

京都大学人文科学研究所国際研究ミーティング実施報告書

1. 国際研究ミーティングの名称

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)

2. 主宰責任者氏名

加藤 隆宏 (東京大学大学院 人文社会系研究科・准教授)

手嶋 英貴 (龍谷大学 法学部・教授)

天野 恭子 (京都大学 文学研究科・人文学連携研究者)

石井 美保 (京都大学 人文科学研究所・准教授)

3. 開催日時等およびプログラム (講演者名または報告者名を明記してください)

① 日時: 2024年 3月 19日 10:30~18:30

場所: 京都大学文学研究科 第2講義室

内容:

セッション1: Development and Utilization of OCR for Indian Scripts

(インド文字 OCR の開発と活用)

講演 The Paleographic Database Indoskript - Design and Future Applications,

Oliver Hellwig (University of Zurich, Research Fellow)

講演 Development of Devanagari OCR: From Typeset to Handwritten Letters

(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 (東京大学准教授)、友成有紀 (東京大学特任研究員)

セッション2: 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)

講演 Efficiency in Text Reading through OCR and Text-mining

(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 (名古屋大学准教授)

講演 Creating Vedic Texts and the Challenge of Identifying the Same Words

(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 (東京大学助教)

講演 Visual Analytics of Intertextual Relationship Using a Mantra Index

(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 (大阪成蹊大学准教授)、天野恭子 (京都大学)

コメント: 永崎 研宣 (人文情報学研究所、主席研究員)

② 日時: 2024年 3月 20日 9:30~16:30

場所: 京都大学文学研究科 第2講義室

内容:

セッション3: Analysis of Internal Structure and Chronology Using Vedic 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)

講演 Dating the Vedic Corpus, Oliver Hellwig (University of Zurich, Research Fellow)

講演 (共著) Analysis of Similarity in Yajurvedic Texts (ヤジュルヴェーダ文献間類似度の分析)

Background of Similarity Analysis in Yajurvedic Texts

(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)

A Corpus Linguistic Analysis of Intertextuality in Vedic Literature using TRACER and Stylo(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)

Reassessment of Similarity Measures for Sanskrit: Word2Vec and Transformers
(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、博士課程、プロジェクト研究員)

Evaluation of Analysis Results (分析結果の評価)、天野恭子(京都大学)

講演 VL2: Visualization of Linguistic Layers in Vedic Literature

(ヴェーダ文献における言語層の可視化)

夏川浩明(大阪成蹊大学准教授)、天野恭子(京都大学)

4. 概要(400 字程度)

本国際ミーティングの主催者の一人である加藤隆宏(東京大学)は、国内外の大学図書館や写本資料館が所蔵するサンスクリット語資料(刊本および写本)の多くが、十分に研究に利用されていない状況を鑑み、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)によって、デーヴァナーガリー文字 OCR を開発し、サンスクリット文献群をデータベース化することを目指している。また、本研究の研究分担者である岩崎陽一(名古屋大学)は、テキストデータ構築に関わる諸問題について取り組み、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」(2022-2023FY)としても研究を進展させている。一方で、天野恭子(京都大学)は、サンスクリット文献の中で特に古い層を形成するヴェーダ文献について、その言語の分析によって、時代・地理的特徴を抽出し、当時の社会の形成・発展の議論を活性化させることを目指し、科研費国際共同強化研究 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)を展開している。このプロジェクトで試みている統計的手法を用いた研究の基礎として、天野はチューリヒ大学 Oliver Hellwig と共に、サンスクリット文献の形態素解析付きデータを作成し、Digital Corpus of Sanskrit (DCS) に蓄積している。Hellwig は、2000-2005 に Deutsche Forschungsgemeinschaft による Indoskript「インドの文字」プロジェクトの中心となって、文字の読み取りと資料化に取り組んだ経験があり、また現在は、ドイツ研究教育省プロジェクト"Chron-BMM - Bayesian Mixture Models für die Datierung von Textkorpora"「テキストコーパスの年代測定のためのベイズ混合モデル」を率いて研究を進めている。今後、日本のみならず世界で活用されるデータベースを構築するために、加藤・岩崎プロジェクトおよび天野・Hellwig プロジェクトのメンバーが共に課題を議論し、それぞれが蓄積してきた研究手法や知見を共有することが本国際ミーティングの目的であった。サンスクリット文献の分野において情報学的手法を展開する研究者はまだ数が少ないが、そのような研究の方向性に興味を持つ研究者が一同に会し、非常に有意義な議論を行うことができた。

これらのプロジェクトが最終的に目指すのは、サンスクリット文献データが広く研究に活用されることであることから、現在まだ人文情報学的研究手法に触れていないサンスクリット文献研究者とも、データ利用の可能性について議論することも本国際ミーティングの重要な目的であった。そのことから、本研究所課題公募班班(一般 A 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))も共催として参画し、当研究班からも数名のサンスクリット文献の専門研究者の参加を得た。文献学の側からも積極的な発言や提言が相次ぎ、非常に有意義な議論となった。

5. 参加者(別紙「参加状況」も記載してください。)

<学外>

狩野恭(神戸女子大学教授)

師茂樹(花園大学教授)

永崎研宣(人文情報学研究所主席研究員)

西村直子(東北大学准教授)

