Benefits of Machine Readable Curricula

Nigel Ward, Link Affiliates

Nick Nicholas, Link Affiliates

Date: 10 March 2010

This version:
http://www.linkaffiliates.net.au/Publications/BenefitsMachineReadableCurricula_Mar10.html
Latest version:
http://www.linkaffiliates.net.au/Publications/BenefitsMachineReadableCurricula/
Previous version:
http://www.linkaffiliates.net.au/Publications/BenefitsMachineReadableCurricula_Mar10.html

Purpose

This is an executive level briefing on the potential benefits of machine readable curricula, as is currently being trialled by Curriculum Corporation (http://www.curriculum.edu.au/) in partnership with the Achievement Standards Network (http://www.achievementstandards.org/).

Background

- Machine readable curriculum techniques move away from managing curricula as prose documents. Instead, they impose a lightweight and flexible structure over a curriculum: decomposing it into a collection of *statements* (e.g. learning areas, strands, outcomes). Each statement is individually identified, described and related using constructs understandable by computers.
- U.S. and European initiatives are developing and using machine readable curricula.
- Curriculum Corporation is engaged in work supporting the national curriculum work of ACARA through IT infrastructure. As part of this work, they are investigating the potential advantages of machine readable curriculum, building on an approach used in the U.S.

Assumptions

- Uptake of the national curriculum in schools will be gradual: in the short to medium term, jurisdictional curricula will co-exist with the national curriculum.
- Teachers will wish to discover resources and contextualise learning using outcomes from both national and local curricula.
- Some agencies will extend the curriculum with more detailed descriptions. (For example, the curriculum may specify numeracy, but an online assessment may be described as specific to division.)

Summary of advantages of machine readable curricula

In the **short term**, machine readable curriculum enables:

1. Learning resources to be easily and unambiguously tagged with relevant learning outcomes. For example, learning content can be tagged with the learning outcomes it might support, an assessment can be tagged with the learning outcomes it assesses.

2. A student's progress to be easily and unambiguously mapped to the curriculum. For example, evidence in a student's e-Portfolio can be tagged with relevant learning outcomes.
3. Lesson plans and learning paths to be easily and unambiguously mapped to the curriculum. For example, a learning path proposed by a professional body can nominate a sequence of outcomes recommended to realise desirable competencies.
4. The curriculum to drive content discovery: teachers want to find online resources matching particular curriculum outcomes they are teaching.
5. Other e-learning applications to be configured to use the curriculum structure to share information. For example, a portal can be configured to show learning content relevant to a learning outcome.

More complex use of machine readable curricula could also provide the following **longer term** advantages:

1. Managing versions of the curriculum over time becomes easier. For example, it is possible to update individual statements in the curriculum without having to re-publish a whole document; the relationships between new and old statements can be tracked; and the description of an outcome can still be discovered and used even after it has been superseded.
2. Statements in different curricula can be compared, and unambiguously mapped to each other. The mapping is more straightforward with identified statements than with a traditional curriculum document. The mapping need not be restricted to "outcome A is the same as outcome B": mapping can encompass similarity, specialisation ("outcome A is a special case of outcome B"), prerequisites, alternates, and whatever else may be appropriate.
3. Mapping between curricula allows gap analysis: once the distinct outcomes of two curricula are identified and correlated, any gaps in coverage of one or the other can be identified more clearly. The comparison can be substantially automated.
4. Agencies can unambiguously extend the curriculum to the level of granularity needed. For example, the curriculum may specify numeracy, but a publisher's piece of assessment may be described as specific to division. The publisher can specify division as a specialisation of numeracy.

ASN machine readable approach

Curriculum Corporation has been working with the Achievement Standards Network (ASN) to explore the applicability of ASN's technical solution to the Australian context. ASN is already in use in the United States, where curriculum alignment between the States is a pressing issue for content providers who must describe their content in multiple state curricula. (The U.S. has no overarching national curriculum.)

The ASN approach imposes a lightweight and flexible structure over a curriculum: decomposing it into a collection of *statements* (e.g. learning areas, strands, outcomes). Each statement is individually identified, described and related using constructs understandable by computers.

