

Kirishitan Texts in the Digital Age デジタル時代におけるキリシタン史料

James Harry Morris, MTheol, PhD, FRAS
モリス、ジェームズ・ハリー

Waseda University, Waseda Institute for Advanced Study
早稲田大学 高等研究所



Project Rationale and Aims

Roman Catholic missionaries in Japan during the 16th and 17th centuries created a vast body of literature including letters and printed texts known as *Kirishitan-ban* キリシタン版 written in Japanese, Latin, Portuguese, and Spanish. Following the outlawing of Christianity in 1613, the number of texts on Christianity grew with the publication of *Haiyasho* 排耶書 (anti-Kirishitan treatise) and the composition of materials related to religious surveillance.

Scholars, students, and others interested in exploring and analyzing this vast body of texts with digital tools and methods face numerous challenges, particularly if they are only casually engaged with the digital humanities or only just learning how to integrate digital methods into their work or studies. These challenges are based on:

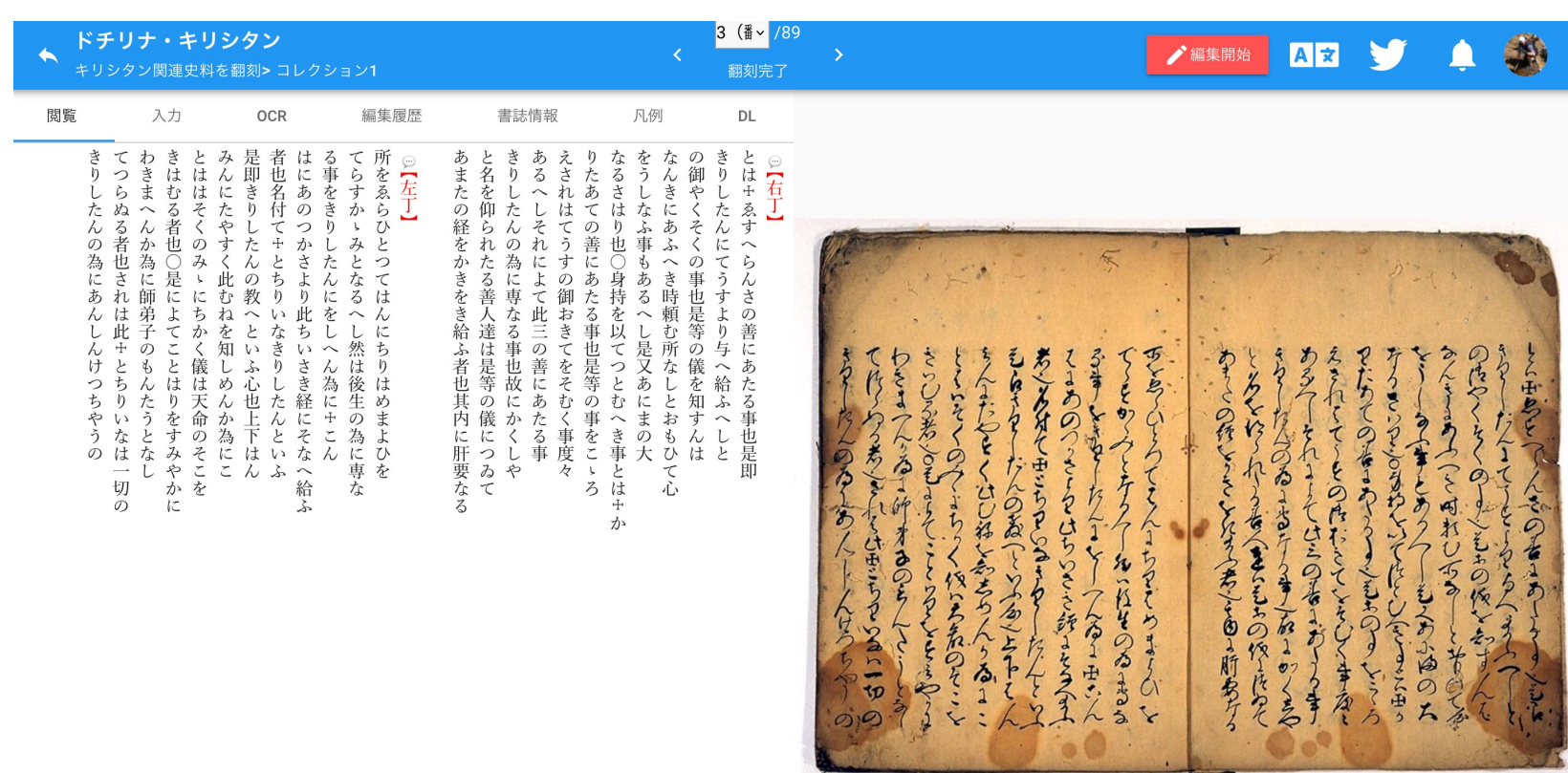
1. The nature of the texts.
 - I. Texts in different languages and scripts.
 - II. Multiple language and script combinations in a single text.
 - III. Multidirectional text.
 - IV. Use of multiple types of printing and composition (moveable type, woodblock, handwritten).
2. The limitations of digital resources and tools.
 - I. Lack of digitized sources and transcriptions.
 - II. Paucity of dictionary resources.
 - III. Poor quality OCR.

