

図書館におけるアルファベット至上主義の現状と展望

Challenging Anglocentric Library Infrastructures: Findability of Resources in Japanese Script in Europe as an Example

Kiyoka Koizumi and Dr. Cosima Wagner | The 34th EAJRS Conference at Sofia University, Bulgaria



Universiteit
Leiden
The Netherlands

UNIVERSITÄTS
BIBLIOTHEK

Discover the world at Leiden University

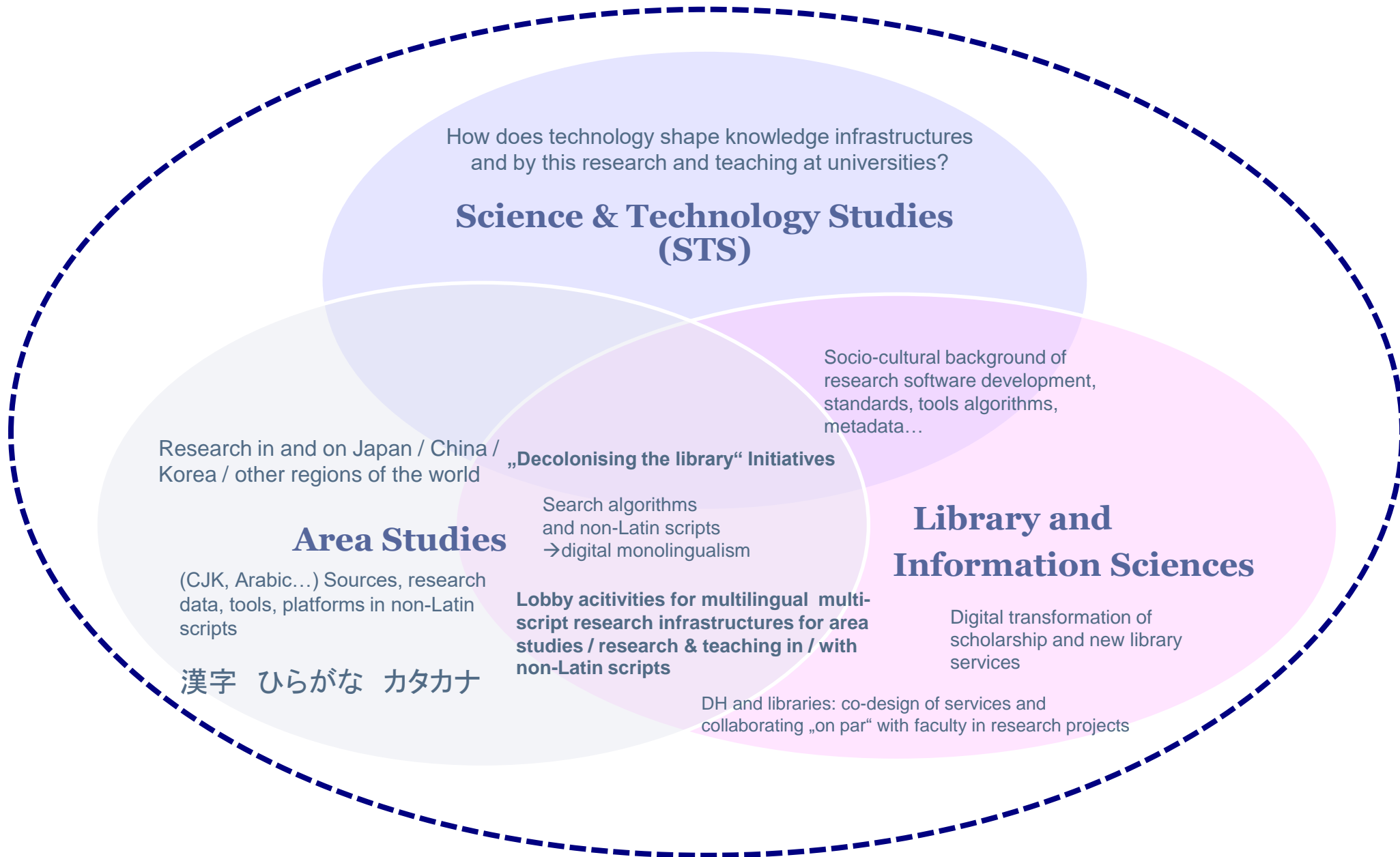
Index

- Introduction
- Why multilingualism matters?
- Issues of multilingualism & multiscryptuarism in libraries
- Examples in issues in findability
- What can & should be done?
- Discussion
- Selected bibliography