ASN differs from other machine readable approaches in that it uses emerging best practices for publishing and connecting structured data on the Web (the Linked Data Initiative http://linkeddata.org/). The ASN solution relies on three major components:

- *Resolvable Persistent Identifiers* (URIs) for curriculum outcomes.
- A framework based on the Dublin Core Metadata Initiative's (DCMI) syntax-independent abstract information model (DCAM).
- Use of *Semantic Web* technologies (in particular RDF) to align and describe curriculum outcomes.

The identifiers allow discrete curriculum statements to be identified unambiguously and reusably. This leads to immediate benefits wherever curriculum outcomes need to be referred to in an e-learning context: enabling curriculum-driven discovery of learning resources, and machine-readable formulation of lesson plans and learning paths. Because the identifiers are linked to online descriptions of the curriculum outcomes, users can still recover the original curriculum statement.

The ASN framework enables the curriculum to be described using interoperable metadata that is familiar to resource cataloguers, and includes Australian specific descriptions (such as Schools Online Thesaurus subject terms) while also supporting international contexts.

Other benefits of the machine-readable approach are longer term. The Semantic Web provides a flexible mechanism for mapping curricula to other curricula—in particular, State curricula to the National curriculum, and earlier to later versions of a curriculum. Once such mapping is in place, it is possible to realise new use cases, such as gap analysis, evaluating assessment, negotiating different granularities and scopes of curricula, and flexible approaches to metadata.

Finally, by making the identifiers persistent, it becomes possible to access and use historical data on curricula, and to introduce change management.

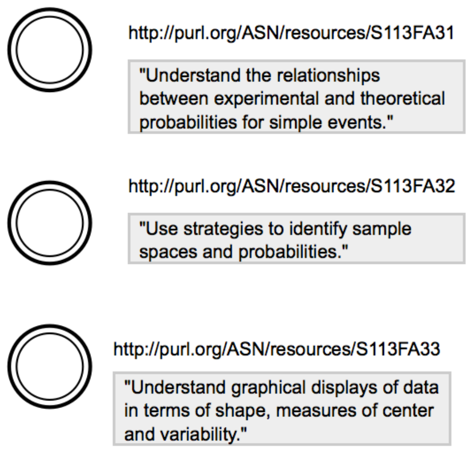
Immediate benefits: Identifiers

ASN analyses the textual statements of State curricula, identifies discrete statements of outcomes and skills developed, and assigns each a persistent identifier (PURL). Each distinct outcome in a curriculum has an associated URL, which resolves to an online description of the outcome.

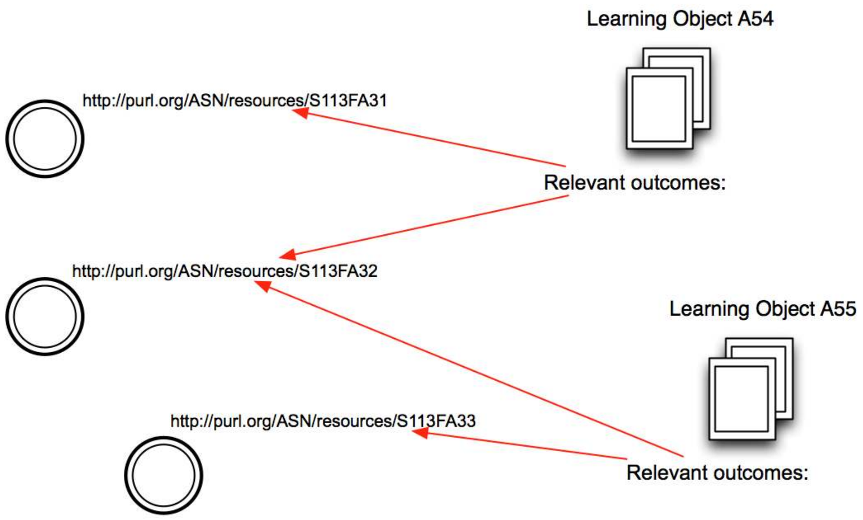
The persistent identifier achieves the following benefits:

- Distinct outcomes can be identified succinctly and unambiguously, without having to parse the text of the outcome statement.

