LD4PE: A Competency-based Guide to Linked Data Principles and Practices

Stuart A. Sutton
Dublin Core Metadata Initiative, USA
sasutton@dublincore.net

Michael D. Crandall
University of Washington, USA
mikecran@uw.edu

Marcia Zeng
Kent State University, USA
mzeng@kent.edu

Thomas Baker
Dublin Core Metadata Initiative, Germany
tom@tombaker.org

Abigail Evans
University of Washington, USA
abievans@uw.edu

Sean Dolan
Kent State University, USA
sdolan5@kent.edu

Joseph Chapman
D2L, Ltd., USA
Joseph.chapman@D2L.com

David Talley
Precise Recall, USA
dtalley@preciserecall.com

Michael Lauruhn
Elsevier Labs, USA
M.Lauruhn@elsevier.com

Abstract

The IMLS-funded project *Linked Data for Professional Education* (LD4PE) has prototyped a competency-based referatory of Learning Resources related to the design, implementation, and management of Linked Data (a referatory is a website that points to and describes learning resources.) The project developed: 1) A "Competency Index for Linked Data" (*Index*), which characterizes the Linked Data field in terms of formally identified competencies usable for tagging resources, as a basis for self-guided learning or for designing curricula. 2) A tool set for creating metadata about Learning Resources, packaging user-selected Learning Resources in Saved Sets, and creating learning trajectory maps for expressing personal or curriculum-based learning journeys through the competencies. 3) A catalog of Learning Resources mapped to the competencies. 4) A project website, http://explore.dublincore.net. 5) Best Practices describing policies reusable by other projects using competency-based description and discovery of learning resources.

Keywords: Linked Data, Learning Resources, competency frameworks, ASN-DL, competency-based teaching and learning.

1. Introduction

Understanding Linked Data standards and practices has become a key requirement for information professionals in galleries, libraries, archives and museums (GLAM). Major national libraries and bibliographic services are leading the trends toward publishing authority files, catalogs, datasets, and bibliographic standards as Linked Data, and toward aggregating Linked Data from external sources such as dbpedia and MusicBrainz to enhance discovery and retrieval. Broad initiatives such as LODLAM and OpenGLAM promote the integration of Linked Data across galleries, libraries, archives, and museums.

Cultural memory institutions find themselves on shaky ground as this paradigm shift pushes the need for competent professionals from national centers toward local institutions. The challenge of acquiring new competencies extends to teachers of the next generation of professionals and trainers who provide continuing professional development.

This urgent need to develop Linked Data competencies in the professional workforce is driving major initiatives to provide Learning Resources about the underlying standards and model of Linked Data such as the EU's Euclid project, LOD2, School for Data, PlanetData, Open Data

https://doi.org/10.23106/dcmi.952137899

Institute, Lean Semantic Web, Cloudera, GATE, *Linked Data Cookbook*, and *Cookbook for Translating Data Models to RDF Schemas*. The products of these initiatives range from curricular structures and full courses to simple "recipes"—brief packages of "how-to" videos and step-by-step instructions that address specific, but frequently unarticulated learning outcomes. These scattered initiatives and their resources can be easy to find by those who already know what they are looking for, but everyone else struggles to put the available resources into a context.

The website of the LD4PE project helps alleviate this struggle by supporting the structured discovery of online Learning Resources that have been made available by both open and commercial providers. At its heart is a competency framework for Linked Data practice that supports the tagging of Learning Resources according to the specific competencies they address. The competency framework, itself expressed in RDF, leverages Linked Data technology by assigning global identifiers (URIs) to statements of competence. These URIs, when used as tags in metadata about Learning Resources, map the Learning Resources to nodes in the *Index*.