The “*Kirishitan-ban* in the Digital Age: A Study of the Opportunities and Limitations of Applying Digital Methods to *Kirishitan-ban*” project seeks to critically analyze the opportunities and challenges that these texts create for scholars working in the digital age, survey the limitations of pre-existing digital tools, and develop new tools and resources.

Creating Machine-Readable Data

At present machine-readable versions of most *Kirishitan-ban*, *Haiyasho* and related texts do not exist, but these are essential for anyone interested in conducting a digital textual analysis.

We are hosting a project on the collaborative transcription platform *Minna de honkoku* みんなで翻刻 to address this need. *Minna de honkoku* allows anyone to try transcribing pre-modern Japanese texts with the help of other users and AI tools. Our project includes 7 *Kirishitan-ban*, 23 *Haiyasho*, and 1 related text of which 7 have been completed so far. The resultant transcriptions can be downloaded in a variety of formats and used freely by any researcher.



Another method for creating machine-readable transcriptions is to use optical character recognition (OCR) technology. Sadly, OCR tools designed to work with pre-modern Japanese struggle with *Kirishitan* texts. Experimenting with *Miwo* on *Kirishitan-ban* returned error rates of 15.1% to 33.3%. Relying on these tools to quickly transcribe texts is, therefore, not a feasible option at present.

I created experimental OCR models with *Transkribus* for *Kirishitan-ban* written in Latin script. These models have achieved error rates of 1.29% and 2.68% and can be shown to be much more accurate than off the shelf OCR tools such as *ABBYY FineReader*. Comparative results are displayed in the following table:

Transkribus Model (Validation Set)	ABBYY FineReader 11.0 (Internet Archive)	Google
Fido no ribai vo tōri motomoru va von imaximè de gozà-	Kdo no ribal vo tōri motomoru va vonimaximedegozd-	Fido no ribai vo tori motomoru va von imaximè de gozd
Saiban mēlaruru damiō no cacurete tōtta mōno vo caie-	Saiban mēlaruru damid no cacurete totta mono voaic-	Saiban mēlaruru damiō no cacurete tōtta mono vo caie
Māta fito no. vie vo iafui xi, lōno cōto vōba varū fata tō-	Mata fito no vie vo jafui xi, funo coto voba varu fatato-	Māta fito no vie vo jasui xi, sono cōto vōba varū fata to

We hope to develop these models for public release in the future.

Special Characters

The difficulties associated with working with texts related to Christianity extend down to the most basic level — characters and the way they appear on a computer screen. Texts often feature multiple languages and scripts which can create issues for OCR platforms and digital analysis. However, the issues are even more basic than this – sometimes it isn't even possible to render certain characters on a computer.