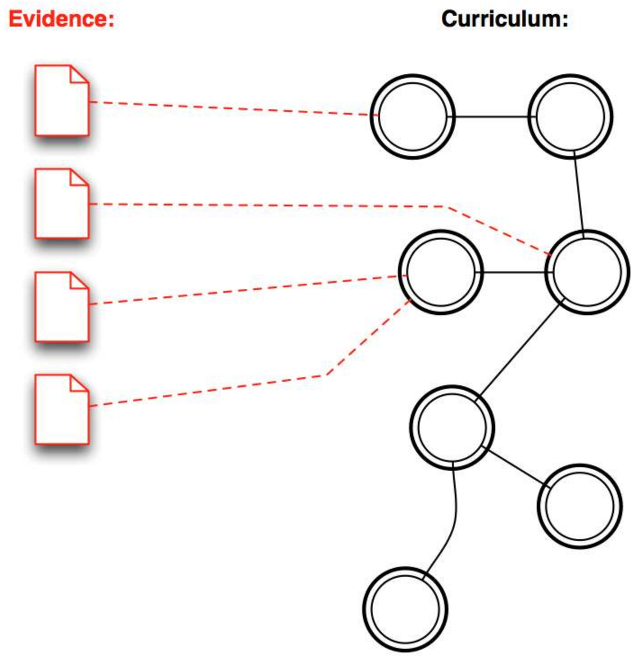
(Outcomes, Mathematics curriculum, North Carolina)



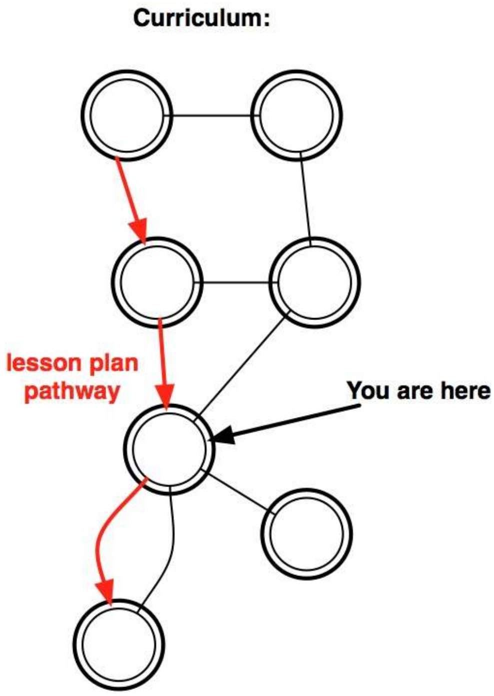
- Learning resources can be tagged as satisfying outcomes efficiently and unambiguously: the unique identifier is included in the learning object metadata. That means that learning objects can now be searched according to what curriculum outcomes they satisfy. This solution has already been taken up by the IMS Common Cartridge specification for describing learning content, and is being piloted.



- A student's progress through the curriculum can be machine readable, which can be used to integrate different systems dealing with the same student's progress. For example, a Learning Management System can populate the e-portfolio of a student with evidence of achievement automatically linked to the outcomes realised.



- Third parties can nominate a sequence of identified outcomes as a desirable path through the curriculum. A teacher's lesson plan can draw on nominated curriculum outcomes, and the identifiers used to retrieve relevant resources. A professional body can propose a learning path, recommending a sequence of curriculum outcomes to realise desirable competencies (for example to attain a trade licence or membership of a professional body).



- The identifier for a curriculum outcome provides easy access to the original curriculum statement: the curriculum outcome's URI can be used to retrieve information about the outcome using a web browser.