Deliverables of the LD4PE project include:

1. *Competency Framework*. RDF-modeled "Competency Index for Linked Data" (*Index*) based on the Achievements Standards Network Description Language (ASN-DL) for describing formally promulgated competencies and benchmarks.
2. *Toolkit*. An openly available, Web-based tool set to support the management of the *Index*; the generation of RDF metadata about Learning Resources; the packaging and arrangement of selected Learning Resources by users in Saved Sets; and the creation of learning trajectory maps expressing curricular structures or personal learning journeys superimposed over the competency framework through the integration of these elements as WordPress custom posts and taxonomies on the LD4PE website.
3. *Learning Resource Descriptions*. Metadata about Learning Resources mapped to the competencies and benchmarks of the *Index* in support of competency-based resource discovery by teachers, trainers and learners.
4. *LD4PE Website.* A website (http://explore.dublincore.net), ownership of which is being transferred to the Dublin Core Metadata Initiative (DCMI) as part of its educational agenda.
5. *Best Practices*. Readily accessible best practice documentation for all LD4PE development and maintenance processes, from the creation of a community-based competency framework development to the creation of metadata about Learning Resources and learner trajectories.

2. Competency Index

Competency frameworks defining what a learner should know and be able to demonstrate underpin labor market credentialing mechanisms including college and university degrees, educational certificates, industrial certification, occupational licenses and even emerging micro-credentialing through digital badging. Learning environments strive for a tight coupling between these controlling frameworks, the resources necessary to achieve the learning objectives embodied in the frameworks, and what is actually assessed in terms of learner competence (Sutton & Golder, 2008).

2.1. *Index* Development

Government and quasi-governmental agencies, commercial entities and professional organizations develop most such competency frameworks under the tight editorial control of selected experts. In contrast to such highly constrained contexts, the LD4PE approach aims at crowd-sourcing expertise both to develop and to assess the project's competency framework, with synthesis and finalization of the published framework in the hands of a small, volunteer Editorial Board comprised of subject matter experts in the domain of Linked Data. During the project period, members of the project team personally solicited feedback at a number of conferences, workshops and other events, and widely disseminated a call for comments on a public version of the *Index*.

https://doi.org/10.23106/dcmi.952137899

Projects such as schema.org are demonstrating that useful outcomes can be achieved through forms of crowdsourcing when version control systems created for "social coding", such as Github, support processes for community input; this approach is discussed further in Section 2.5.

Figure 1 illustrates the various components in the LD4PE competency-creation process.

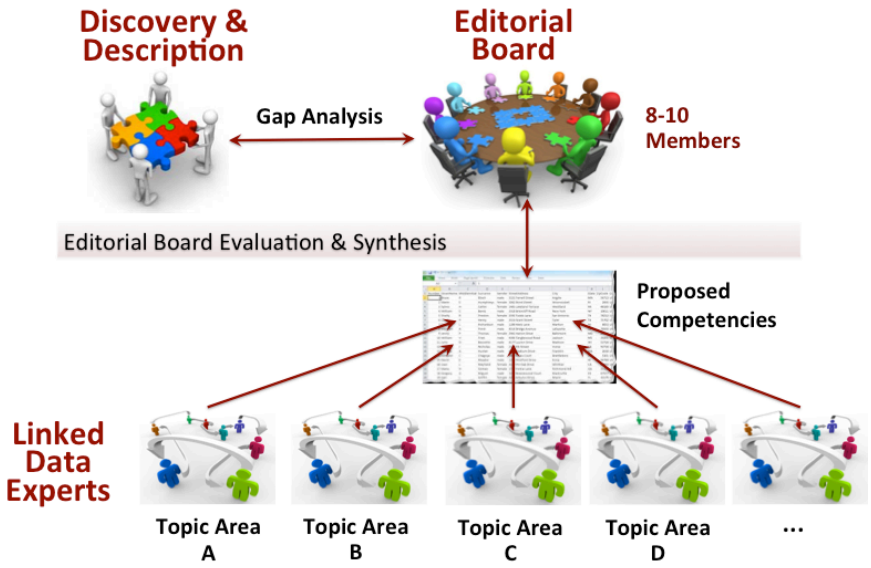

FIG. 1. Crowd-sourcing expertise for the *Index*.

Discovery and Description. The process of discovering, describing, and mapping Learning Resources to competencies involves continuous assessment of the coverage of the *Index* in terms of:

1. *orphan competencies*—competencies found in Learning Resources that cannot be mapped to the *Index*; and
2. *misaligned granularity*—where imprecise mapping options reveal a mismatch in granularity between *Index* competencies and Learning Resources.