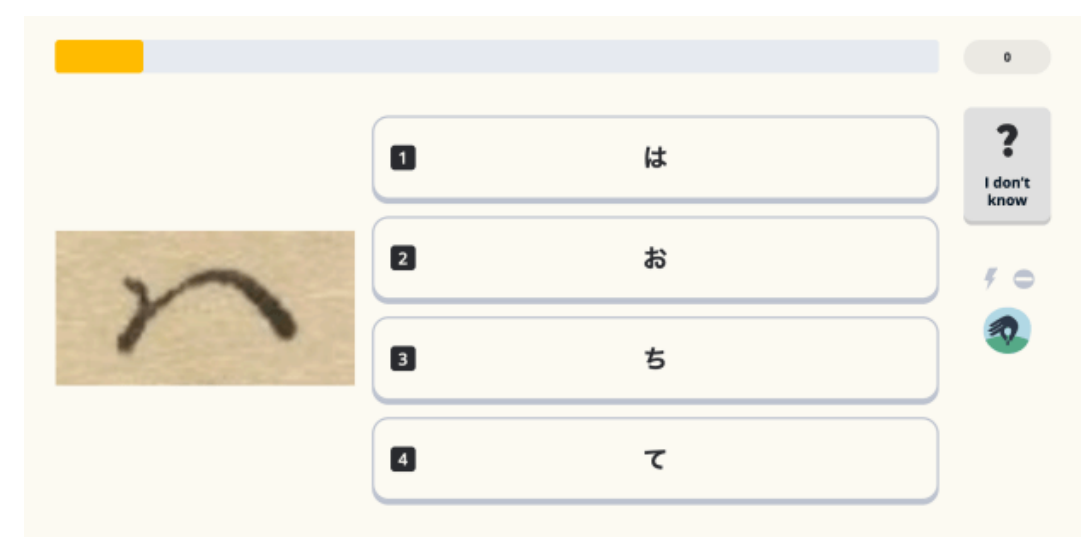
/apo	ꝑꝝ	/ds2	ꝢꝢ	/jxo	ꝯꝯ	/pu	ꝑꝰ
/apt	ꝑꝝꝰ	/ds3	ꝢꝢꝰ	/ma	ꝰꝰꝰ	/sm	ꝰꝰꝰ
/bta	ꝰꝰꝰ	/ds4	ꝢꝢꝰꝰ	/pa	ꝰꝰꝰ	/sma	ꝰꝰꝰꝰ
/bto	ꝰꝰꝰꝰ	/fo	ꝰꝰ	/pe	ꝰꝰ	/xao	ꝰꝰꝰꝰ
/de	ꝰꝰꝰ	/jao	ꝰꝰꝰ	/po	ꝰꝰ	/xo	ꝰꝰꝰ
/de2	ꝰꝰꝰꝰ	/jo	ꝰꝰ	/pp	ꝰꝰ	/x1	ꝰꝰꝰ
/dpo	ꝰꝰꝰꝰ	/jo2	ꝰꝰꝰ	/pt	ꝰꝰ	/x2	ꝰꝰꝰ
/ds	ꝢꝢ	/js	ꝰꝰ	/pta	ꝰꝰꝰ	/x3	ꝰꝰꝰꝰ

As part of the project, we created a font package that allows some special characters (ligatures) to be input into a word processor. This represents the first time that the public can input these characters into their computers and opens the possibility for new types of digital analysis on the usage of these characters.

Educational Resources

There is a huge lack of educational resources for those interested in learning how to read *Kirishitan-ban*. This presents particular difficulties for students, however, a paucity of dictionary resources, for example, also influences tasks such as text segmentation which is an essential first step for digital analysis and the type of textual analysis that we can perform. This lack of resources, therefore, affects both education and research. Experiments conducted on four major online Japanese dictionaries found that they contained less than 32% of specialised terms found in *Kirishitan-ban*. As such, one long-term goal of the project is to create an online dictionary of *Kirishitan* terminology.

To increase the number of educational resources available for learning how to read *Kirishitan-ban* we have released two flashcard decks for individual and classroom use and have trialed these amongst students learning early modern Japanese palaeography.



