About RIF-CS

RIF-CS structure | Obligation and repeatability of RIF-CS elements | Vocabularies for Registry Schema | Syntax

The Registry Interchange Format: Collections and Services (RIF-CS) schema is a data interchange format that supports the electronic exchange of collection and service descriptions. It is the metadata format required by the Research Data Australia (RDA) Registry to enable descriptions to be harvested automatically for display and discovery in Research Data Australia.

RIF-CS is based on the international standard, ISO 2146:2010 *Information and documentation -- Registry services for libraries and related organisations*.

The advantage of the ISO 2146 information model is its strong support for a federated registry service that contains descriptive and administrative metadata for collections and related services, parties and activities, and also supports the expression of relationships between those entities. RIF-CS only includes those elements of ISO 2146 which are necessary for a collections registry, and is not fully binding to the standard.

This Content Providers Guide describes encoding requirements to prepare valid RIF-CS records for submission to the RDA Registry. The Registry applies a number of structural and content rules about information being supplied in order to ensure its validity and broad fitness for purpose. Completion of mandatory metadata elements is required but not sufficient to create a quality set of information about research data collections. See Metadata for Impact: make RIF-CS work for you for more information.

Further technical documentation and information about the governance of RIF-CS can be found at RIF-CS schema:

- Schema Guidelines
- Overview Diagram
- Controlled Vocabularies
- Schema Documentation

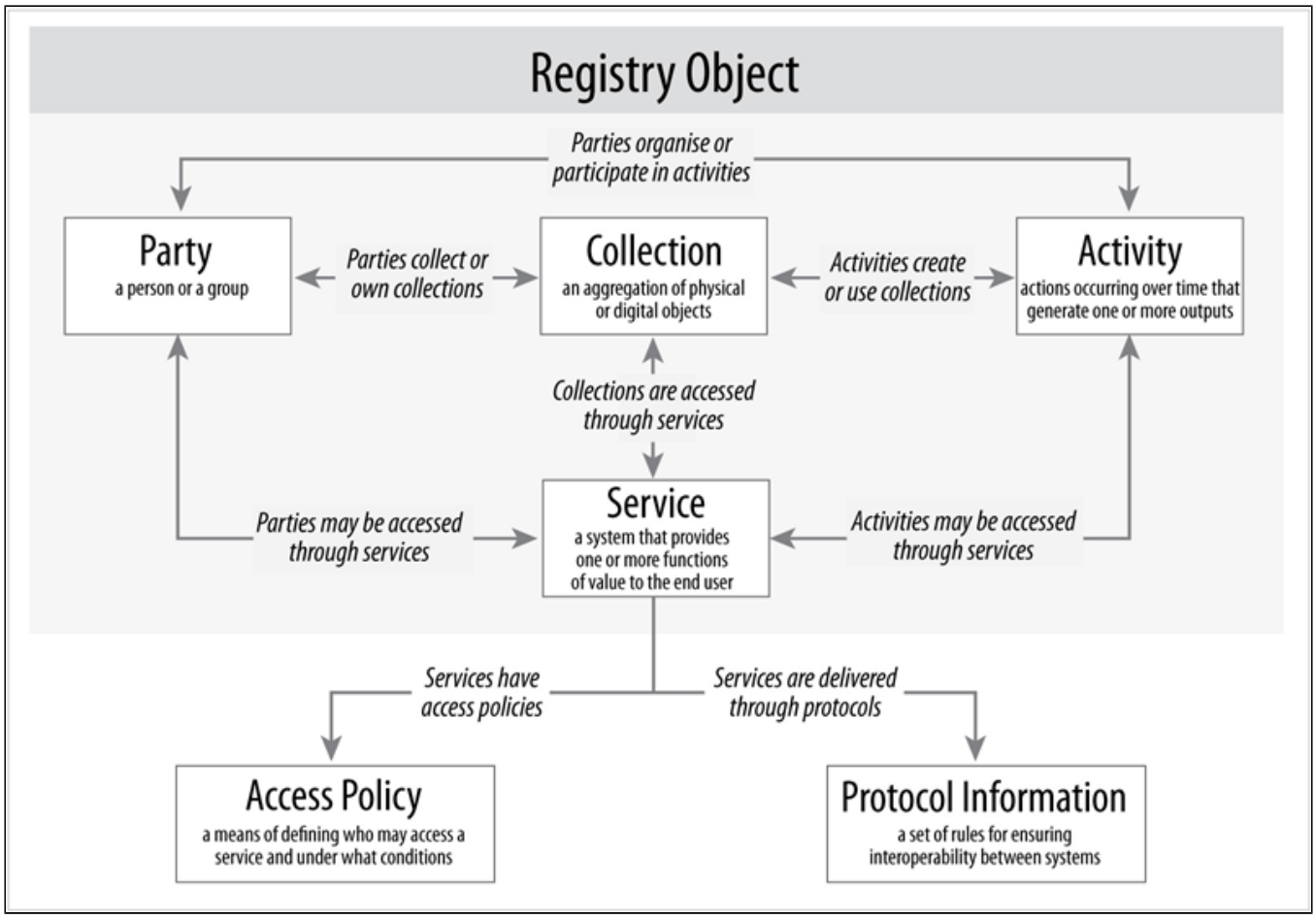
If your organisation creates or captures metadata using another schema such as ANZLIC, ISO19115, CKAN or Dublin Core, a crosswalk to RIF-CS may be available to automate the conversion of your existing records to valid RIF-CS on ingest to the RDA Registry. Assistance with crosswalks is available.