Linked Data Expert engagement through both solicited and unsolicited proposals for new competencies provides a mechanism through a version-controlled submission process for subsequent review and synthesis into the evolving *Index* by the *Editorial Board*, a small group of domain experts who volunteer their time.

2.2 *Index* modeling

The *Index* itself is modeled as an RDF graph using the Achievements Standards Network Description Language (ASN-DL). The ASN-DL is comprised of two resource classes, *StandardDocument* for competency frameworks described as wholes, and *Statement* for individual competency assertions in the document. Object properties are used both to replicate the original hierarchical or graph modeling of the canonical document and to express the semantic relationship between individual assertions.

The ASN-DL is extensible through addition of new properties and subproperty refinements to support domain-specific profiles. Value vocabulary namespaces relevant to the jurisdiction of the competency framework may be specified and encoded as RDF using the W3C standard Simple Knowledge Organization System (SKOS).

https://doi.org/10.23106/dcmi.952137899

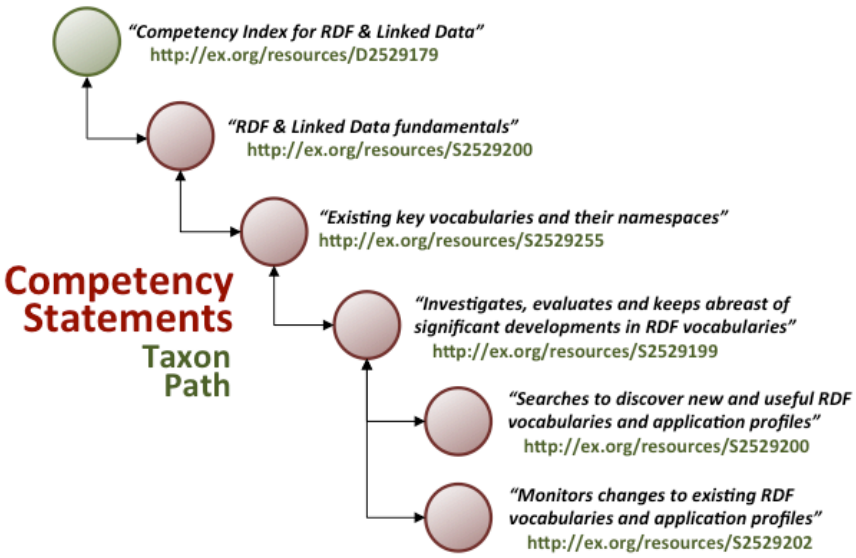

FIG. 2. Example Index ASN-DF modeling

2.3 *Index* scope

One of the major challenges during the early months of the project was operationalizing the term "Linked Data" in such a way that the scope of both the developing *Index* and the Learning Resource to be described and mapped were tractable within the context of the two-year grant. "Linked Data" denotes a method for publishing structured data on the web that builds on a set of open standards and technologies for expressing and querying of Linked Data, such as HTTP, URI, RDF, SPARQL, OWL, and SKOS. The concept of Linked Data is roughly characterized in terms of four practices recommended by Sir Tim Berners-Lee: the use of URIs as names for things; the use of HTTP URIs that allow people to find information about those names; the provision of that information in a machine-readable form; and the inclusion of links to other relevant resources (Berners-Lee, 2006). Extending the *Index* and the body of Learning Resources described in LD4PE metadata to encompass all underlying and enabling technologies for Linked Data would have been unrealistic as the scope for a single competency framework. The project struggled with the issue of where to draw the line. The community aspect of the project is meant to provide a continuous feedback channel on this issue and foster an evolving sense of the boundaries of the *Index*.

2.4 Outcome

The outcome of this process is a community-developed *Index* that describes a set of learning objectives and outcomes in terms of relevant knowledge, skills, practices, and habits of mind necessary to learn successful Linked Data practice, from design and modeling through implementation and maintenance.

