

# Metadata Architecture and Applications

UNC iSchool Summer Session 2, Instructor: Sam Oh (SKKU iSchool, Korea)

## Course Description:

The course covers fundamentals of metadata, metadata building blocks including design and implementation details, metadata services, the role of metadata in the Web of Data, and metadata outlook in research. The main objective of this course is to equip students with essential knowledge needed to create sound and interoperable application profiles (APs), which constitute an important infrastructure for enhanced information service to users and to understand the role of linked data for metadata interoperability. Implementations of metadata schemas and APs will be done using XML technology. The course focuses on how to achieve syntactic and semantic interoperability among diverse metadata schemas, APs and linked data.

## Course Objectives:

The purpose of this course is to provide students with conceptual and technical knowledge regarding how to design interoperable metadata schemas and application profiles. The objectives of this course are as follows:

- Students will acquire in-depth understanding of XML technology.
- Students will acquire in-depth understanding of most widely used metadata schemas and application profiles (APs) available today.
- Students will be able to use XML technology to design complex APs using well-established metadata standard terms.
- Students will learn how to use XML namespace and URIs to design interoperable metadata schemas and APs.
- Students will learn how to implement APs using Oxygen XML Editor.
- Students will not only understand published schemas (MODS, PREMIS, VRA, textMD, MIX, etc), but also are able to extend standard schemas to design new APs.
- Students will understand RDF (Resource Description Framework) and LD (Linked Data) to enhance the connectivity of metadata.
- Students will also learn the essence of ontology modeling using RDF/OWL.

## Course Expectations and Requirements:

I anticipate that students will grasp essential concepts and technical details necessary for effective metadata design and management in the Semantic Web environment. In this course students will be required to use Oxygen XML editor heavily and the essence of TBC ontology editor.

## Course Format:

The entire class will be conducted in F2F mode.

## Textbook:

Marcia Zeng and Jin Qin. (2016). Metadata. 2<sup>nd</sup> Edition. ISBN-13: 978-1555709655

## Grading:

- AP Project 1 (DC Application Profile): 10%
- AP Project 2 (VRA or LOM Application Profile): 20%
- AP Project 2 (MODS Application Profile): 20%
- AP Project 4 (RDF/OWL Modeling): 20%
- Final Project: (METS AP): 30%

## Instructor Bio:

Sam Oh is a professor at Sungkyunkwan University (established in 1398), Seoul Korea and an affiliate professor at UW iSchool. Prior to joining SKKU, he taught at the UW iSchool for 4 years as an assistant professor. His teaching and research interests include metadata and ontology design, data modeling, social data analytics, big data analytics, and knowledge management (KM). He has extensive consulting experiences in the areas such as metadata/ontology design, social data analytics for companies and government sectors in Korea. He was recently elected as iCaucus chair (2018-2019). He chaired TC46/SC9 (Identification & Description) for 6 years and currently serves as a chair of ISO/IEC JTC1 SC34 (Document Description & Processing Languages). He also serves as a member of DCMI (Dublin Core Metadata Initiative) governing board. He has also teaching experiences at other iSchools such as Syracuse University, University of Pittsburgh, University of Texas Austin, and University of North Carolina Chapel Hill.

