

Exploring the Potential of Utilizing Data from Sanskrit Literature

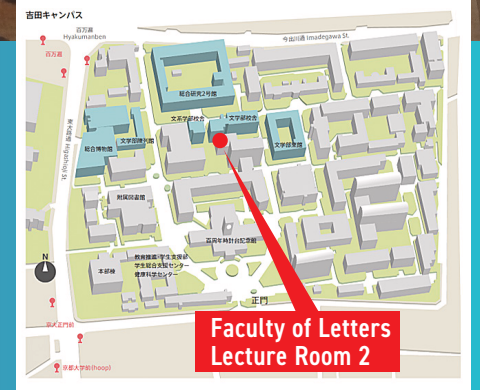
サンスクリット文献のデータ利用の可能性

March 19-20, 2024 in person and online

Lecture Room 2, Faculty of Letters,
Kyoto University (Yoshida Main Campus)

2024年3月19日(火)、20日(水・祝) オンライン併用

京都大学文学研究科 第2講義室(京都大学吉田キャンパス、本部構内)



In recent years, techniques in the digital domain such as digitalization of materials and cross-referential search in a database have become commonplace in literary studies. However, in the field of Sanskrit studies, there remain numerous challenges in creating digital data and conducting data analysis due to the unique characteristics of the language and materials. In this workshop, we aim to discuss, from the perspectives of both digital humanities and literary studies, effective ways of organizing data and methods of utilizing data that are relevant to studies in Sanskrit literature. Our aim is to explore the potential for future research that opens up new perspectives and new frontier in this field.

資料のデジタル化や横断検索など、情報学的研究手法は文献学にとって身近なものになった。その一方で、サンスクリット文献学においては、その言語や資料の特殊性に対応したデジタルデータの作成やデータ分析について、未だ課題が多い。本ワークショップでは、サンスクリット文献学にとって有効なデータの在り方、データ利用の方法について、人文情報学と文献学双方の視点から議論を行い、新たな視点、新たな領域を拓く今後の研究の可能性を見出したい。

Sessions

March
19
(Tue)

Development and Utilization of OCR for Indian Scripts

インド文字OCRの開発と活用

Creating and Utilizing Data of Sanskrit Texts and Non-Textual Data

サンスクリット語データの作成と活用、テキストデータ以外のデータの活用

March
20
(Wed)

Analysis of Internal Structure and Chronology Using Sanskrit Corpus, and its Visualization

サンスクリットコーパスを用いた文献構造分析と年代分析、およびその可視化

Please register using the Google Form on the official website of the project.

URL: <https://ancientindia-datascience.jp>

Please choose between on-site participation or online participation. (A Zoom meeting link will be sent via email.)

Participation is free of charge.

本プロジェクトウェブサイト上のGoogleフォームより参加登録をお願いいたします。

URL: <https://ancientindia-datascience.jp>

現地参加かオンラインかを選択してください。(ご登録後e-mailにてZoomミーティングのリンクを送付いたします。)

参加費は無料です。

After the conclusion of the first day on the 19th, a reception will be held in Lecture Room 1 of the Faculty of Letters (next to the workshop venue). (We will collect a fee from those who attend.) We encourage you to join the reception as well.

Registration will be accepted until the end of the workshop, but for those wishing to attend the reception, please apply by March 5th.

19日(第一日目)終了後、文学研究科 第1講義室(ワークショップ会場の隣の部屋)で懇親会を開催いたします。(ご参加の方から懇親会費をもらい受けます)懇親会にもぜひご参加ください。

参加登録はワークショップ終了まで受け付けいたしますが、懇親会に参加ご希望の方は、準備の都合上3月5日までに申し込みをお願いいたします。

Program

March 19 (Tue) 10:30-18:30

10:30-10:40 Opening Remarks **Kyoko Amano**(Kyoto University) 天野 恭子(京都大学)

Development and Utilization of OCR for Indian Scripts

インド文字OCRの開発と活用

10:40-11:20 The Paleographic Database Indoskript - Design and Future Applications **Oliver Hellwig** (University of Zurich)