The structure of the CI is as follows:

- *Topic cluster*
  - **Topic**
    - Competency: Tweet-length assertion of knowledge, skill, or habit of mind
      - *Benchmark: Action demonstrating accomplishment in related competencies*

This structure is implemented in the LD4PE triplestore, and expressed on the Learning Resources page of the LD4PE website. There are six Topic Clusters in the current version:

1. Fundamentals of Resource Description Framework
2. Fundamentals of Linked Data

https://doi.org/10.23106/dcmi.952137899

3. RDF vocabularies and application profiles
4. Creating and transforming RDF Data
5. Interacting with RDF Data
6. Creating Linked Data applications

Under these clusters, there are 30 topic groups, which contain 95 competencies. In the following example, under the topic "RDF serialization", there are two competencies, each of which is associated with a benchmark.

- **RDF serialization**
  - Distinguishes the RDF abstract data model and concrete serializations of RDF data.
    - *Expresses data in serializations such as RDF/XML, N-Triples, Turtle, N3, Trig, JSON-LD, and RDFa.*
  - Understands RDF serializations as interchangeable encodings of a given set of triples (RDF graph).
    - *Uses tools to convert RDF data between different serializations.*

One can think of competencies as expressing the knowledge to be imparted by instructors, covered in tutorials, or acquired by self-learners, and benchmarks as the basis for homework assignments, exercises, or exams.

2.5 Further Development of the Competency Index

Like the evolving technologies of Linked Data itself, the Competency Index for Linked Data will always be a work in progress. Over the course of the project, technologies for supporting a crowdsourced approach to the maintenance of the Index have also evolved. As of September 2017, the "canonical" version of the index has been installed in a Github repository under management of DCMI (in both a Chinese and English version). Github is a "social coding" platform originally created to support the collaboration of programmers in the development of open-source software such as Linux, but is increasingly being used for collaboration on other open-source endeavors, such as the maintenance of metadata vocabularies such as Schema.org.

The LD4PE Competency Index Editorial Board will continue its work in this new context, using Github's facilities for allowing collaborating editors, or even members of the public, to propose changes or additions to the Index. By navigating to https://dcmi.github.io/ldci/D2695955/, anyone can click on the button "Edit in Github", after which they will be guided through the process of submitting a "pull request" -- a request that the changes be integrated into the master copy of the Index. Github will display the differences between the master copy and the copy with proposed changes, side-by-side, for consideration by the Editorial Board; support discussion threads about aspects of the proposal; and accepted changes will be merged into the master copy.

LD4PE and DCMI have not yet established any policy about the URIs used to identify competencies or to their persistence, so the ASN-based URIs hitherto used may be replaced by new URIs, such as PURLs. DCMI also wants to experiment with new approaches to keeping translations of the Index synchronized as changes are made. The Chinese translation has been set up its own Github repository and for now, changes to the English version will be communicated to the Chinese translators in the form of "diffs" -- side-by-side comparisons showing changes approved and published by the Editorial Board.

Although no formal assessments have been made of the Competency Index to date, the development process itself has provided substantial input from subject matter experts as well as users. We expect that further use and adoption of the Index will contribute to its maturation and development over time.

https://doi.org/10.23106/dcmi.952137899

3. Toolkit

The LD4PE project is committed to the use of open standards and, where appropriate, adaptation of existing tools and toolsets, implementing anew only where necessary. The LD4PE toolkit is a set of open tools that enable the expression and use of the *Index* and associated Learning Resources through the LD4PE website. Key areas of project implementation are light-weight, browser-based editors for creation of Learning Resource descriptions, competency descriptions, and controlled vocabularies. On the LD4PE website, custom WordPress plug-ins manage integration of Learning Resource descriptions as instances of custom post-type and represent competency descriptions and vocabularies as custom taxonomies. Custom plug-ins also extend WordPress functionality to allow authenticated users to save collections of Learning Resource descriptions and organize competency statements into "learning maps" sequenced to support curriculum development.

3.1 Metadata Generation Tools

To support the generation of metadata describing Learning Resources and for authoring the *Index* and its competency assertions, the project has developed an open toolkit consisting of two browser-based editors implemented using AngularJS. The first of these editors is designed for creating RDF Learning Resource descriptions based on an application profile of the Learning Resource DCMI (schema.org) schema.