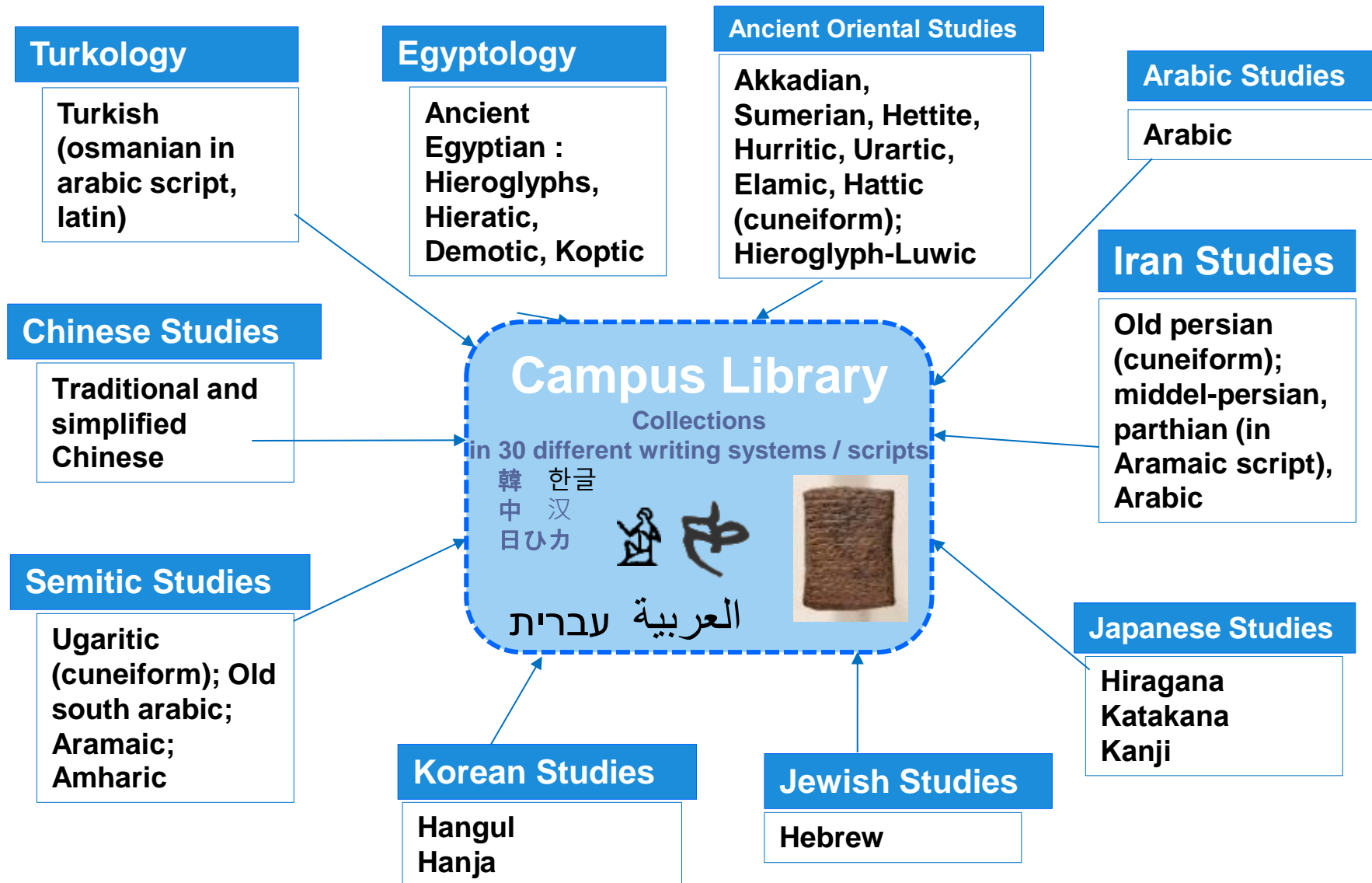


Introduction





Freie Universität Campus Library as a multilingual Area Studies knowledge hub



Leiden Institute for Area Studies (LIAS) and IIAS

LIAS is comprised of three research schools with expertise in different areas:

- **School of Asian Studies (SAS)**

Chinese Studies
Japanese Studies
Korean Studies
South Asian & Tibetan Studies
Southeast Asian Studies

- **School of Middle Eastern Studies (SMES)**

Arabic Studies
Assyriology
Egyptology
Hebrew & Aramaic Studies
Papyrology
Iranian and Central Asian Studies
Turkish Studies

- **School of Religious Studies (LUCSoR)**

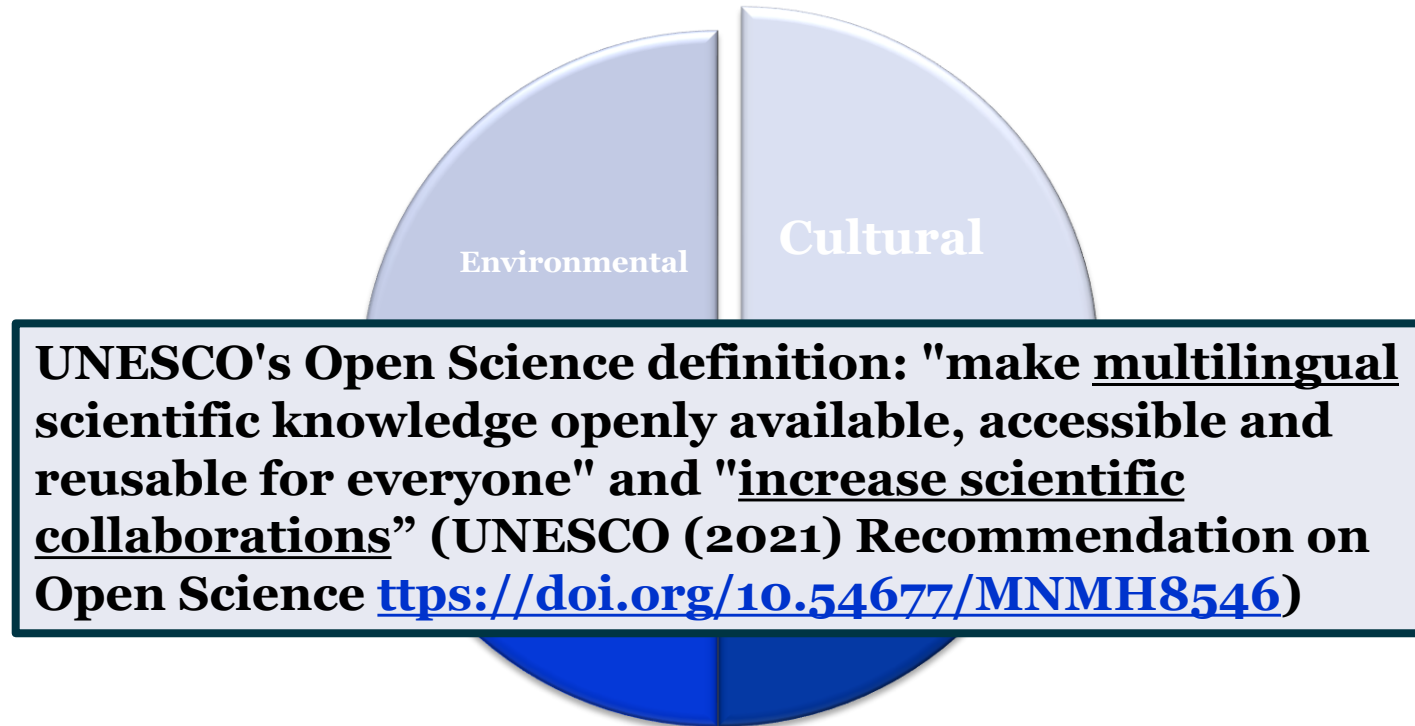
Religious Studies



Why multilingualism matters?



Culture and Language as a Dimension of Sustainability



“...the meaning of language and discourse in sustaining cultural heritage and the world’s languages, and cultural diversity. It has been acknowledged that culture constitutes a part of social sustainability, and covers such aspects as social justice and equity, social participation, economic needs and work, and, above all, awareness of the range of sustainability ideas.”