Examples of ASN identified curriculum statements:

(Outcomes, Mathematics curriculum, North Carolina)

Identifier	Outcome	Grade
http://purl.org/ASN/resources/S113FA31	Understand the relationships between experimental and theoretical probabilities for simple events.	6
http://purl.org/ASN/resources/S113FA32	Use strategies to identify sample spaces and probabilities.	6
http://purl.org/ASN/resources/S113FA33	Understand graphical displays of data in terms of shape, measures of center and variability.	6

Long-Term benefits: Persistence

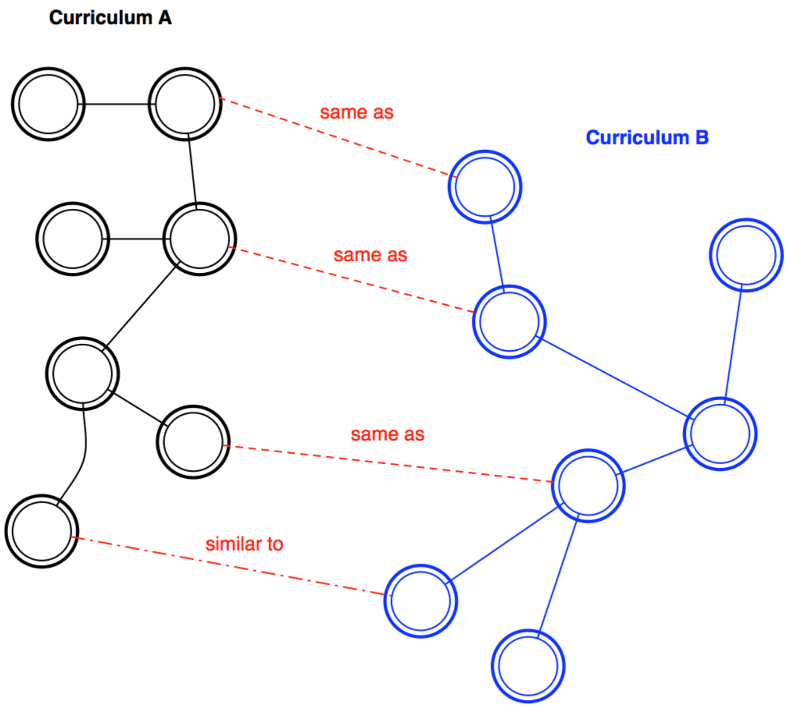
The identifier for a curriculum outcome can be made persistent: the identifier can still be used, and resolved to an online description, even after it has been superseded in the current curriculum.

This makes change management of curricula possible. If a revised statement is felt to be different from the original, it is assigned a distinct identifier, and the revision history can be tracked. The identifier remains resolvable after the original curriculum has been superseded, which is important for historical contexts, e.g. determining a student's or employee's qualifications.

Making a curriculum machine-readable, and disaggregating it into discrete statements, makes it a different kind of object to manage than the prose text of traditional curriculum documents. In particular, change management of individual outcomes becomes critical, as so many systems come to depend on the curriculum. That requires accountability to be in place for curriculum changes, and to be accessible to users of various systems.

Long-Term benefits: Mapping curricula

Once identifiers of outcomes are available, the outcomes in different curricula can be compared, and mapped to each other. The mapping is more straightforward with identifiers than with the source text of curricula.



Aligning and mapping curricula to each other under the machine-readable approach relies on Semantic Web technologies. The *Semantic Web* allows metadata about resources, and the relations between resources, to be expressed in a semantically rich way, allowing more powerful searches to be carried out. RDF in particular allows relations between two resources to be expressed as a graph using identifiers for the resources related, and for the relation between them. An intelligent search engine can reason about the various resources related through RDF, and come up with relations that have not been coded explicitly.

