School of Information Studies Faculty Articles

Information Studies (School of)

2020

Linking, Mapping, Matching, and Change: Contemporary Use of Ranganathan's Three Planes of Work in Classification Activity

Wan-Chen Lee
University of Wisconsin - Milwaukee, wclee@uwm.edu

Recommended Citation

Wan-Chen Lee – University of Washington, United States

Linking, Mapping, Matching, and Change:
Contemporary Use of Ranganathan's Three Planes of Work in Classification Activity

Abstract:
Scholars have identified interoperability issues in mapping metadata in a linked data environment (Zeng 2019). This study builds on previous research and proposes a creative use of Ranganathan's (1989) three planes of work in classification activity. By extending the application of the three planes of work to the linked data environment, we can use this conceptual model as an analytical tool to highlight particular mapping challenges. This paper uses three cases to show how discrepancies between the idea plane, verbal plane, and notational plane may cause mapping issues. Further, we can see that mapping issues are not limited to differences between metadata standards. The three planes of work can highlight mapping issues that are caused by changes at different planes of the same metadata. The challenges presented in this study complement the known mapping issues, and contribute to the discussion of interoperability in linking, mapping, matching, and change in metadata.

1.0 Introduction

The library community has been linking and mapping metadata before the linked data era. For instance, the testing project for Virtual International Authority File started in 1998, linking authority records created by international institutions. Catalogers can look up correlations between the Library of Congress Classification and Dewey Decimal Classification using Classification Web since 2004. Linking and mapping metadata can improve interoperability, lower maintenance cost, and increase the use of authority data both within and beyond the library community (OCLC 2019). Today, linked data shapes and supports the linking and management of metadata. It provides new approaches for authority control. For years, authority control has been text-based. Entities with the same name are differentiated through textual labeling and qualifiers. With linked data, we can manage identities by assigning unique identifiers, which is less language-dependent. While identity management does not replace text-based authority control, major metadata creators and aggregators like the Library of Congress and OCLC have been adding unique identifiers to entities to enhance their authority data. However, regardless of whether metadata carry unique identifiers, linking metadata is not without concerns. One challenge is that links between metadata may have different meanings, and the meanings may not be clear to users. A link can represent *linking*, which shows the linked metadata as related. This general sense of linking can cover many kinds of relationships (Green 2001). A link can also represent *mapping*, which often indicates a functional equivalence relationship between the linked metadata. When two terms are mapped, they are treated as equivalent functionally. That is, one term can stand in for another term. However, this does not guarantee semantic equivalence. An example is posting up a narrower term. In a standard, we may see USE cross reference that instructs users to use a broader term (e.g., dogs) for a narrower term (e.g., corgis). In this case, the two

terms are treated as equivalent functionally, but not equivalent semantically. We can also see *mapping* between terms from standards with different levels of specificity. A link can also represent *matching*, which indicates semantic equivalence. Variant forms of the same Library of Congress Subject Heading (LCSH) is an example of matching. Since these linking types may all be represented by the same expression: a link, without explicit specifications, the meanings of links may be ambiguous. Adding to the complexity of links, the distinctions I have outlined here, between linking, mapping, and matching are not always acknowledged and used consistently. People may use different categories for linking types, such as exact match, partial match, etc. Also, these linking types could refer to different semantic relationships (e.g., hierarchical, equivalence, associative), and there is no one-to-one relationship between a linking type and a semantic relationship. Recognizing the ambiguity of links, some projects link metadata with pre-defined relationships (e.g., DCMI metadata terms). This clarifies the meanings of links, but users may have to take extra steps to access the scope notes of the represented relationships.

Meaning changes over time is another concern for linked metadata. Assigning unique identifiers to linked metadata enables easier updates for the preferred form of an entity, such as name changes. Nonetheless, the updated form only represents the updated meaning of an entity. Users cannot trace the history of meaning changes or name changes of an entity. Without contextual information, users would not know which links of an entity were created before or after meaning changes, and whether the links were re-evaluated.

One other concern for linked metadata is unclear or inconsistent linking, mapping, and matching criteria. One example is a pilot project I observed in my ethnographic fieldwork[1]. The project is an attempt of a group of librarians to explore mapping LCDGT (Library of Congress Demographic Group Terms) to LCSH. The group leader drafted mapping criteria. Members go through all LCDGT terms, and use the criteria to search for matches or closest matches in LCSH. In the mapping process, members surface different aspects of mapping, including concept, text string, and types of heading. When there is an exact concept match in LCSH for a LCDGT term, more complexities follow. For instance, the matched LCSH may or may not use the identical text string. Also, a LCDGT term may match with a variant form of a LCSH or a former heading. How could we distinguish and present the *different types of exact concept match* to users? This project shows how sophisticated mapping criteria may be. When linking metadata, if we only show links between metadata without clear explanation of the criteria, we risk using a set of criteria that differ from users' expectations.