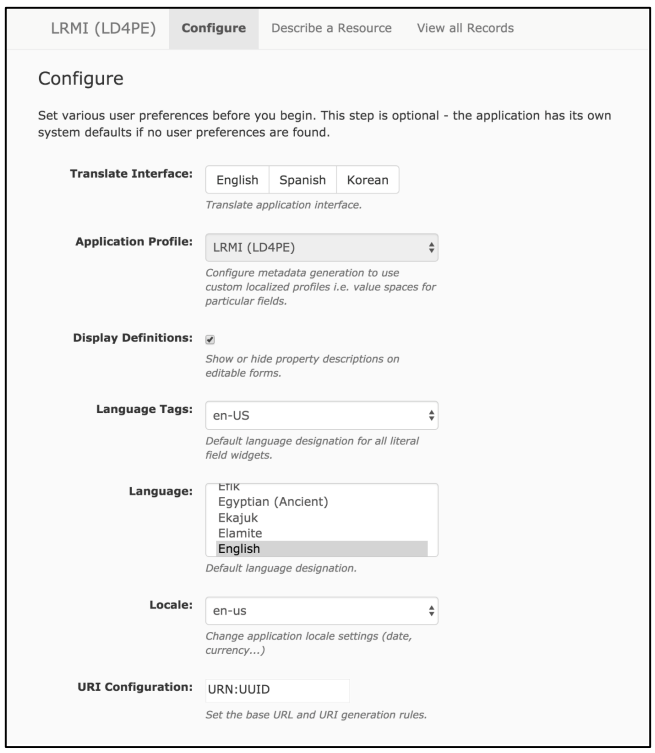

FIG. 3. LD4PE Competency Framework editor.

The second editor can be used to create *Index* metadata based on the ASN Description Language (ASN-DL) for competency framework modeling and description, following the editorial process described earlier, or some other version of that process suitable to another knowledge domain.

https://doi.org/10.23106/dcmi.952137899

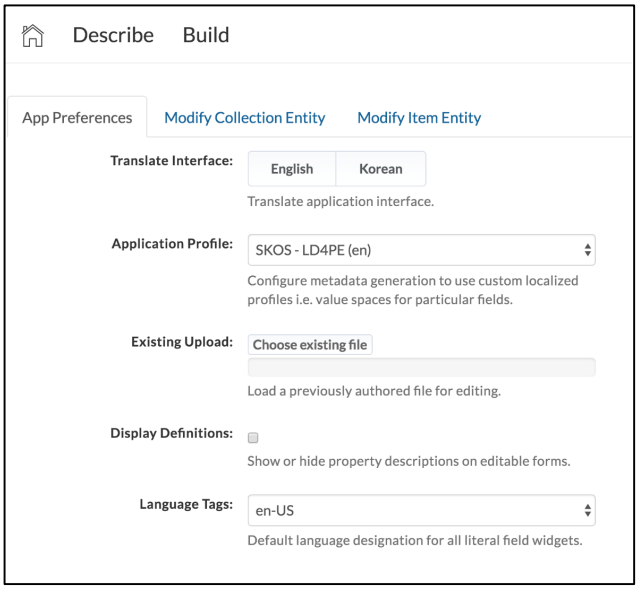

FIG. 4. LD4PE Learning Resource editor.

These two editors are key components of the reusable design followed by the project. While the LD4PE project has focused exclusively on the domain of Linked Data, we anticipate that others may wish to create similar competency frameworks for other areas, and catalog Learning Resources related to those domains using these tools. The tools and custom WordPress plug-ins will be available for download from Github to allow their reuse in other contexts

3.2 Saved Sets and Learning Maps Tools

While the *Index* defines a set of competencies, it neither prescribes any competencies as "core" nor defines a logical sequencing of its components. In other words, the *Index* does not, in itself, define a curriculum by prescribing a specific learning trajectory or map through the set of competencies. Instead, the LD4PE project provides tools to enable teachers, trainers, and learners to map their own pathways through the *Index* graph, through creation of Saved Sets of Learning Resources and *Learning Maps* of competencies. Saved Sets and Learning Maps traverse, or overlay, the competency nodes of the *Index*.