Week	Topic	Reading	Project
Day 1 6/20 (M)	Course Overview FRBR, Metadata Basic, XML Schema I		
Day 2 6/21 (T)	XML Schema II Simple Metadata Schema using Oxygen DC Simple using Oxygen	Zeng and Qin (2016) Chapter 1	
Day 3 6/22 (W)	DC Advanced using Oxygen	Zeng and Qin (2016) Chapter 2	
Day 4 6/23 (R)	DC AP using Oxygen LOM, LOM AP using Oxygen	Zeng and Qin (2016) Chapter 3	
Day 5 6/24 (F)	VRA VRA AP using Oxygen	Zeng and Qin (2016) Chapter 4	
Day 6 6/27 (M)	MODS MODS AP using Oxygen	Zeng and Qin (2016) Chapter 5	P1: DC AP
Day 7 6/28 (T)	MIX   TextMD   PREMIS Lab for AP Design	Zeng and Qin (2016) Chapter 6	P2: VRA or LOM AP
Day 8 6/29 (W)	METS, METS AP using Oxygen RDF/OWL	Zeng and Qin (2016) Chapter 7	
Day 9 6/30 (R)	Linked Data Basic (Karen Coyle) RDF/OWL Ontology Modeling I	Zeng and Qin (2016) Chapter 8	P3: MODS AP
Day 10 7/1 (F)	RDF/OWL Ontology Modeling II Linked Data Advanced	Zeng and Qin (2016) Chapter 9	
7/4 (M)	No class: National Holiday		P4: RDF/OWL Modeling
Day 11 7/5 (T)	Presentation of Final Projects by students (20 minutes): METS AP and RDF/OWL Modeling	Zeng and Qin (2016) Chapter 10	P5: Final Project: METS AP

# Additional Reading List

## Day 1

[Mandatory]

- (1) NISO. 2004. Understanding Metadata. Bethesda, MD:NISO Press. [online]  
<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

[Optional]

- (1) Duval, Erik, Wayne Hodgins, Stuart Sutton, and Stuart L. Weibel. 2002. "Metadata Principles and Practicalities." D-Lib Magazine 8 (4). doi: 10.1045/april2002-weibel. [online]  
<http://www.dlib.org/dlib/april02/weibel/04weibel.html><http://www.dlib.org/dlib/april02/weibel/04weibel.html>
- (2) Gilliland, Anne J. 2008. "Setting the Stage." In Introduction to Metadata: Pathways to Digital Information, edited by Murtha Baca. Online Edition (Version 3.0). Los Angeles: Getty Research Institute. [online]  
[http://www.getty.edu/research/publications/electronic\\_publications/intrometadata/setting.html](http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html)
- (3) Weibel, Stuart, Jean Godby, Eric Miller, and Ron Daniel. 1995. OCLC/NCSA Metadata Workshop Report. [online]  
<http://dublincore.org/workshops/dcl/report.shtml>

## Day 2

[Mandatory]

- (1) Baker, Thomas. 2012. "Libraries, Languages of Description, and Linked Data: A Dublin Core Perspective." Library Hi Tech 30 (1):116–33.

[Optional]

- (1) Harper, Corey, 2010. "Dublin Core Metadata Initiative: Beyond the Element Set." Information Standards Quarterly 22 (1):21–28. [online] [http://www.niso.org/publications/isq/free/FE\\_DCMI\\_Harper\\_isqv22no1.pdf](http://www.niso.org/publications/isq/free/FE_DCMI_Harper_isqv22no1.pdf)
- (2) Heery, Rachel, and Manjula Patel. 2000. "Application Profiles: Mixing and Matching Metadata Schemas." Ariadne no. 25. [online] <http://www.ariadne.ac.uk/issue25/app-profiles/>
- (3) Manola, Frank, Eric Miller, and Brian McBride. 2014. "RDF 1.1 Primer." W3C Working Group Note 25, February 2014. [online] <http://www.w3.org/TR/rdf11-primer/>

## Day 3

[Mandatory]

[Optional]