Kamińska, Anna M., Łukasz Opaliński, and Łukasz Wyciślik. 2022. "The Landscapes of Sustainability in the Library and Information Science: Systematic Literature Review" *Sustainability* 14, no. 1: 441, p.18-19, <https://doi.org/10.3390/su14010441>

Europeana report (2020)

Benefits of multilingual digital cultural heritage:

“The biggest benefit was noted as being that people can **access the information of other cultures and language groups.**

Followed by the point that **multilingualism promotes socially inclusive societies and mutual understanding of different cultures.**” (p.4)



Benefits, challenges and solutions for
multilingual digital cultural heritage: a report
from the Finnish Presidency Europeana meeting

February 2020



Europeana meeting under Finland's Presidency of EU Council - Espoo, Finland, 24-25 October 2019
Sebastian Ter Burg - CC BY

Source:https://pro.europeana.eu/files/Europeana_Professional/Publications/Output%20report%20from%20Finnish%20Presidency%20on%20multilingual%20digital%20cultural%20heritage..pdf

Multilingual ≠ multiscriptual

Nederlands and English



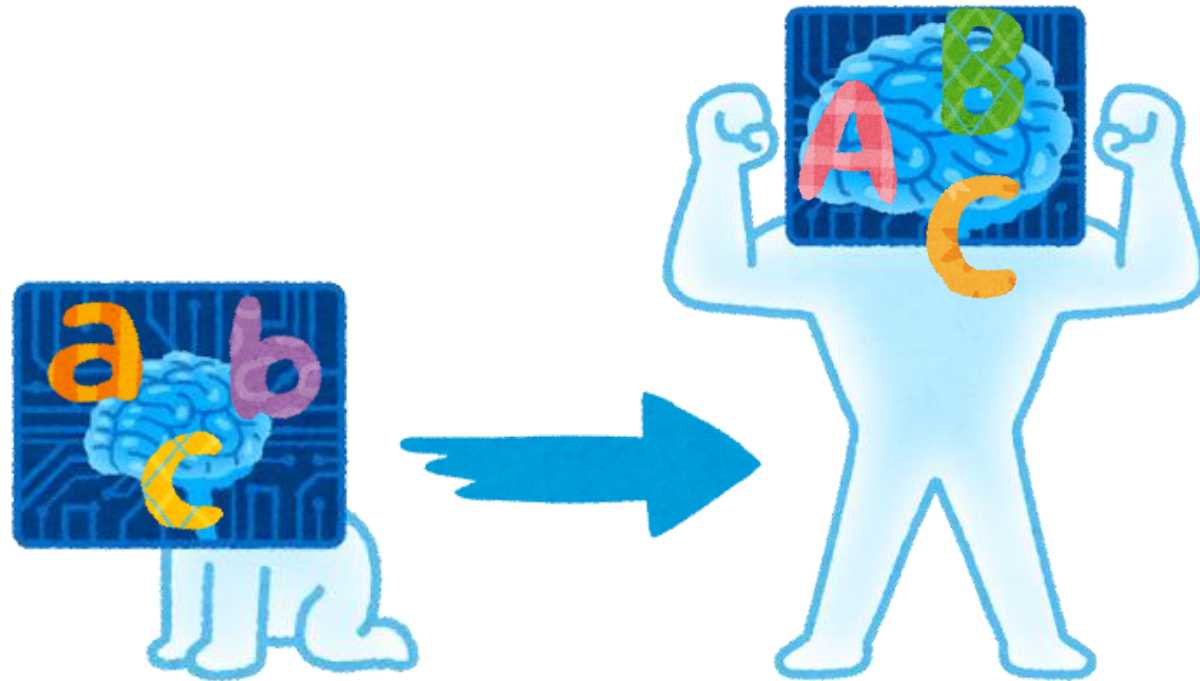
日本語と English

Meanwhile the reality is...



アルファベットによる、アルファベットのための

“[T]he meta-linguistic codes and algorithms of electronic communication which are revolutionizing almost every facet of knowledge and production, of information and projection, are founded on a sub-text, on a linguistic ‘pre-history’, which is fundamentally **Anglo-American** (in the ways in which we may say that Catholicism and its history had a foundational Latinity). **Computers and databanks chatter in ‘dialects’ of an Anglo-American mother tongue.**”



Steiner, George. (1998). *After Babel*. Aspects of Language and Translation. Oxford: Oxford University Press.