This is accomplished through custom plug-ins that extend basic WordPress functionality to allow authenticated users to define *Saved Sets* of Learning Resource descriptions for future access, as well as adding and removing Learning Resources from their sets. A similar tool allows users to sequence sets of competency statements to define Learning Maps representing logical pathways for instruction. Both tools give users the option to open sets or maps for public access, allowing other users to take those curated collections as starting points to extend and adapt for their own purposes.

The saved learning trajectories or pathways may be identified as formal curriculum structures or as personalized trajectories created by instructors or learners as evidence of progress. They may also function as suggested paths forward in the learning process or as guides to other instructors or learners in creating their own trajectories.

https://doi.org/10.23106/dcmi.952137899

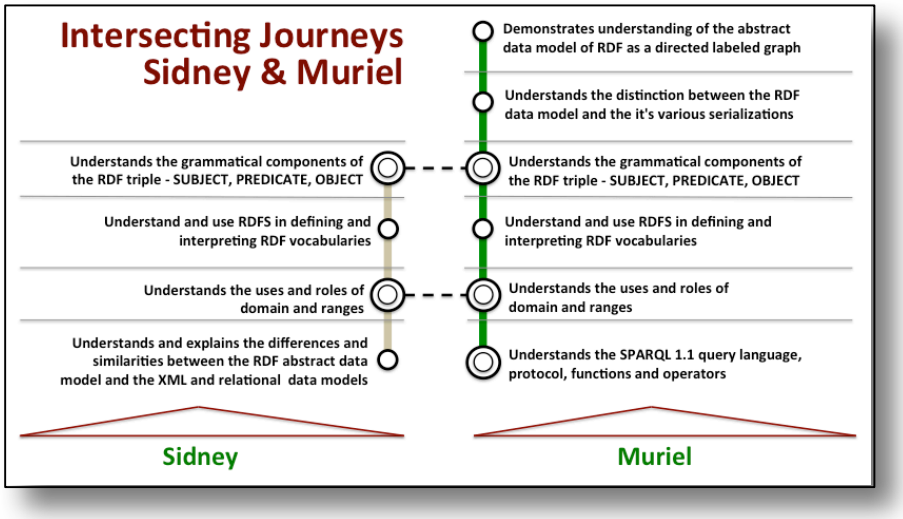

FIG. 5. Learning trajectory maps.

Research has shown that such maps reveal the macrostructure of the body of information or knowledge within a field, making the context of that information or knowledge more apparent and useful to learners (O'Donnell et al., 2002; Hall & O'Donnell, 1996; Hall & O'Donnell, 2010). Similarly, learning trajectory overlays contextualize competencies and learning outcomes in a "larger picture" of learning expectations, outcomes, and personal learner progressions.

Thus, LD4PE stands in sharp contrast to other projects in not prescribing a single curricular point of view but in providing instead the means for instructors, trainers and learners to chart multiple, diverse pathways for learning—pathways defined as public or private, individual or collective, prescribed or exploratory. Each Saved Set provides a different roadmap for discovering and traversing lesson plans, how-to recipes, webinars, and tutorials that have been described and aligned to the competency nodes of the *Index* graph.

3.3 Sample data set for teaching

Although not formally part of the tool kit, OCLC has contributed a large static triplestore (a subset of their WorldCat Linked Data focused on the Library Science domain) that is available for use in the creation of stable examples and assessments against the competencies that can be used in teaching situations. This avoids the difficulties inherent in using live examples which can change rapidly from minute to minute, and allows instructors to create reusable activities for learners with known outcomes. A tutorial has also been created to help novice users take advantage of this data set and construct queries against the triple store (http://explore.dublincore.net/related/oclc-dataset/).