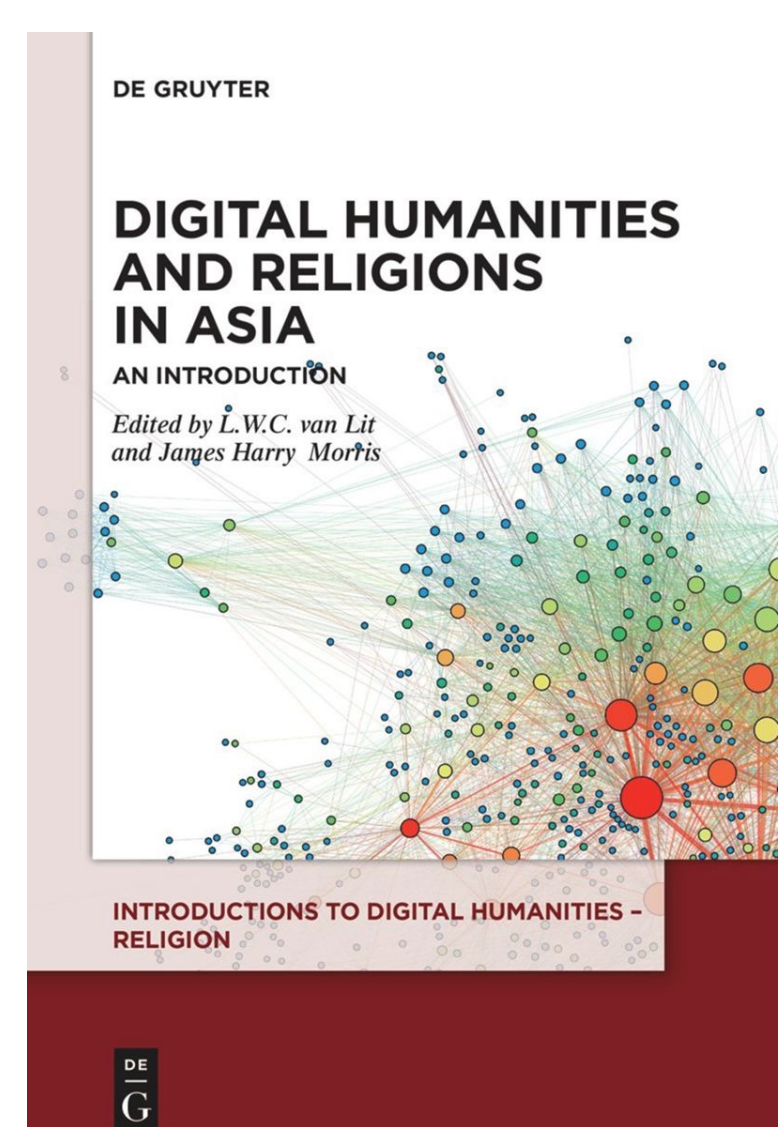
Publications

“A new way to input ligatures found in *Kirishitan-ban*,” *Manuscript Studies* (Forthcoming).

“Measles Epidemics during the Edo Period: An Analysis using DANJURO ver. 7.0,” *Hungarian Journal of English and American Studies* (Forthcoming).

“Missiology in the Digital Age: Challenges and Opportunities for the Study of Pre-Modern Christianity in Japan,” in *Digital Humanities and Religions in Asia: An Introduction*, edited by L.W.C. van Lit and J.H. Morris (Berlin: De Gruyter, 2023), 205-230.

「明治三年の高佐村から非人番が登場する『切支丹宗門御改帳』について」 早稲田大学高等研究所紀要 (Under Review). 共著者: アンナ・ドゥリーナ (京都大学)。



Funding and Collaborators

Funding

DNP文化振興財団のグラフィック文化に関する学術研究助成 (2023年度～2024年度)
早稲田大学特定課題(研究基盤形成) (2023年度)

Collaborators

Joseph Bills (Keio University)
みんなで翻刻 — <https://honkoku.org/>
大曲都市 — <https://www.tosche.net/fonts>



SCAN ME