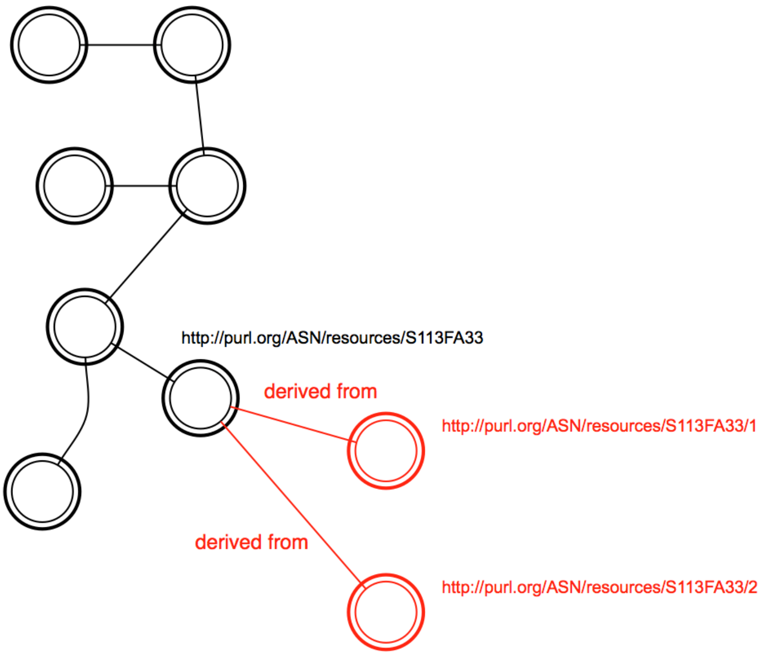
This allows an open-ended and powerful vocabulary for aligning outcomes with each other: the alignment need not be restricted to "same as", but can encompass similarity, specialisation, prerequisites, alternates, and whatever else may be appropriate. These relations can be explored and searched through intelligent search engines.

Long-Term benefits: Granularity and metadata

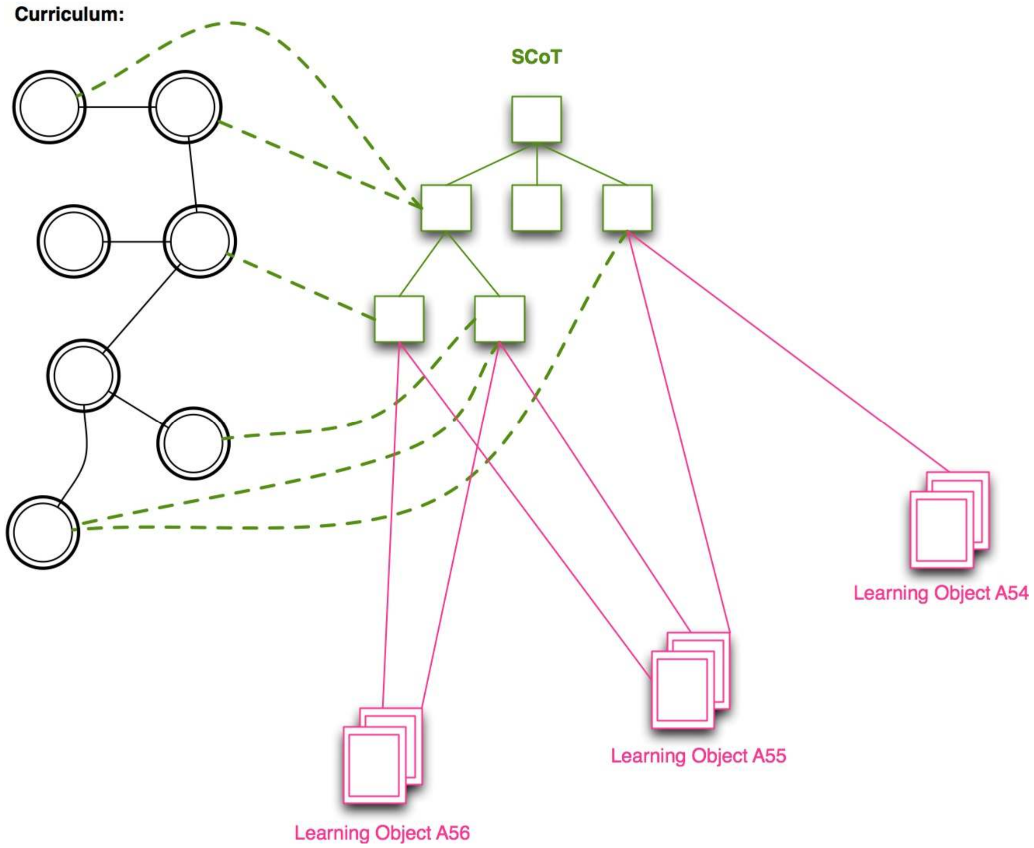
Because the relations expressed through RDF are dynamic, differences in granularity and scope of outcomes between different agencies can be addressed, without dismantling the existing repertoire of curriculum outcomes. If an ad hoc outcome is needed to describe a piece of assessment more specifically, it can be formulated in a way that search engines and other systems can deal with. This is done through derived statements: more specific curriculum outcomes used to describe resources, but not present in the published curricula. RDF relates those derived statements back to the published curriculum. Because they are related back, a search for the published curriculum outcome can still discover these more specific outcomes.

