgement

This work was initially funded by the U.S. National Science Foundation 1443062

Metadata Standards, Concepts, Dialects, and Recommendations

From Gordon, S., & Habermann, T. (2018). The influence of community recommendations on metadata completeness. In *Ecological Informatics* (Vol. 43, pp. 38–51). Elsevier BV. <https://doi.org/10.1016/j.ecoinf.2017.09.005>

Communities that recognize the need for metadata typically address that need using one of several approaches: they either use a metadata standard proposed by a related community or organization, or they develop a community standard. In most cases, they also include a standard representation for the metadata. We refer to these representations as *metadata dialects*. These metadata dialects include concept names, definitions and associated structures. A *concept* is a general, dialect-independent term for describing a documentation entity, i.e., an element or attribute defined in XML or json. Typically, the communities or organizations that develop the standard also develop a set of recommendations for metadata content. We refer to these as *metadata recommendations*.

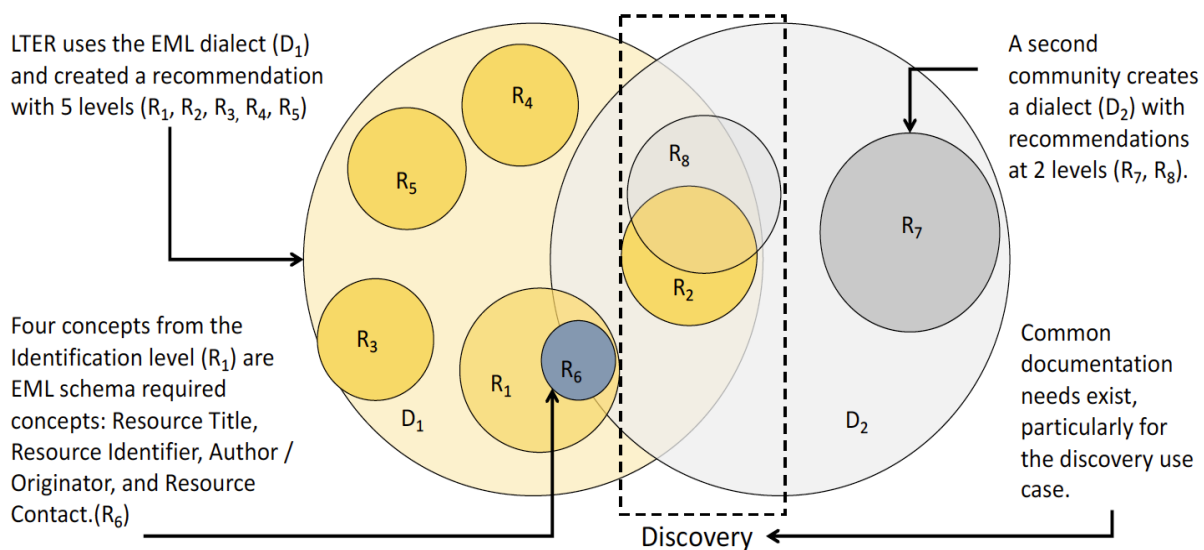


Figure 1. Metadata dialects and recommendations.

The relationship between dialects and recommendations is illustrated in Figure 1 using Long-Term Ecological Research/EML, and FGDC/CSDGM as examples. LTER uses the EML dialect (D_1) and their recommendation has five levels (R_1, R_2, R_3, R_4, R_5). All the concepts in the recommendation are included in the dialect.

In some cases, the recommended concepts are required by the XML schema used to implement the dialect, illustrated as special case R₆. There are many similar examples of metadata dialects and recommendations. When another community, like FGDC, creates a second dialect (D₂) with recommendations at two levels (R₇, R₈), there is typically overlap between the dialects (most often for discovery content) and the recommendations, e.g. R₂ and R₈ in Figure 1.