信賀加奈子(国際仏教学大学院大学附属図書館職員)
古宇田亮修(淑徳大学研究員)
平野克典(中村元東方研究所専任研究員)
柴崎麻穂(中村元東方研究所専任研究員)
須藤龍真(早稲田大学高等研究所講師)
塩田宝樹(大同大学非常勤講師)
榊和良(北海道武蔵女子短期大学教授)
中村史(小樽商科大学教授)
伊澤敦子(中央大学兼任講師)
川尻道哉(東海大学准教授)
加藤隆宏(東京大学准教授)
岩崎陽一(名古屋大学准教授)
大島智靖(大阪大学非常勤講師)
夏川浩明(大阪成蹊大学准教授)
宮川創(国立国語研究所助教)
塚越柚季(東京大学助教)
鄭弯弯(名古屋大学助教)
友成有紀(東京大学特任研究員)
山城貢司(東京大学研究員)
矢崎長潤(広島大学日本学術振興会特別研究員)
谷口力光(東京大学博士課程学生)
坪田さより(大阪大学博士課程学生)
竹崎隆太郎(東京大学博士課程学生)
辻春樹(東京大学博士課程学生)
Kölbl Max(大阪大学博士課程学生)
乙川文英(一般)
丸小野壮太(一般)

<学外、海外>

Yanchevskaya, Nataliya(Princeton University 教授)
Erhard, Franz Xaver(Leipzig University 教授)
Rossi Paola(University of Milan 教授)
Srinivasa Acharya(MAHE Manipal University 准教授)
DiSimone, Charles(Ghent University 准教授)
Oliver Helwig(University of Zurich 研究員)
Serbaeva Saraogi, Olga(University of Zurich 研究員)
Villamor Herrero Efrain(University of Salamanca 研究員)
Gowtham R(University of Patanjali 助教)
Acri Andrea(パリ高等研究実習院 助教)
張倩倩(福州外語外貿学院講師)
Falqui, Diletta(University of Cagliari 研究員)
京極祐希(Leipzig University 博士課程学生・プロジェクト研究員)
Weber, Elia(Freie Universität Berlin 博士課程学生)
Lin, Chia-Wei(Université de Lausanne 博士課程学生)
Kolbe, Martin(一般)
Nicholson, Brennan(一般)

<学内>

喜多千草(文学研究科教授)
宮崎泉(文学研究科教授)
Somdev Vasudev (文学研究科教授)
天野恭子(文学研究科・人文学連携研究者)
南谷奉良(文学研究科准教授)
PAN Tao(文学研究科特定准教授)
井狩彌介(文学研究科名誉教授)
芳原綾子(文学研究科非常勤講師)
藤本花音(文学研究科助教)
廣瀬勤(文学研究科博士後期課程)
岡崎智加(文学研究科博士後期課程)
山田修平(文学研究科修士課程)

6. 助成金の使途等

招へい旅費(Oliver Hellwig) 178,336 円
会議中に提供するコーヒー代 21,664 円

7. その他(成果や今後の展開等、自由に記載してください)

参加者それぞれが、発表や議論から新しい研究のアイデアを得たが、具体的に例を挙げると、今回の発表をきっかけとして天野と Hellwig が新しい文献分析の方法を考案し、その方法を用いた研究について 2024 年 12 月に Kathmandu で開催される第19回国際サンスクリット学会に発表を申し込むことを合意した。

参加状況

区分	機関数	参加人数					延べ人数				
		総計	海外研究者	40歳未満	35歳以下	大学院生	総計	海外研究者	40歳未満	35歳以下	大学院生
人文研所属 (内女性)	0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
京大内(人文研を除く) (内女性)	1	12 (5)	0 (0)	5 (2)	4 (2)	3 (1)	21 (8)	0 (0)	9 (3)	7 (3)	6 (2)
国立大学 (内女性)	6	15 (3)	0 (0)	10 (1)	10 (1)	5 (1)	30 (6)	0 (0)	20 (2)	20 (2)	10 (2)
公立大学 (内女性)	0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
私立大学 (内女性)	10	10 (3)	0 (0)	3 (0)	0 (0)	0 (0)	18 (5)	0 (0)	6 (0)	0 (0)	0 (0)
大学共同利用機関法人 (内女性)	1	1 (0)	0 (0)	1 (0)	1 (0)	0 (0)	2 (0)	0 (0)	2 (0)	2 (0)	0 (0)
独立行政法人等公的研究機関 (内女性)	0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
民間機関 (内女性)	2	3 (1)	0 (0)	0 (0)	0 (0)	0 (0)	5 (2)	0 (0)	0 (0)	0 (0)	0 (0)
外国機関 (内女性)	13	17 (6)	17 (6)	11 (4)	8 (3)	3 (1)	34 (11)	34 (11)	21 (7)	15 (5)	6 (2)
その他 (内女性)	0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
計	33	58 (18)	17 (6)	30 (7)	23 (6)	11 (3)	110 (32)	34 (11)	58 (12)	44 (10)	22 (6)
【その他の参加状況】											

※()内には、女性数を記載

※受入機関、受入人数、延べ人数を区分に応じて記入してください。

※外国人、若手研究者(40歳未満)、若手研究者(35歳以下)、大学院生の人数はそれぞれ受入人数、延べ人数に対しての内数を記入してください。

※受入人数、延べ人数については上段に総数を下段に()で女性の内数を記入してください。

※「学内」の所属機関数は「学部数」等を記入してください。

※【その他の参加状況】には「その他」区分に計上した、具体的な所属等を記載してください。無所属の場合は機関数0とカウントし、この欄の記載不要。

(例) 高校教員

※受入人数及び延べ人数の算出方法は、以下の例に基づき算出してください。

(例) 国際研究ミーティングに参加者2人が3回参加した : 受入人数2人、延べ人数6人