4. Architecture

The LD4PE technical architecture consists of a WordPress public interface on top of a triplestore managing Learning Resource descriptions, the *Index* itself and controlled vocabularies for selected statements. As of September 2017, the *Index,* hitherto available through the ASN-D2L open repository and reflected on the LD4PE WordPress website as a custom taxonomy, is henceforth available at https://dcmi.github.io/ldci/, backed by the DCMI Github repository https://github.com/dcmi/ldci/. The editing tools function as open, stand-alone applications that can be used independently of the WordPress instance and the project's triplestore.

https://doi.org/10.23106/dcmi.952137899

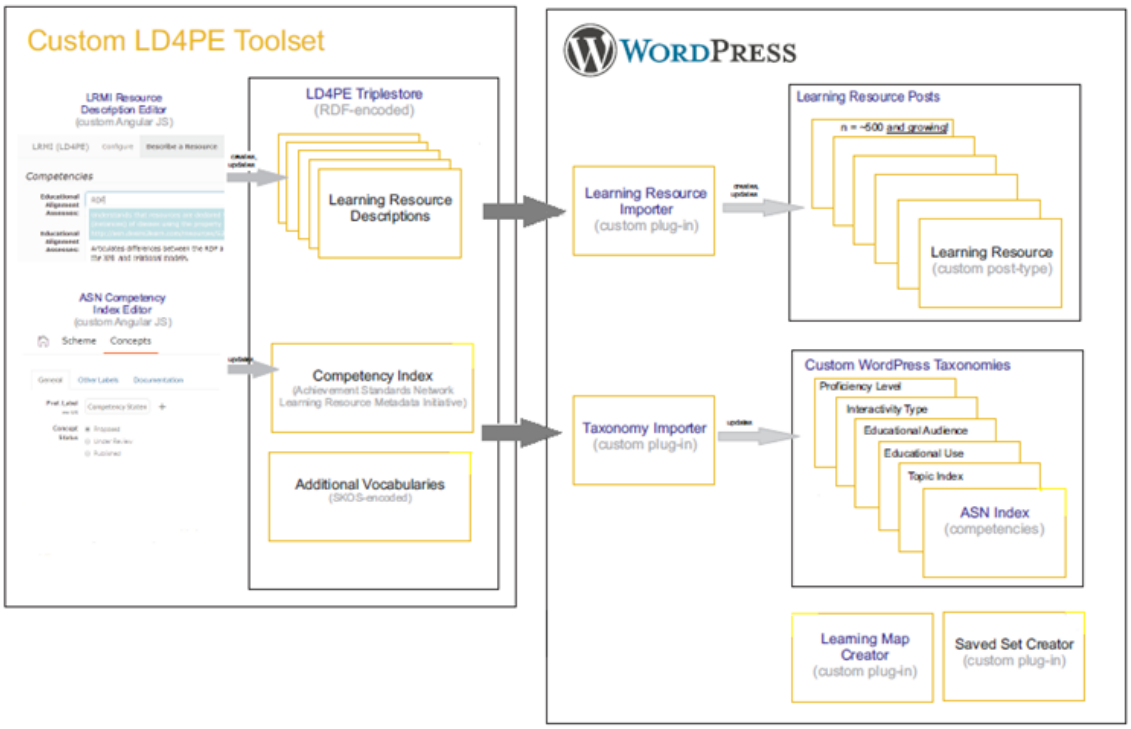

FIG. 6. LD4PE architecture.

The LD4PE website primarily provides competency-based browse access to Learning Resources. In addition, it provides a publication venue both for learning trajectory maps and for select Learning Resources such as a pilot set created by project partners Sungkyunkwan University Institute of Information and Management (Korea), OCLC, Elsevier, Synaptica, and Access Innovations. A key asset donated by OCLC is a very large static triplestore publicly available as a resource for developing replicable exercises, tests, and examples for teaching Linked Data principles.

5. Best Practice Documentation and Dissemination Activities

Publicly available guidelines are published on the LD4PE website covering best practices in:

1. Community development and management of competency frameworks, including a description of the process used by the editorial board to define and organize the competencies and benchmarks, and achieve a consistent style across the range of entries. (English version [https://dcmi.github.io/ldci/]; Chinese version [https://dcmi.github.io/ldci-zh/]);
2. Creation of useful metadata descriptions of Learning Resources using the Learning Resource DCMI application profile and alignment of Learning Resources to relevant competencies, including instructions for using the publicly available schema tools (http://explore.dublincore.net/related/share-our-tools/);
3. Development of learning pathways to create guided navigation through the competencies by teachers and learners, for use in curricula or training sessions aimed at various audiences and knowledge levels (http://explore.dublincore.net/theory/learning-pathways-as-transit-maps/); and
4. A sample data set that can be used for teaching purposes, along with a tutorial that can be used to help novice users take advantage of this data set and construct queries against the triple store (http://explore.dublincore.net/related/oclc-dataset/).