Disrupting Digital Monolingualism (2020)

- “Discussion around digital multilingualism is still fragmented (and sometimes marginalised), but momentum is growing across different fields to disrupt monolingual/**Anglo-centric** or **linguistically / geoculturally exclusionary** approaches to language technology.”
- “Access to open and **multidisciplinary** data, services and tools needs to be made easier and expanded to cover a wider range of languages. Training is also needed to lower the barrier to entry for those wishing to work with new digital tools in less technically-oriented fields of study.”
- “Digital multilingualism is not just about linguistics, or language. It has a **cultural and socio-political dimension** which is crucial in studying increasingly transcultural and translingual dynamics and in helping us to understand what are ultimately human-designed and complex (digital) cultural artefacts.”



Paul Spence. (2021). Disrupting Digital Monolingualism: A report on multilingualism in digital theory and practice (1.0). Disrupting Digital Monolingualism (DDM). Language Acts & Worldmaking project. <https://doi.org/10.5281/zenodo.5743283>

The sustainability of libraries and knowledge infrastructures (especially in the **Global North**) should be challenged...

... with regard to embracing more linguistic and geocultural diversity as well as the **accessibility** and **discoverability** of their collections!



“...language is one of the most important **barriers** that pose an obstacle to the sustaining of global knowledge diversity.”

Kamińska, Anna M., Łukasz Opaliński, and Łukasz Wyciślik. 2022. "The Landscapes of Sustainability in the Library and Information Science: Systematic Literature Review" *Sustainability* 14, no. 1: 441, p.18-19. <https://doi.org/10.3390/su14010441>

Issues in library specific context



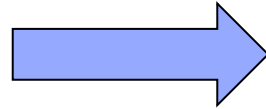
“Barriers to access that hinder area studies research” (Pitman) = challenges of area studies librarian’s work

- Existence of “hidden” collections,



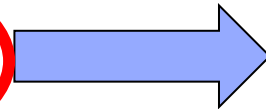
Lack of expertise, lack of positions for area studies in libraries, differing publication cultures

- Inappropriate classification schemes,



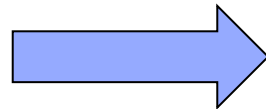
Finding workarounds, inaccessibility of sources in non-Latin scripts...time budgets...

- Variations in cataloguing, metadata, and technical standards for foreign language collections, and



Constant rising awareness activities and the need to bring together all departments concerned (power structures/hierarchies within libraries...)

- License restrictions



Lack of availability of electronic resources from the regions

→ Lack of area studies librarian positions (incl. acknowledgement of qualifications, training)



Payment gap of subject librarians & liaison librarians

Source: Pitman, Lesley (2015): Resource discovery and systems: Barriers to access in area studies. In: *ibid.*, *Supporting Research in Area Studies: A Guide for Academic Libraries*. Chandos Publishing, p. 69-80. <https://www.sciencedirect.com/science/article/abs/pii/B9781843347903000068>

“CJK”

Search Configurations for Different Languages

[https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_\(English\)/010Getting_Started/050Alma_User_Interface_%E2%80%93_General_Information/Searching_in_Alma/Search_Configurations_for_Different_Languages](https://knowledge.exlibrisgroup.com/Alma/Product_Documentation/010Alma_Online_Help_(English)/010Getting_Started/050Alma_User_Interface_%E2%80%93_General_Information/Searching_in_Alma/Search_Configurations_for_Different_Languages)

CJK Languages

Chinese and Korean Characters

Alma does hiragana to katakana transliteration, traditional Chinese to simplified Chinese transliteration, and splits words into bigrams and unigrams. See the [ICU Transform Filter](#) for more information.

Alma also does Hanja to Hangul transliteration. The sorting is unique to the Korean language.