Besides the concerns discussed above, previous studies such as the AAT-Taiwan project identified mapping and translation issues in developing the Chinese language *Art & Architecture thesaurus* (Chen, Zeng, & Chen 2016). Likewise, Zeng (2019) reviews research, standards, and projects, and discusses approaches to address interoperability issues in metadata mapping. These studies present categorizations of mapping issues and provide suggestions to address these issues. Building on previous research, this paper proposes a creative use of Ranganathan's (1989) three planes of

---

[1] The ethnographic fieldwork started in September 2015, and ended in November 2019. I shadowed a cataloger at an academic library to explore cultural influences in cataloging practices. Through participatory observations, informal interviews, and taking field notes, I captured rich cataloging scenarios of the cataloger's applications of international and U.S. standards to catalog materials in various formats and languages, and her interactions with other librarians.

work. By extending its application from classification work to the linked data environment, we can use it as an analytical tool to discuss issues of linking, mapping, matching, and change in metadata.

2.0 Cases

This section will apply Ranganathan's (1989) three planes of work to three cases to analyze linking, mapping, and matching issues. The three planes are idea plane, verbal plane, and notational plane. Ideas are originated from the minds of their creators, and communicated through language. Language is the medium for the communication of ideas. The ambiguities of language (e.g., homonyms) are embedded in the representations of ideas. Recognizing the ambiguities of natural language in verbal plane, the notational plane represent disambiguated meanings or help arrangement. Notations may include language, symbols, and numbers (e.g., class numbers). The three planes provide a structure to break down the levels of abstraction of the concepts described in knowledge organization actions, such as classification and cataloging.

2.1 Eugenics in DDC, CCL, and NDC: discrepancies between the three planes

Tennis (2012) examines the subject ontogeny (i.e., the life of a subject over time) of *eugenics* in all editions of the Dewey Decimal Classification (DDC). The study demonstrates meaning changes of *eugenics*, and how those changes were reflected in the scheme over time. Using the three planes of work, we can view this as an example of changes in the idea plane (i.e., the definitions of *eugenics*) leading changes of the notational plane (i.e., the class numbers), while the verbal plane (i.e., the verbal expression "eugenics") remains the same. Based on this study, I examine the subject ontogeny of *eugenics* in the New Classification Scheme for Chinese Libraries (CCL) in Taiwan and the Nippon Decimal Classification (NDC) in Japan, and bibliographic records with *eugenics* as a subject heading (Lee 2016). Through analyzing the titles and co-assigned subject headings in the bibliographic records, I capture different meaning changes of *eugenics* in CCL and NDC (Lee 2018). However, these changes at the idea plane were not reflected in the other two planes. The subject name *eugenics* remains the same, and the class numbers for *eugenics* in both schemes are relatively static. If we compare the ontogeny of *eugenics* in DDC, CCL, and NDC, we can identify two factors that may lead to issues in mapping metadata. First, the meaning of a concept may change and diverge in different languages and schemes. This makes mapping metadata in different languages more challenging. Further, when the idea plane is not in sync with the other two planes, there is a risk of misrepresentation and imprecise mapping. The three planes of work can help us identify subtle changes in the idea plane.

2.2 Vernacular title or non-Latin script title: change at the verbal plane

Through the aforementioned ethnographic fieldwork, I observed a case of change at the verbal plane. The cataloger Q [pseudonym], whom I observed in the field, retrieved a bibliographic record of a Japanese book using the library catalog. In the record, the title in the original script was labeled *vernacular title*, and the Romanized form of the title was labeled *title*. Q explained that the term *vernacular* is discriminative[2]. The East

---

[2] According to Merriam-Webster Online dictionary (2020), *vernacular* was first known and used as "using a language or dialect native to a region or country rather than a literary, cultured, or foreign language." The origin *verna* means "slave born in the household."

Asian Libraries and the cataloging community in the U.S. have stopped using this term for more than a decade. Q sent a proposal to change the name of this metadata attribute, and it was changed to *non-Latin script title*. This is an example of how the verbal plane may shape people's understanding and reaction to the idea plane of a concept. While both *vernacular title* and *non-Latin script title* may refer to the same concept, the nuance of an expression may carry different meanings and lead to different interpretations over time. Failing to account for this could cause issues in linking and managing metadata.