- (1) Baca, Murtha, Patricia Harpring, Elisa Lanzi, Linda McRae, and Ann Whiteside, eds. 2006. "VIII. Database Design and Relationships." Cataloging Cultural Objects, A Guide to Describing Cultural Works and Their Images, 20–27. Chicago: American Library Association. [online] [http://cco.vrafoundation.org/index.php/toolkit/cco\\_pdf\\_version/](http://cco.vrafoundation.org/index.php/toolkit/cco_pdf_version/).
- (2) Greenberg, Jane. 2002. "Metadata Generation: Processes, People and Tools." Bulletin of the American Society for Information Science and Technology 29 (2):16-19. [online] <http://www.asis.org/Bulletin/Dec-02/greenberg.html>.
- (3) NISO Framework Advisory Group. 2007. A Framework of Guidance for Building Good Digital Collections. 3rd ed. [online] <http://www.niso.org/publications/rp/framework3.pdf>.
- (4) Shreeves, Sarah, Jenn Riley, and Liz Milewicz. 2006. "Moving Towards Shareable Metadata." First Monday 11 (8). [online] <http://firstmonday.org/ojs/index.php/fm/article/view/1386/1304>.
- (5) Manola, Frank, Eric Miller, and Brian McBride. 2014. "RDF 1.1 Primer. W3C Working Group Note 25 February 2014." <http://www.w3.org/TR/rdf11-primer/>.
- (6) Powell, Andy, Mikael Nilsson, Ambjörn Naeve, Pete Johnston, and Thomas Baker. 2007. "DCMI Abstract Model." [online] <http://dublincore.org/documents/abstract-model/index.shtml>.

## Day4

[Mandatory]

[Optional]

- (1) Caplan, Priscilla. 2003. Metadata Fundamentals for All Librarians, 25–53. Chicago: American Library Association.
- (2) Coyle, Karen, and Thomas Baker. 2009. "Guidelines for Dublin Core Application Profiles." DCMI Recommended Resource. [online] <http://dublincore.org/documents/profile-guidelines/>.
- (3) Greenberg, Jane. 2001. "A Quantitative Categorical Analysis of Metadata Elements in Image-Applicable Metadata Schemes." Journal of the American Society for Information Science and Technology 52 (11):917.
- (4) Hillmann, Diane I, and Elaine L. Westbrook Eds. 2004. Part I. Project-Based Implementations. Metadata in Practice. Chicago: American Library Association.

- (5) Sutton, Stuart A. 1999. "Conceptual Design and Deployment of a Metadata Framework for Education Resources on the Internet." *Journal of the American Society for Information Science* 50 (13):1182–92.
- (6) Weibel, Stuart. 2005. "Border Crossings, Reflections on a Decade of Metadata Consensus Building." *D-Lib Magazine* 11 (7/8). [online] doi:10.1045/july2005-weibel.<http://www.dlib.org/dlib/july05/weibel/07weibel.html>.

### Day5

[Mandatory]

[Optional]

- (1) Costello, Roger L. 2006. "XML Schemas: Best Practices." [online] <http://www.xfront.com/BestPracticesHomepage.html>.
- (2) D7.I.3 — Study on Persistent URIs, with Identification of Best Practices and Recommendations on the Topic for the MSs and the EC. 2012. Phil Archer, Stijn Goedertier, and Nikolaos Loutas. Deliverable released by European Commission ISA Programme. [online] <http://joinup.ec.europa.eu/sites/default/files/D7.I.3%20-%20Study%20on%20persistent%20URIs.pdf>.
- (3) Manola, Frank, Eric Miller, and Brian McBride. 2014. "RDF 1.1 Primer." W3C Working Group Note 25 February 2014. [online] <http://www.w3.org/TR/rdf11-primer/>.
- (4) Refsnes Data. 2007. "XML Schema Tutorial." W3Schools. [online] <http://www.w3schools.com/schema/default.asp>.
- (5) Srivastava, Rahul. 2007. "XML Schema: Understanding Namespaces." Oracle Technology Network. [online] <http://www.oracle.com/technetwork/articles/srivastava-namespaces-092580.html>.
- (6) van der Vlist, Eric. 2001. "Using W3C XML Schema." O'Reilly Media. [online] <http://www.xml.com/pub/a/2000/11/29/schemas/part1.html>.
- (7) W3C. 1999. XML Schema Requirements. XML Schema Working Group. [online] <http://www.w3.org/TR/NOTE-xml-schema-req>.
- (8) W3C. 2004. XML Schema Part 0: Primer Second Edition. W3C Recommendation 28 October 2004. <http://www.w3.org/TR/xmlschema-0/>.