Japanese Characters

For institutions that have the Japanese searching setup for Repository Search, Browse Bib Headings, and Browse Auth Headings, Alma performs the following:

- Punctuation removal
- Normalization between Hiragana and Katakana
- Iterated character normalization
- Normalization of variant Kanji characters

Not specified

CJK Punctuation Handling

For institutions that have the Chinese, Hong Kong, Japanese, or Korean searching setup, all punctuation marks are removed during indexing when they appear within CJK text. However, the punctuation remains when you are searching. This helps to ensure that the best results are retrieved. Note that the display of CJK content continues to show the punctuation.

See the known search issue related to punctuation that is described in the [note](#) in the [Using Advanced Search](#) section.

Example 1 頻出！

Original materials do not contain 旧字体 but records do

- 說 ↔ 説

(小説 = novel, 概説 = overview, 天動説 = Ptolemaic theory)

- 戸 ↔ 戶

(江戸時代 = Edo period, 戶籍 = census, 神戸市 = Kobe city)

- 録 ↔ 錄

(記録 = record, document, history, 目錄 = inventory, catalog, 付録 = supplement)



Example 2

Original material contains 旧字体

『昭和戦時期の娯楽と検閲』(2021)

閱 ↔ 閱



LDR	03254cam a2200601 i 4500
001	9940181202402711
005	20230428110313.0
008	210201s2021 ja b 001 0 jpn c
010	__ la 2020453223
020	__ la 9784642039017
020	__ la 4642039015
035	__ la (OCoLC)1235741397
035	__ la (OCoLC)on1235741397
040	__ la JaToTRC b eng e rda c TRCLS d ZCU d OCLCO d OCLCF d OCLCO d CGU
042	__ la pcc
043	__ la a-ja--
050	_.4 a ML3501 b .K365 2021
066	__ c \$1
084	__ a 772.1 2 njb/9
100	1_ a Kaneko, Ryōji, d 1984- 0 (DLC)no2021103136 e author. 4 aut 1_ a 金子龍司, d 1984- e author. 4 aut
240	10 a Senjika Nihon ni okeru goraku tōsei no kenkyū 10 a 戦時下日本における娯楽統制の研究
245	10 a Shōwa senjiki no goraku to ken'etsu / c Kaneko Ryōji cho. 10 a 昭和戦時期の娯楽と検閲 / c 金子龍司著.
264	_.1 a Tōkyō : b Yoshikawa Kōbunkan, c 2021. _.1 a 東京 : b 吉川弘文館, c 2021.
300	__ a 5, 287, 4 pages ; c 22 cm
336	__ a text 2 rdacontent
337	__ a unmediated 2 rdamedia
338	__ a volume 2 rdacarrier

Enhancement Request of Harvard University Library to the commercial knowledge infrastructure provider ExLibris in March 2023 regarding the Latin-script bias of the discovery system Primo:

- “8152 Primo VE: **improve support for diversity and equity** through automatic handling of non-Latin data in 880 fields
- 8153 Display non-Latin script fields in CDI, **important for equity and inclusion**”

→ Problem remains: lack of visibility/accessibility of and access to research or sources in non-Anglophone languages and non-Latin scripts in library catalogues

→ Barrier for supporting international, interdisciplinary and open research!

→ Problem now addressed as an issue of “support for diversity and equity”

Source:

<https://wiki.harvard.edu/confluence/display/LibraryStaffDoc/HOLLIS+Known+Issues+and+Enhancement+Requests#HOLLISKnownIssuesandEnhancementRequests-PleasevoteformtheseenhancementrequestsInIdeasExchange>

What can & should we do?



What can & should be done? (1)

- Academic infrastructure strategies must refer to multilingualism and multi-scriptuality as elementary components of diversity and equity policies.
- Multilingualism and multi-scriptuality must be added to the specifications/requirement lists of library IT software & platforms.
- Libraries must give more attention to build multilingually enabled digital information landscapes in order to pave the way for a culture of globally networked thinking.

What can & should be done? (2)

- Use of open source systems instead of commercial locked-in offers from data brokers.
- Continuously enhance catalogue and discovery systems to accommodate all scripts relevant to the collection.
- Create personnel resources and expertise through further education and training programs.
- Consistent application guidelines for transliteration.

What can & should be done? (3)

- Multilinguality and -scriptuality must play a greater role in decision-making bodies of libraries
- Embedding multilinguality as a topic in library education & training
- More opportunities for career changers i.e. from area studies disciplines in libraries
- Including multilinguality/multiscriptuality in project and infrastructure development as a requirement from the outset

Discussion and Q&A



How can EAJRS libraries and librarians collaborate in order to co-design more sustainable, multilingual, geo-cultural diverse knowledge infrastructures for (digital) scholarship in Japan(ese) Studies?