Example:

http://purl.org/ASN/resources/S113FA33	Understand graphical displays of data in terms of shape, measures of center and variability.	Original statement
http://purl.org/ASN/resources/S113FA33/1	Understand graphical displays of data in terms of shape.	Derived statement
http://purl.org/ASN/resources/S113FA33/2	Understand graphical displays of data in terms of measures of center.	Derived statement



RDF also allows third parties to create new metadata about curriculum statements, without being constrained by the structure of the original outcome descriptions. For example, Curriculum Corporation has been exploring associating subject terms from the Schools Online Thesaurus (ScOT) with curriculum statements. Given that Curriculum Corporation learning content is already tagged with ScOT terms, this provides a convenient path for mapping between curriculum outcomes and curriculum content.



Use cases

The combination of these two approaches (identifiers + semantic web) makes possible a range of use cases that are difficult without machine readable curricula, and which can enhance the learner's interaction with the curriculum substantially. These include:

- The curriculum can be incorporated explicitly into content classification and discovery: the curriculum has a direct impact on how teachers design and assess their courses.
- Gap analysis of curricula against each other: once the distinct outcomes of two curricula are identified and correlated, any gaps in coverage of one or the other can be identified more clearly; the comparison can be substantially automated.
- Transfer of assessment between states: a student's assessment according to one state's curriculum criteria can be translated to another state's, once the state curricula are aligned to each other and to the national curriculum. Again, this translation can be substantially automated, and any required remedial work can be identified quickly.
- The explicit change management of curricula at the level of individual outcomes means that updates to courses can reflect updates to the curriculum more closely, without guesswork or the risk of slippage.

Advantages of semantic web over traditional metadata approaches

This section is more technical. It provides more detail on the Semantic Web approach to resource description.

Traditional metadata approaches describe a resource using a hierarchical structure contained in a metadata record. For example an XML binding of IEEE LOM (Learning Object Metadata) record for a resource might contain the following statements:

```
<identifier>
      <catalog>TLF</catalog>
      <entry>L7855</entry>
</identifier>
<IntendedEndUserRole>learner</IntendedEndUserRole >
<Relation>
      <kind>IsPartOf<kind>
      <resource>
            <identifier>
                  <catalog>TLF</catalog>
                  <entry>L7853</entry>
            </identifier>
      <resource>
</Relation>
```

These statements say that the resource is identified as TLF L7855, that it is intended for a learner, and that it is part of another resource called TLF L7853.

"RDF triples" are the information structure underlying the Semantic Web. Instead of using hierarchical structured metadata records, RDF treats metadata at its most basic level: as simple propositions—sentences containing claims about the resource. The propositions are a combination of three things:

- the *subject*: what the proposition is about.
- the *predicate*: what property is being claimed.
- the *object*: what the value of the property is.

So, for example, the LOM XML fragment above might be expressed in RDF as

```
<http://tlf.edu.au/L7855>
    <http://ieee.org/lom/IntendedEndUserRole>
    "Learner"

<http://tlf.edu.au/L7855>
    <http://ieee.org/lom/relation/IsPartOf>
    <http://tlf.edu.au/L7853>
```

One difference between the RDF and LOM XML versions of this metadata is that each RDF statement explicitly identifies the resource being described (the subject) using a URI (`http://tlf.edu.au/L7855`). This is because, unlike the LOM XML, RDF doesn't require that the statements are collocated in a metadata record: the statements can be stored in different documents and can even be made by different parties.

