Introduction

Benjamin Riesenberg

The purpose of this issue of *Library Technology Reports* is to provide a brief introduction to metadata application profiles (MAPs) and to present information about what they do, how they are put together, and how they are changing. In providing examples of MAPs currently in use and development, the issue also shows how they fit into the larger context of library work, specifically within technical services work taking place in cataloging and metadata units at many institutions.

Two Broad Categories

One challenge in discussing metadata application profiles is taking into account the diversity of practices that they can be used to define. When you consider that each collection of information resources is different, that resources and their descriptions are managed and accessed using myriad software applications, and that information is stored in a growing and evolving array of data formats, it is easy to see that permutations of these factors add up to an overwhelming variety of scenarios.

In much of the discussion here, this wide variety of metadata implementations has been simplified into two broad groups: linked-data implementations and everything else. Both categories can involve different descriptive practices, data formats, and software applications, but considering practices as they fall into these two broad categories is helpful in discussing many aspects of metadata and metadata modeling, and it is often necessary to differentiate between them when discussing MAPs. For example, the ways that entities—the resources that are described by metadata—are handled (or not) in linked-data applications versus non-linked-data applications can be quite different.

For readers who want to acquire basic knowledge of the Resource Description Framework (RDF), the following texts offer helpful starting points. The "RDF 1.1 Primer" in particular is an accessible introduction to the basic concepts and provides helpful examples.

Schreiber, Guus, and Yves Raimond, eds. "RDF 1.1 Primer." W3C Working Group Note, World Wide Web Consortium. June 24, 2014. https://www.w3.org/TR/rdf11-primer.

Cyganiak, Richard, David Wood, and Markus Lanthaler, eds. "RDF 1.1 Concepts and Abstract Syntax." W3C Recommendation, World Wide Web Consortium. February 25, 2014. https://www.w3.org/TR/rdf11-concepts.

A certain level of familiarity with essential linked-data concepts is assumed in these chapters, but in-depth knowledge of data models, implementations, syntaxes, and applications is not required. The Resource Description Framework (RDF) is a primary—but not the only—data model underlying linked-data and Semantic Web practices, and this term is used often throughout the issue.

Metadata

A MAP may go by different names—variations such as *application profile* or simply *profile*, or alternate terms such as *data dictionary* or even *data model*. A MAP can provide a model—that is, a detailed definition of the standards, tools, and practices used to provide description and access for information resources. A MAP is not only detailed, but also complete, incorporating the three primary things that most metadata

```
<metadata>
  <record>
    <displayedAt>https://upload.wikimedia.org/wikipedia/commons/3/39
    /PDX _ palms.jpg</displayedAt>
    <title>Northwest U.S. Palm Trees</title>
    <creator>Riesenberg, Benjamin</creator>
    <location>Portland, Oregon</location>
    <date>2016-04-17</date>
    <type>StillImage</type>
  </record>
</metadata>
```

Figure 1.1
An XML metadata instance, showing metadata elements and values

```
@prefix ex: <http://example.org/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix schema: <https://schema.org/> .
@prefix dct: <http://purl.org/dc/terms/> .
@prefix dce: <http://purl.org/dc/elements/1.1/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

ex:a _ photo rdf:type schema:MediaObject .
ex:a _ photo schema:contentUrl
<https://upload.wikimedia.org/wikipedia/commons/3/39/PDX _ palms.jpg> .
ex:a _ photo dct:title “Northwest U.S. Palm Trees”@en .
ex:a _ photo dce:creator “Riesenberg, Benjamin”@en .
ex:a _ photo dct:spatial <http://vocab.getty.edu/tgn/7014273> .
ex:a _ photo dct:date “2016-04-17”^^xsd:date .
ex:a _ photo dct:type <http://purl.org/dc/dcmitype/StillImage> .
```

Figure 1.2
An RDF metadata instance, showing triples with subjects, predicates, and objects

instances—that is, a given quantity of metadata—include: the entities described, the properties used to describe them, and the values provided for these properties. While each of these will be covered in detail in subsequent chapters, it would be difficult to begin describing the MAP-making process without first saying a bit about each.

Entities

Metadata describes entities. Records in a library's online public-access catalog describe printed books and other physical media on the shelves, e-books available through a web browser, and many more types of resources; metadata in a digital-collections platform describes images, sound recordings, documents, and more; attributes and values in an online shopping website describe products for sale.

In non-RDF metadata implementations, the entities themselves may be somewhat difficult to find in a metadata instance. In some digital-collections platforms, for example, the entity—for example, a digital image—doesn't really exist in the metadata, only in the user interface. We may export metadata from such a system and not really see the entity, only the statements made about it. An XML metadata export, for example, may look something like figure 1.1.

In RDF implementations, the entities being described are somewhat easier to find. Because RDF uses a “triple” structure to make statements, with each triple including a subject, predicate, and object, we can clearly identify the entity being described in each triple. This may be easier or harder based on the serialization or syntax used to encode the data, but in the example in figure 1.2, which uses Turtle syntax, we can see a subject—an entity—quite clearly in each statement.[1]

Properties

Properties, which may also be referred to using terms such as *attribute* and *field*, are descriptive elements

(elements is another term used here and elsewhere to refer to them) used to make statements about entities. The nature of properties, like that of entities, also differs in linked-data and non-linked-data implementations. Most user interfaces presenting metadata will provide textual labels for properties so that users can understand the resource descriptions. But exports or other views of a metadata instance underlying these displays will vary widely.

In an RDF metadata instance, each property must be identified by either a Uniform Resource Identifier (URI) or an Internationalized Resource Identifier (IRI; the term we will most often use here). Best practices for creating and publishing RDF vocabularies require that these IRIs can be looked up (dereferenced) via the World Wide Web to retrieve information about the resources they identify. User interfaces presenting RDF data to humans will most likely display textual labels for properties, but in the underlying data these properties will be identified by machine-actionable IRIs. In figure 1.2 we see an entity, property, and value in each triple. Entities, properties, and some values are IRIs—often prefixed names, such as rdf:type, used in place of full IRIs—while other values are textual or numeric.

Values

The properties we use to describe entities are somewhat meaningless unless we provide values for them. MAPs can define values in a metadata instance by putting constraints in place. Examples of these constraints are requirements that all books have one and only one value for the title property, that values for a creator property be recorded using a "[last name], [first name]" format, or that creator values be taken from the Library of Congress Name Authority File wherever available.

Depending on the nature of an implementation, these constraints may include human-readable instructions for creating or transcribing values, specifications for formats to which values must adhere, or specific sources from which values must be taken.

Note

1. Eric Prud'hommeaux and Gavin Carothers, eds., "RDF 1.1 Turtle: Terse RDF Triple Language," W3C Recommendation, World Wide Web Consortium, February 25, 2014, https://www.w3.org/TR/turtle.