### Day6

[Mandatory]

- (1) DCMI Usage Board. 2012. DCMI Metadata Terms. [online] <http://dublincore.org/documents/dcmi-terms/>
- (2) Library of Congress. 2015. Outline of Elements and Attributes in MODS. [online] <http://www.loc.gov/standards/mods/mods-outline.html>

### Day7

[Mandatory]

- (1) Library of Congress. 2014. VRA Core 4.0 Introduction. [online] [http://www.loc.gov/standards/vracore/VRA\\_Core4\\_Intro.pdf](http://www.loc.gov/standards/vracore/VRA_Core4_Intro.pdf)
- (2) Library of Congress. 2007. VRA Core 4.0 Element Outline. [online] [http://www.loc.gov/standards/vracore/VRA\\_Core4\\_Outline.pdf](http://www.loc.gov/standards/vracore/VRA_Core4_Outline.pdf)

[Optional]

- (1) Library of Congress. 2007. VRA Core 4.0 Element Description and Tagging Examples. [online] [http://www.loc.gov/standards/vracore/VRA\\_Core4\\_Element\\_Description.pdf](http://www.loc.gov/standards/vracore/VRA_Core4_Element_Description.pdf)
- (2) Library of Congress. 2007. Restricted Schema Type Values. [online] [http://www.loc.gov/standards/vracore/VRA\\_Core4\\_Restricted\\_schema\\_type\\_values.pdf](http://www.loc.gov/standards/vracore/VRA_Core4_Restricted_schema_type_values.pdf)
- (3) Library of Congress. 2007. VRA Core 4.0 Schemas - Unrestricted Version. [online] <http://www.loc.gov/standards/vracore/vra.xsd>
- (4) Library of Congress. 2007/ VRA Core 4.0 Schemas - Restricted Version. [online] <http://www.loc.gov/standards/vracore/vra-strict.xsd>

### Day8

[Mandatory]

- (1) Library of Congress. 2015. PREMIS Data dictionary for Preservation metadata. [online] <http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>

[Optional]

- (1) Baker, Thomas, Pierre-Yves Vandenbussche, and Bernard Vatan. 2013. "Requirements for Vocabulary Preservation and Governance." *Library Hi Tech* 31 (4): 657-668.
- (2) Berners-Lee, Tim. 2006. "Linked Data." Last modified June 6, 2009. [online] <http://www.w3.org/DesignIssues/LinkedData.html>.

- (3) Deliot, Corine. 2014. "Publishing the British National Bibliography as Linked Open Data." Library Association of Ireland (LAI) 2014 Annual Seminar "Linked Data for Libraries," 6th November 2014, Dublin, Ireland. [online] [http://www.bl.uk/bibliographic/pdfs/publishing\\_bnb\\_as\\_lod.pdf](http://www.bl.uk/bibliographic/pdfs/publishing_bnb_as_lod.pdf).
- (4) Fons, Ted, Jeff Penka, and Richard Wallis. 2012. "OCLC's Linked Data Initiative: Using Schema.org to Make Library Data Relevant on the Web." *Information Standards Quarterly* 24 (2–3):29–33. [online] [http://www.niso.org/apps/group\\_public/download.php/9408/IP\\_Fons-et-al\\_OCLC\\_isqv24no2-3.pdf](http://www.niso.org/apps/group_public/download.php/9408/IP_Fons-et-al_OCLC_isqv24no2-3.pdf).