https://doi.org/10.23106/dcmi.952137899

Through international professional conferences, the LD4PE team members have reached out to various communities that may benefit from the Competency Index and the Learning Resources compiled by the LD4PE. These include the schools of information science (iSchools), the galleries, libraries, archives and museums communities, and scholarly publishing community. Conference presentations and training sessions have been given in the United States, Europe, and Asia. In addition, extensions of the Learning Resources and *Index* to other languages has been encouraged and supported, as demonstrated by the first fully translated Competency Index to Chinese in 2016, rolled out in a well-attended seminar in Asia. The tool set of LD4PE supports the configuration of non-English languages for the resource description editor as well, providing the opportunity for inclusion of Learning Resources in non-English languages in the future.

6. Conclusion

While LD4PE focuses specifically on skills, knowledge, and professional practice in the area of Linked Data, nothing precludes the extension of its competency-based approach to other areas of library and information science (LIS), archives, and museum curricula. Similar competency frameworks could be developed describing practice in knowledge organization systems, cataloging, and organizational management. Describing a knowledge domain in terms of competencies necessary for mastery provides a solid scaffolding for teaching and learning, and our hope is that others will take advantage of the tools and practices embodied in this project to develop similar offerings in other areas.

This approach will be appropriate wherever the goal is for learners to achieve competence within a defined set of knowledge, skills, and habits of mind. DCMI is exploring ways to sustain and grow the products and assets of this work moving forward because it sees the project's conceptualization and its outcomes (including best practice documentation related to the framework itself) as generalizable to all areas of principled metadata design and best practice, while also providing a guiding template for future development of DCMI's education and training agenda.

Acknowledgments

The work described in this poster was partially funded by the Institute of Museum and Library Services (IMLS). Content partners Sungkyunkwan University Institute of Information and Management (Korea), OCLC, Elsevier, Synaptica, and Access Innovations were generous in donating their time and expertise to the project.

References

Berners-Lee, Tim (2006). Design Issues: Linked Data. Retrieved June 28, 2015 from http://www.w3.org/DesignIssues/LinkedData.html.

Crandall, M.D., Tennis, J., Sutton, S.A., Baker, T. & Talley, D. (2013). Planning a platform for learning linked data. In Moen, W. & Rushing, A. (Eds.), *International Conference on Dublin Core and Metadata Applications*. Retrieved June 28, 2015 from http://dcpapers.dublincore.org/pubs/article/view/3693/1916.

Hall, R.H. & O'Donnell, A. (1996). Cognitive and affective outcomes of learning from knowledge maps. *Contemp. Educ. Psychol. 21*, 94-101.

Hall, R.H., Hall, M.A. & Saling, C.B. (2010). The Effects of Graphical Postorganization Strategies on Learning from Knowledge Maps. J. Exp. Educ. 67(2):101-112 (DOI:10.1080/00220979909598347).

O'Donnell, A.M., Dansereau, D.F. & Hall, R.H. (2002). Knowledge maps as scaffolds for cognitive processing. *Educ. Psychol. Review 14*, 71-85.

Sutton, S. A. & Golder, D. (2008). Achievement Standards Network (ASN): An application profile for mapping K-12 educational resources to achievement standards. In Heike Neuroth (Ed.), *Proceedings of the International Conference on Dublin Core and Metadata Applications*, (pp. 69-79). Retrieved, June 28, 2015 from http://dcpapers.dublincore.org/ojs/pubs/article/view/920/916.

https://doi.org/10.23106/dcmi.952137899

Ward, N. & Nicholas, N. (2010). Benefits of machine readable curricula. Retrieved, June 28, 2015 from http://www.australiancurriculum.edu.au/static/docs/Benefits%20of%20a%20Machine%20Readable%20Curricula.pdf

https://doi.org/10.23106/dcmi.952137899