2.3 LCNAF and Wikidata: notational plane and structural interoperability

In the field, Q shared their observations of mapping between the Library of Congress Name Authority File (LCNAF) and Wikidata. When editing LCNAF records, catalogers can add links to other resources that describe the same entity. For instance, in the LCNAF for *Twain, Mark, 1835-1910* (Library of Congress 2020), catalogers can add a link to the Wikidata entry for *Mark Twain* (Q7245 2020). In this case, the two sources describe the same person, with the same verbal expression, using different notations. The issue is, the LCNAF is identity based, while Wikidata is person based. For people who publish works using multiple identities, each identity has its own LCNAF record. The records of the same person are linked to one main record, which serves as a hub and links to different identities of the person. Wikidata collocates all identities of a person under one page. If we search for Samuel Clemens in Wikidata, we will be directed to the Mark Twain page. Hence, when linking a LCNAF record that represents *one identity of a person* to a Wikidata entry of *a person with multiple identities*, the link does not connect two records with the same scope. Mapping issues may occur even if the idea plane and verbal plane are identical. How could we present the structural distinctions so users do not assume an equivalence relationship between the records? How to clarify the meanings of links? If we access the LCNAF record for *Twain, Mark, 1835-1910* through the LC linked data service, we see the link to Mark Twain's Wikidata page is under *closely matching concepts from other schemes*, which indicates a non-equivalence relationship. The Wikidata page for Mark Twain lists the Library of Congress authority ID for both *Twain, Mark, 1835-1910* and *Clemens, Samuel Langhorne, 1835-1910*. Users may infer the relationships between these identities from other metadata on the page, which specify that Mark Twain is *also known as* Samuel Langhorne Clemens. While both systems indicate the differences between the records, the meanings of the link are not explicitly clear. Could we improve this, maybe at the notational plane?

3.0 Conclusion

Ranganathan's three planes of work in classification activity remains relevant in the linked data environment. By extending its application, we can use it as an analytical lens to examine issues of linking, mapping, matching, and change in metadata. Also, it accounts for meaning changes of the same concept over time, which may undermine semantic interoperability if not reflected in metadata linking. This complements studies that focus on linking issues between different metadata standards. Further, we can interpret and address linking issues by identifying discrepancies between the three planes. On one hand, following Ranganathan (1989), we expect the idea plane to lead the change of its expressions in the other two planes. Through the *eugenics* case, we see the risk of misrepresenting concepts when the three planes are not in sync. On the other hand, in the latter two cases, we recognize that change in the idea plane is not the only force of change for the verbal plane and notational plane. Change in these two planes,

under the premise of remaining in sync with the idea plane, may help address linking issues.

References

Chen, Shu-jiun, Marcia Lei Zeng, and Hsueh-hua Chen. 2016. "Alignment of Conceptual Structures in Controlled Vocabularies in the Domain of Chinese Art: A Discussion of Issues and Patterns." *International Journal on Digital Libraries* 17, no.1: 23-38. Doi:10.1007/s00799-015-0163-1

Green, Rebecca. 2001. "Relationships in the Organization of Knowledge: An Overview." In *Relationships in the Organization of Knowledge*, edited by C. A. Bean & R. Green. Springer Netherlands, 3-18.

Lee, Wan-Chen. 2016. "An Exploratory Study of the Subject Ontogeny of Eugenics in the New Classification Scheme for Chinese Libraries and the Nippon Decimal Classification." *Knowledge Organization* 43, no. 8: 594-608.

Lee, Wan-Chen. 2018. "Across Borders: The Concept and Applications of Eugenics in the CCL and the NDC." *Association for Information Science and Technology (ASIS&T)*, Vancouver, Canada.

Library of Congress. "Twain, Mark, 1835-1910." Last modified January 24, 2020. http://id.loc.gov/authorities/names/n79021164

Online Computer Library Center (OCLC). "VIAF: Convenient Access to Name Authority Files." Last modified December 20, 2019. https://www.oclc.org/en/viaf.html

Ranganathan, Shiyali Ramamrita. 1989. *Prolegomena to Library Classification*. New Delhi, India.: Ess Ess Publications.

Tennis, Joseph T. 2012. "The Strange Case of Eugenics: A Subject's Ontogeny in a Long-lived Classification Scheme and the Question of Collocative Integrity." *Journal of the American Society for Information Science and Technology* 63, no. 7: 1350-59. https://doi.org/10.1002/asi.22686

"Q7245." Wikidata. February 7, 2020. https://www.wikidata.org/w/index.php?title=Q7245&oldid=1108375903

Merriam-Webster.com Dictionary. "Vernacular." Last modified February 3, 2020. https://www.merriam-webster.com/dictionary/vernacular.

Zeng, Lei. 2019. "Interoperability." Knowledge Organization 46, no.2: 122-46. https://doi.org/10.5771/0943-7444-2019-2-122