11:20-11:45 break

11:45-12:45 Development of Devanagari OCR: From Typeset to Handwritten Letters
Takahiro Kato (University of Tokyo) **Yuki Tomonari** (University of Tokyo)
デーヴァナーガリーOCRの開発:活字から手書き文字へ 加藤 隆宏(東京大学) 友成 有紀(東京大学)

12:45-13:15 Discussion

13:15-15:00 lunch break

Creating and Utilizing Data of Sanskrit Texts and Non-Textual Data

サンスクリット語データの作成と活用、テキストデータ以外のデータの活用

15:00-16:00 Efficiency in Text Reading through OCR and Text-mining **Yoichi Iwasaki** (Nagoya University)
OCRとテキストマイニングを活用した文献読解の効率化 岩崎 陽一(名古屋大学)16:00-17:00 Creating Vedic Texts and the Challenge of Identifying the Same Words **Yuzuki Tsukagoshi** (University of Tokyo)
ヴェーダ文献テキストの作成とその困難 —「同じ語」とはなにか— 塚越 柚季(東京大学)

17:00-17:15 break

17:15-18:00 Visual Analytics of Intertextual Relationship Using a Mantra Index
Hiroaki Natsukawa (Osaka Seikei University) **Kyoko Amano** (Kyoto University)
マントラインデックスを利用した文献間の影響関係の視覚的分析 夏川 浩明(大阪成蹊大学) 天野 恭子(京都大学)18:00-18:30 Discussion Comments by **Kiyonori Nagasaki**(International Institute for Digital Humanities) 永崎 研宣(人文情報学研究所)

18:45-21:00 reception (Faculty of Letters, Lecture Room 1) (文学研究科 第1講義室)

March 20 (Wed) 9:30-16:30

Analysis of Internal Structure and Chronology Using Sanskrit Corpus, and its Visualization

サンスクリットコーパスを用いた文献構造分析と年代分析、およびその可視化

9:30-10:30 Dating the Vedic Corpus **Oliver Hellwig** (University of Zurich)

10:30-10:45 break

Analysis of Similarity in Yajurvedic Texts

ヤジュルヴェーダ文献間類似度の分析

10:45-11:00 Background of Similarity Analysis in Yajurvedic Texts **Kyoko Amano** (Kyoto University)
ヤジュルヴェーダの文献間類似度分析の背景 天野 恭子(京都大学)11:00-11:45 A Corpus Linguistic Analysis of Intertextuality in Vedic Literature using TRACER and Stylo **So Miyagawa** (NINJAL)
ヴェーダ文献におけるTRACERとStyloを用いた間テキスト性のコーパス言語学的分析 宮川 創(国立国語研究所)11:45-12:30 Reassessment of Similarity Measures for Sanskrit: Word2Vec and Transformers **Yuki Kyogoku** (Leipzig University)
サンスクリット文献に有効な類似度推定の方法論再評価: Word2Vecと Transformers 京極 祐希(Leipzig University)12:30-12:40 Evaluation of Analysis Results **Kyoko Amano** (Kyoto University)
分析結果の評価 天野 恭子(京都大学)

12:40-13:10 Discussion

13:10-15:00 lunch break

15:00-16:00 VL²: Visualization of Linguistic Layers in Vedic Literature
Hiroaki Natsukawa (Osaka Seikei University) **Kyoko Amano** (Kyoto University)
ヴェーダ文献における言語層の可視化 夏川 浩明(大阪成蹊大学) 天野 恭子(京都大学)

16:00-16:20 Discussion

16:20-16:30 Closing Remarks

[主催] 科研費基盤研究B「デーヴァナーガリー文字OCRの実用化と文献データベースの利活用に向けた応用研究」(2023-2026FY) 研究代表者:加藤隆宏(東京大学)
国際共同研究強化B「ヴェーダ文献における言語層の考察とそれを利用した文献年代推定プログラムの開発」(2021-2026FY) 研究代表者:天野恭子(京都大学)

[共催] 研究所共同研究拠点利用公募研究班「インドにおける『循環的存在論』の形成」代表:手嶋英貴(龍谷大学)
科研費挑戦的研究(萌芽)「インド古典文献研究の基礎作業支援環境の構築—ウダヤナ文献研究における実践を通して」(2022-2023FY) 研究代表者:岩崎陽一(名古屋大学)

[助成] 京都大学人文科学研究所共同拠点利用、国際ミーティング助成