Another difference is that the predicate, and sometimes the object of an RDF statement are also identified with URIs. This allows the metadata to be self-documenting: an important aspect of the Semantic Web is that anything talked about has a URI that users can look up in a web browser to find useful information. To illustrate, if terms from the intended end user role vocabulary are assigned URIs, then the first RDF statement can be re-expressed as

```
<http://tlf.edu.au/L7855>
    <http://ieee.org/lom/IntendedEndUserRole>
    <http://ieee.org/lom/UserRoleVocab/Learner>
```

A user can then look up the `http://ieee.org/lom/UserRoleVocab/Learner` URI to discover a definition of "Learner".

Conventional metadata approaches capture resource descriptions in well-structured XML documents, which impose hierarchical structure and strict type-checking. This approach works well where content is managed in repositories and metadata is strictly controlled. This approach, however, has largely been ignored on the "Open Web", and such metadata records are usually ignored by search engines such as Google.

In contrast, the Semantic Web allows anyone to publish metadata statements about a resource, and to create networks of linked statements. The SPARQL RDF query language allows very powerful queries to be made across these networks of linked statements. For example, "what are all the curriculum outcomes and learning resources related to the topic of astronomy?" In fact, logical inferencing becomes possible (e.g. through rules on how properties relate to each other), so that queries can exceed the capabilities of relational databases.

Another advantage of RDF is that it is extensible: new statements about existing resources can be made by anyone. Additionally, statements can be related to each other, so for example, it is possible to indicate that the curriculum outcome "Understand graphical displays of data in terms of shape" is a refinement or derivation of the curriculum outcome "Understand graphical displays of data in terms of shape, measures of center and variability".

As a result of these advantages, metadata initiatives like Dublin Core and ISO Metadata for Learning Resources are starting to use RDF alongside or even instead of XML schemas.

For machine-readable curricula, the network of relations between curriculum outcomes is complex and extensible, which makes an RDF approach essential:

- The large number of curriculum outcomes, and the ambiguity of the natural language expressing them, requires large scale use of identifiers, so that each outcome can be treated as its own resource.

- Outcomes from different curricula are not always merely equivalent: outcome A might be a generalisation of outcome B, it might be similar, a prerequisite, and so forth. This requires a flexible vocabulary of properties, and powerful inferencing queries to negotiate their logical relations.
- Outcomes are organised into structures, such as year levels and subject areas. These structures are best expressed as logical networks (trees), which RDF is suited to.
- Curriculum outcomes can be reorganised into different sequences, as required by local learning plans or suggested by professional bodies. RDF allows the flexibility of putting more than one structure over outcomes.
- The open-endedness of RDF allows different kinds of metadata to be mixed in with curriculum outcomes. For example, it is very easy to extend an RDF description of a curriculum outcome with subject terms from the Schools Online Thesaurus (ScOT), which allows a search engine to link the curriculum outcomes directly to curriculum content tagged with the same ScOT terms.

Further Reading

- Briefing Paper, Curriculum Description (Subproject, Technical Standards project): http://linkaffiliates.net.au/Activities/BriefingPapers/BP-Curriculum.html. Scenarios, business requirements met through the machine-readable approach.
- http://blog.linkaffiliates.net.au/2009/07/20/national-curriculum-machine-readable/: Blog overview of National Curriculum work, including why machine readable curricula are useful. Less detailed version of the picture painted in the Briefing Paper, though with more detail on ASN.
- http://blog.linkaffiliates.net.au/2009/08/18/ims-global-meeting-curriculum-standards/: Blog posting on IMS Common Cartridge work with ASN, using curriculum identifiers to tag content. Includes discussion of using RDF to navigate the varying granularity from different agencies of curriculum outcomes.

Acknowledgements

This briefing paper was prepared by Link Affiliates as part of the *Technical Standards for Digital Education project* funded by the Australian Government's Department of Education, Employment and Workplace Relations (DEEWR).

The authors wish to thank Helen Smith (Curriculum Corporation), Nick Weidemann (Curriculum Corporation) and Diny Golder (JES & Co.) for their feedback on drafts of this paper.