RIF-CS structure

There are four different kinds of ISO 2146 objects. These are:

- Collection: an aggregation of physical or digital objects;
- Party: a person or group;
- Activity: something occurring over time that generates one or more outputs; and
- Service: a physical or electronic interface (for example, an RSS feed) that provides its users with benefits such as work done by a party or access to a collection or activity.

These objects can be related to one another, as the following diagram shows:



Each object can be related to any number of other objects (zero, one or many). For example, a collection record may be related to several party records that describe researchers and/or institutions.

Each RIF-CS document can describe one or many registry objects.

Each object can be either a collection OR an activity OR a party OR a service.

Each object is described by one or more elements.

Obligation and repeatability of RIF-CS elements

Obligation refers to whether a particular RIF-CS component is required or optional.

The number of elements that each object can have (zero, one or many) is defined by the Schema Documentation (these rules are enforced as part of validating RIF-CS XML documents at the point of ingestion into the RDA Registry).

For example, a collection object may have two identifiers: a persistent identifier (such as a DOI) and a local repository identifier. These can both be recorded against the same object in the RIF-CS document.

Attributes cannot be repeated within an element. If necessary, attributes can be repeated by adding another instance of their element, with the exception of the following:

- activity, collection, party or service (only one per metadata record)
- date modified (only one per metadata record)
- date accessioned (only one per metadata record).

Full details of repeatability are contained in the Schema Documentation.

Vocabularies for Registry Schema

Each RIF-CS element has attributes which refine that element's broader meaning. Usually these attributes are 'Types', for example 'Identifier Type' or 'Description Type'.

There is a suggested vocabulary (a standard list of terms) for most RIF-CS attributes in the RDA Registry, as consistent terminology improves the precision of searching.

However, for most elements, the RDA Registry also allows user-defined vocabularies, so that if you have useful local terms already stored in existing systems, e.g. those appropriate to specific disciplines and communities, these can be used. Ad hoc use of user-defined terms is not recommended.

Versioning of vocabularies

The RIF-CS vocabularies are not subject to version control. Vocabulary changes or the addition of new vocabulary types occur on an ongoing basis, following confirmation by the RIF-CS Advisory Board. User-defined terms are allowed for most elements, therefore use of deprecated vocabularies will not cause any problems for contributors. If you would like a new term added to a vocabulary, email services@ardc.edu.au.

Changes to vocabularies are not linked or tied to RIF-CS versions in the RDA Registry or in Research Data Australia, although those may occur at the same time as changes in the schema.

Syntax

Syntax refers to the structure in which elements must be encoded (or expressed). RIF-CS is an XML schema; XML (eXtensible Markup Language) is a markup language used to electronically encode and exchange information.

The RIF-CS Schema Documentation describes the required structure and content of a valid RIF-CS XML document.

The required syntax of some elements is determined by the Type chosen from the suggested vocabularies. This is particularly important for information such as spatial coordinates, as correct syntax must be used to generate map displays. Refer to the references provided for those Types in this Guide for correct formatting of element values.

XHTML formatting can be provided for text in the Description element and the AddressPart element, to support better display in Research Data Australia.