



2-28-2021

さようなら (Sayōnara)

Follow this and additional works at: <https://scholarsarchive.byu.edu/jeal>

BYU ScholarsArchive Citation

(2021) "さようなら (Sayōnara)," *Journal of East Asian Libraries*: Vol. 2021 : No. 172 , Article 8.

Available at: <https://scholarsarchive.byu.edu/jeal/vol2021/iss172/8>

This Retirements is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in *Journal of East Asian Libraries* by an authorized editor of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

さようなら (Sayōnara)

This is my 116th—and last—blog post. I’m retiring at the end of November, something I’ve deferred as I’ve had such a great time hanging with all of you—staff at our Partner institutions, professionals from all corners of the library, archival, and information technology worlds, and my OCLC colleagues. But it’s time. Of all the ways I know to say “good-bye”, the Japanese *sayonara* is the most wistful: it literally means “if it must be so” (shorter than “parting is such sweet sorrow.”)

You’ve inspired me and taught me so much! I hope I’ve contributed meaningfully to the evolving discussions around metadata, linked data, and multilingual support to improve access to the information our communities want.

I am proud to have been part of the foundation of the Unicode Consortium. My work with the East Asian Character Set (Z39.64) proved that “Han Unification” was feasible— just as we have one code for the character “a” whether it’s used in English, French, German, Tagalog, Indonesian, etc. with different pronunciations, we can have one code for each Chinese (“Han”, or 漢) character common to Chinese, Japanese, and Korean. I advocated strongly for “Han Unification” and wrote a position paper on it in 1991. The [Unicode Chronology](#) highlights the stages of incorporating Han Unification into Unicode 1987-1992; Unicode became an international standard (ISO 10646) in 1993.

Unicode represented an “infrastructure revolution” for all of us. Those of you of a certain age may recall the days when there was a separate character set used in library systems, which included a range of diacritics to be used with other characters, often for use in transliterations of non-Latin scripts. But the data could only be shared and used by other library systems—if you copied/pasted data into another application it came out as gibberish. Non-Latin scripts were each defined by separate national character sets, and unless you used the same national character set, you could not read the text. Unicode, the result of a consortium including major computer corporations, software companies, and research institutions, changed all that. The scope included a far wider range of scripts than any other character set (the latest version includes over 140,000 characters). Because Unicode significantly decreased the costs of developing products for a global market, it was very quickly implemented in software applications. And Unicode included all the “combining diacritics” that libraries had used for decades. We take our ability to read non-Latin scripts in different applications and on websites for granted now.

Library catalogs still do not yet take advantage of the full range of scripts available in Unicode, however. Library users who read languages written in non-Latin scripts should be able to search and retrieve the metadata describing the resources written in those languages using the metadata in that script. Unfortunately, many of these non-Latin script resources are represented in catalogs only by transliteration, a barrier to access. (See my 2015 blog post, “[Transcription vs. Transliteration.](#)”) I’m pleased that OCLC has taken steps to remedy that situation, starting with the languages written in Cyrillic script. My colleagues Jenny

Toves, Bryan Baldus, and Mary Haessig blogged about this work earlier this year in “[кириллица в WorldCat](#)”.

Soon after the adoption of Unicode and (to me) amazingly quick implementation, I started bringing together the managers of technical services to discuss common issues and identify work that was needed to guide future developments that would improve the metadata underpinning the discovery of all the resources curated and managed by libraries, archives, and other cultural heritage organizations. Over the last 27 years this group evolved into the OCLC Research Library Partners [Metadata Managers Focus Group](#), which at one point included representatives from 63 Partner institutions in 12 countries spanning four continents. It spawned six working groups or task forces focused on particular issues and published reports of their investigations, such as [Registering Researchers in Authority Files](#) in 2014 and [Addressing the Challenges with Organizational Identifiers and ISNI](#) in 2016. My meta-synthesis of the Focus Group’s discussions over the last six years was recently published as an OCLC Research Report, [Transitioning to the Next Generation of Metadata](#). The recordings of our November 2020 discussions about the report are available as “past webinars” on the [Works in Progress Webinars](#) web page.

The Focus Group’s intense interest in who was implementing linked data and for what purposes led to a series of “International Linked Data Surveys for Implementers” I conducted between 2014 and 2018. A total of 143 institutions in 23 countries reported one or more linked data project or service. The results of these surveys are shared for the benefit of others wanting to undertake similar efforts on the [OCLC Research Linked Data Survey](#) web page.

One of the most rewarding highlights of my career was collaborating with my OCLC colleagues and OCLC members on “Project Passage,” a linked data Wikibase prototype which served as a sandbox in which librarians from 16 institutions could experiment with creating linked data to describe resources. The project was stimulating, educational, and *fun*! I enjoyed writing up what we learned with some of the participants in the 2019 report, [Creating Library Linked Data with Wikibase: Lessons Learned from Project Passage](#). This work generated another working group, [Archives and Special Collections Linked Data Review Group](#), drawn from the OCLC Research Library Partnership’s rare book, archives, and special collections communities, which explored key issues of concern and opportunities for archives and special collections in transitioning to a linked data environment, summarized in the 2020 OCLC Research Report, [Archives and Special Collections Linked Data: Navigating between Notes and Nodes](#).

I leave behind a set of publications and presentations. The *relationships* I’ve enjoyed with so many talented, inspiring staff within the OCLC Research Library Partnership I’ll treasure. I look forward to seeing what you all do in the coming years to leverage metadata and embed multilingualism into everything you do!

Sayōnara!

Karen Smith-Yoshimura, senior program officer, topics related to creating and managing metadata with a focus on large research libraries and multilingual requirements. Karen retired from OCLC November 2020.

Originally published on “Hanging Together: The OCLC Research Blog”

©2021 OCLC. Accessed February 2021, <https://hangingtogether.org/?p=8739>