## Day9

[Mandatory]

[Optional]

- (1) Bruce, Thomas R., and Diane I. Hillmann. 2004. "The Continuum of Metadata Quality: Defining, Expressing, Exploiting." In *Metadata in Practice*. Eds. Diane I. Hillmann, and Elaine L. Westbrook, 238–56. Chicago: American Library Association.
- (2) Guy, Marieke, Andy Powell, and Michael Day. 2004. "Improving the Quality of Metadata in EPrint Archives." *ARIADNE* no. 38. [online] <http://www.ariadne.ac.uk/issue38/guy>.
- (3) Hillmann, Diane I., Naomi Dushay, and Jon Phipps. 2004. "Improving Metadata Quality: Augmentation and Recombination." *DC-2004 International Conference on Dublin Core and Metadata Applications*, 11–14 October 2004, Shanghai, China. [online] <http://dcpapers.dublincore.org/pubs/article/view/770/766>.
- (4) O'Neill, Edward T., and Diane Vizine-Goetz. 1988. "Quality Control in Online Databases." In *Annual Review of Information Science and Technology*, edited by M. Williams. Vol. 23: 125–56. Medford, NJ: Learned Information.
- (5) Van Hooland, Seth and Ruben Verborgh. 2014. *Linked Data for Libraries, Archives and Museums—How to Clean, Link and Publish Your Metadata*. London: Facet Publishing. Chapters 3–6.
- (6) Zaveri, Amrapali, Anisa Rula, Andrea Maurino, Ricardo Pietrobon, Jens Lehmann, Sören Auer, and Pascal Hitzler. 2013. "Quality Assessment Methodologies for Linked Open Data." Submitted to *Semantic Web Journal*. [online] <http://www.semantic-web-journal.net/system/files/swj414.pdf>.

## Day10

[Mandatory]

[Optional]

- (1) Bernstein, Philip. A., Jayant Madhavan, and Erhard Rahm. 2011. "Generic Schema Matching, Ten Years Later." In *Proceedings of the 37th International Conference on Very Large Data Bases (VLDB 2011)*, edited by H. V. Jagadish et al., 695–701. Seattle: VLDB Endowment.
- (2) Halpin, Harry, Ivan Herman, and Patrick J. Hayes. 2010. "When owl:sameAs Isn't the Same: An Analysis of Identity Links on the Semantic Web." *W3C Workshop—RDF Next Steps*, Stanford University, Palo Alto, CA, USA. June 26–27, 2010. [online] <http://www.w3.org/2009/12/rdf-ws/papers/ws21>.
- (3) Heery, Rachel. 2004. "Metadata Futures: Steps toward Semantic Interoperability." *Metadata in Practice*, edited by Diane I. Hillmann and Elaine L. Westbrook, 257–271. Chicago: American Library Association.
- (4) St. Pierre, Margaret, and William P. LaPlant, Jr. 1998. *Issues in Crosswalking Content Metadata Standards*. Bethesda, MD: NISO Press. [online] [http://www.niso.org/publications/white\\_papers/crosswalk/](http://www.niso.org/publications/white_papers/crosswalk/).

## Day11

[Mandatory]

[Optional]

- (1) Berners-Lee, Tim. 1997. *Metadata Architecture*. [online] <http://www.w3.org/DesignIssues/Metadata.html>.
- (2) Gold, Anna. 2007a. "Cyberinfrastructure, Data, and Libraries, Part I: A Cyberinfrastructure Primer for Librarians." *D-Lib Magazine* 13 (9/10). [online] <http://www.dlib.org/dlib/september07/gold/09gold-pt1.html>.
- (3) Gold, Anna. 2007b. "Cyberinfrastructure, Data, and Libraries, Part II: Libraries and the Data Challenge: Roles and Actions for Libraries." *D-Lib Magazine* 13 (9/10). [online] <http://www.dlib.org/dlib/september07/gold/09gold-pt2.html>.
- (4) Schöch, Christof. 2013. "Big? Smart? Clean? Messy? Data in the Humanities." *Journal of Digital Humanities* 2 (3). [online] <http://journalofdigitalhumanities.org/2-3/big-smart-clean-messy-data-in-the-humanities/>.