Bibliography (selection) 1

Coleman, Catherine Nicole (2020). "Managing Bias When Library Collections Become Data." *International Journal of Librarianship* 5.1: 8–19.

Fiormonte, D. (2021). "Taxation against Overrepresentation? The Consequences of Monolingualism for Digital Humanities". Kim D. and Koh A. (eds), *Alternative Historiographies of the Digital Humanities*. Santa Barbara: Punctum Books, 333–376.

Ghorbaninejad, M, Gibson, N. P. , Wrisley D.J. (2023). "Right-to-Left (RTL) Text: Digital Humanists plus Half a Billion Users". In Gold, M.K. and Klein, L. (eds), *Debates in the Digital Humanities 2023*. Minneapolis: University of Minnesota Press, 47–73.

Glasberg, E. (2024). "What Does Global Language Loss and Justice Look Like?" <https://news.columbia.edu/news/what-does-global-language-loss-and-justice-look>.

Golumbia, D. (2013). "Postcolonial Studies, Digital Humanities, and the Politics of Language". *uncomputing*, May 31.

<https://web.archive.org/web/20160623210924/http://www.uncomputing.org/?p=241>.

Grallert, T. (2022). "Open Arabic Periodical Editions: A Framework for Bootstrapped Scholarly Editions Outside the Global North." *Digital Humanities Quarterly* 16(2).

<https://web.archive.org/web/20231018192515/http://www.digitalhumanities.org/dhq/vol/16/2/000593/000593.html>.

Horvath, A. (2021). "Enhancing Language Inclusivity in Digital Humanities: Towards Sensitivity and Multilingualism." *Modern Languages Open* 1(26), 26. <http://doi.org/10.3828/mlo.voio.382>.

Horvath, Aliz, Cornelis van Lit, Cosima Wagner and David Joseph Wrisley (2023). "Centering Multilingual Users: Thinking through UX Personas in the DH Community." *Digital Studies / Le champ numérique*.

Horvath, Aliz, Cornelis van Lit, Cosima Wagner, and David Joseph Wrisley. (2024). "Towards Multilingually-Enabled Digital Knowledge Infrastructures: a Qualitative Survey Analysis" in Viola, Lorella and Paul Spence (eds.). *Multilingual Digital Humanities*. Routledge.

Bibliography (selection) 2

Jiang, Shuyon (2007a.): “Lost in Translation: The Treatment of Chinese Classics in the Library of Congress Classification”, *Cataloging & Classification Quarterly*, 45(1), p. 3-14, https://doi.org/10.1300/J104v45n01_02

Liu, L. H., Rao, A., and Silverman, C. A. (eds). (2024). *Global Language Justice*. New York: Columbia University Press.

Malínek, V., Umerle, T., Gray, E., Heibi, I., Király, P., Klaes, C., Korytkowski, P., Lindemann, D., Moretti, A., Panušková, C., Péter, R., Tolonen, M., Tomczyńska, A., Vimr, O. (2024). “Open Bibliographical Data Workflows and the Multilinguality Challenge”, *Journal of Open Humanities Data*, 10, 27, 1-14. <http://doi.org/10.5334/johd.190>

Peggy Nzomo, Pamela McKenzie, Isola Ajiferuke & Liwen Vaughan (2021) Towards a Definition of Multilingual Information Literacy (MLIL): An Essential Skill for the 21st Century, *Journal of Library Administration*, 61:7, 897-920, DOI:10.1080/01930826.2021.1972737

Phillipson, R. (2013). *Linguistic Imperialism Continued*. Abingdon: Routledge.Princeton University. (n.d.). “New Languages for NLP.” <https://newnlp.princeton.edu>

Schmidt, Nora. 2020. *The Privilege to Select. Global Research System, European Academic Library Collections, and Decolonisation*. Lund: Lund University: Faculties of Humanities and Theology. DOI: 10.5281/zenodo.4011296

Spence, P. (2021). “Disrupting Digital Monolingualism: A Report on Multilingualism in Digital Theory and Practice”, London: Language Acts and Worldmaking Project. <http://doi.org/10.5281/zenodo.5743283>.

Contact

Dr. Cosima Wagner

Freie Universität Berlin University Library

Tel: +49-30-838-60307

Email: cosima.wagner@fu-berlin.de

Web: <https://www.fu-berlin.de/en/sites/ub/ueberuns/team/wagnerc/index.html>

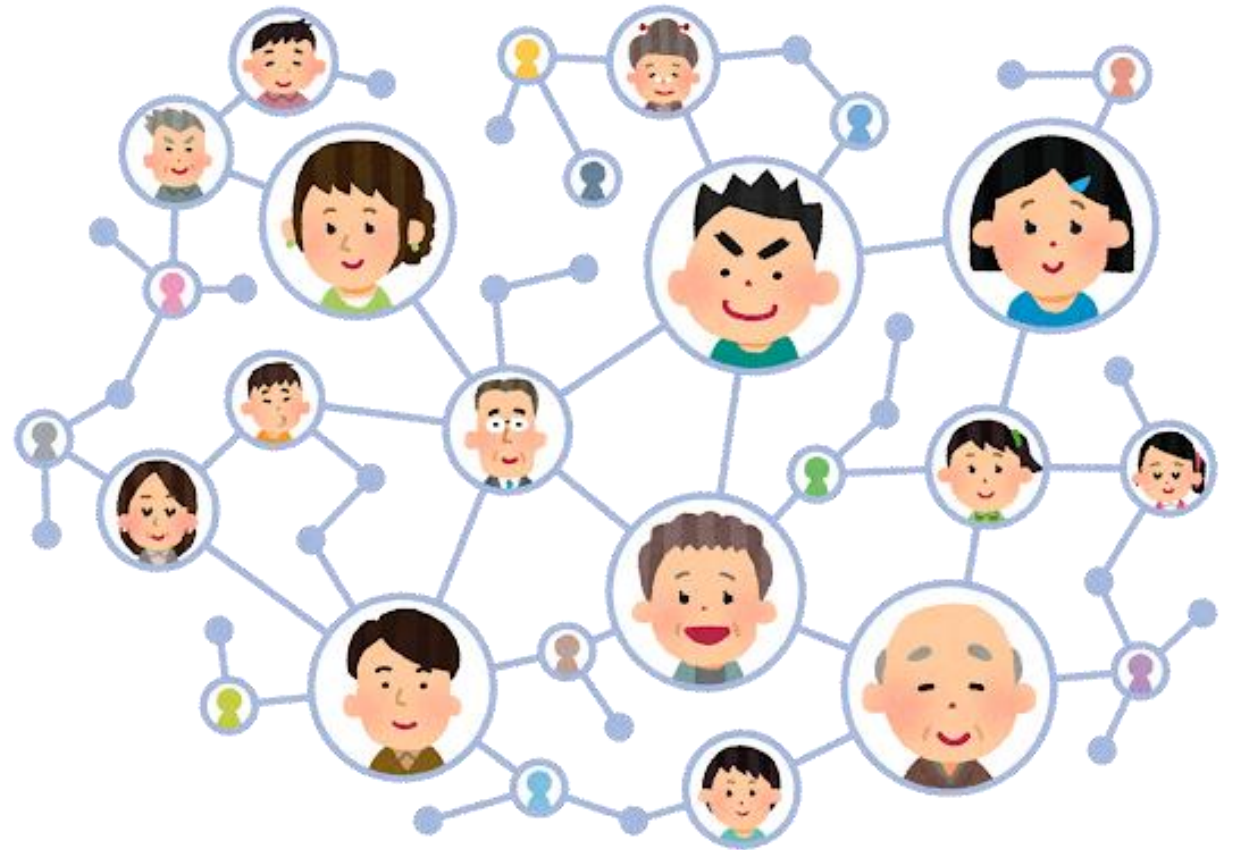
Kiyoka Koizumi

Leiden University Libraries

Tel: +31-71-527-6117

Email: k.koizumi@library.leidenuniv.nl

LinkedIn : <https://www.linkedin.com/in/kiyokakoizumi/>



ご清聴
ありがとうございました



Universiteit
Leiden
The Netherlands

Discover the world at Leiden University