

## 歴史的文書に関わるデジタル図書館研究

### A. メンバー

【事業推進担当者】 前田亮

【RA】 ビルゲサイハン・バドジャルカル

【学内研究協力者】 木村文則、井坪将、大崎隆比古  
久木貴博

【その他】 ハルタルフー・ガルマーバザル、手塚太郎

### B. 研究目的

近年、デジタル図書館やデジタルアーカイブが注目され、さまざまな文化的資料のデジタル化や保存に関する研究が盛んに行われている。しかしながら、それらのコンテンツに対して容易で効率的なアクセス手段を提供するという観点からの研究はまだ多くはない。コンテンツの量が膨大になればなるほど高度なアクセス手段が要求されることは、現在の Web の状況を見ても明らかである。

本研究プロジェクトでは、日本文化に関わる多言語からなる歴史的な文書のデジタル図書館システムを構築し、これらのコンテンツに対して容易で効率的なアクセスを実現する各種技術について研究を行う。

さらに、テキストだけでなく、イメージや映像などの各種メディアから構成される複数のデジタル図書館・アーカイブ・ミュージアムに対して効率的な情報アクセスを実現するためのデータベース横断検索の技術について研究を行う。最終的には、言語・文化・時代の壁を越える横断検索技術の確立を目指す。

### C. 本年度の成果

#### 1) 日本語古典史料のテキスト処理手法の開発

当プロジェクトでこれまで行ってきた日本語古典史料の現代語による検索、テキスト分析および可視化の研究を発展させ、日本語古典史料に対するテキスト処理手法の確立を目指して研究を進めた。

前年度より研究を行っている、文字 N グラムの出現確率を用いた日本語古典史料テキストからの単語抽出の手法について、手法の実装および実際の文書を用いた評価実験を行い、特にテキスト分析において重要な名詞の抽出において一定の精度が得られた。

また、以前から研究を進めている、『兵範記』『玉葉』『吾妻鏡』における注目人物と特定二者との関係の推移の可視化の研究および『兵範記』における地名との共起情報を用いた人物関係の可視化の研究について、Web ブラウザから利用可能なユーザインタフェースを構築し、人文系研究者が容易に分析を行える環境を整えた。さらに、このユーザインタフェースを京都文化研究班の杉橋研究室のメンバーに使用してもらい、提案手法の利用者評価を行った。その結果、可視化結果の信頼性、提案手法の新規性および有用性の観点において概ね肯定的な評価が得られた。

本研究に関しては、著書（分担執筆）2 件、国際会議 2 件、国内会議 2 件の発表を行った。

#### 2) 人文系データベースの横断検索システム

本学アート・リサーチセンターをはじめ、国内外で多数公開されている人文系データベースの相

互利用を図るため、これらのデータベースを横断的に検索するための技術について研究を行った。

具体的には、国内外で公開されている複数の浮世絵画像のデータベースを対象とし、複数のデータベースを同時に検索し、検索結果をその場で統合して表示するプロトタイプシステムを構築した。現時点で対象としているデータベースは、大英博物館、ヴィクトリア・アンド・アルバート博物館、ボストン美術館、アメリカ議会図書館、国立国会図書館、立命館大学アート・リサーチセンターの6データベースであり、合計10万点以上の画像が対象となっている。これらはデータベースによってメタデータが英語もしくは日本語のいずれかで記述されているが、これらに対して辞書や翻字ツールなどの言語資源を用いて人名や専門用語を含むメタデータの翻訳を行うことで、海外のデータベースの日本語による検索、あるいは国内のデータベースの英語による検索を実現した。また、本システムを海外の日本文化研究者に実際に使用してもらい、利用者評価を行った。

本研究に関しては、著書(分担執筆)2件、国際会議3件、国内会議1件の発表を行った。また、2012年7月に開催される国際会議 Digital Humanities 2012 でポスター発表を行う予定である。

## E. 業績一覧

### 〈著書(分担執筆)〉

前田亮, 木村文則, Batjargal Biligsaikhan 「デジタル図書館・アーカイブへの言語・時代・文化横断型の情報アクセス」八村広三郎, 田中弘美編『デジタルアーカイブの新展開』ナカニシヤ出版, pp.150-168, 2012年3月, Akira Maeda, Fuminori Kimura, and Biligsaikhan Batjargal, 'Cross-Lingual, Cross-Chronological, and Cross-Cultural Information Access to Digital Libraries and Archives', Kozaburo Hachimura, and Hiromi T. Tanaka eds., "New Developments on Digital Archives", Nakanishiya Shuppan, pp.324-342, 30 March 2012

### 3) 伝統的モンゴル文字文書のデジタル図書館システム

従来から研究を進めている伝統的モンゴル文字文書のデジタル図書館システムについて、昨年度に公開したシステムの評価実験を中心に行った。まず、現代モンゴル語による119件の問合せを用いた検索精度の評価実験の結果、伝統的モンゴル語文書に対する現代モンゴル語による検索において、実用上十分な精度が得られることを実証した。また、モンゴル語に関する専門家および一般利用者20名のモンゴル人による利用者評価を行った結果、専門家および一般利用者の双方から、システムの有用性および操作性に関して高い評価が得られた。

本研究に関しては、著書(分担執筆)1件の発表を行った。また、本研究に関する論文を学術論文誌に投稿中である。

## D. 論文・学会発表以外の活動の記録

特記事項なし

Biligsaikhan Batjargal, Garmaabazar Khaltarkhuu, Fuminori Kimura, and Akira Maeda, 'Integrated Information Access Technology for Digital Libraries: Access across Languages, Periods, and Cultures', Kuo Hung Huang ed., "*Digital Libraries- Methods and Applications*", 2, pp.23-44, April 2011

#### 〈論文〉

【審査付き】 Fuminori Kimura, Mamoru Yoshimura, and Akira Maeda, 'Term Extraction from Japanese Ancient Writings Using Probability of Character N-grams', *Proceedings of The 2nd International Conference on Culture and Computing (Culture and Computing 2011)*, pp.183-184, Kyoto University (Kyoto, Japan), 22 October 2011

【審査付き】 Biligsaikhan Batjargal, Fuminori Kimura and Akira Maeda, 'Realizing Bilingual and Parallel Access to Ukiyo-e Databases in the World', *Proceedings of The 2nd International Conference on Culture and Computing (Culture and Computing 2011)*, pp.165-166, Kyoto University (Kyoto, Japan), 22 October 2011

【審査付き】 Biligsaikhan Batjargal, Fuminori Kimura, and Akira Maeda, 'Metadata-related Challenges for Realizing Federated Searching System for Japanese Humanities Databases', *The 11th International Conference on Dublin Core and Metadata Applications (DC-2011)*, 6273/2010, pp.80-85, The National Library of the Netherlands (The Hague, Netherlands), 22 September 2011

【審査付き】 Sho Itsubo, Takahiko Osaki, Fuminori Kimura, Taro Tezuka and Akira Maeda, 'Visualization of Co-occurrence Relationships Using the Historical Persons and Locational Names from Historical Documents', *Conference Abstracts of Digital Humanities 2011*, pp.326-329, Stanford University (Stanford, CA), 20 June 2011.

【審査付き】 井坪将, 木村文則, 前田亮「古典史料からの相対的な人物関係の時間的変化の推定と可視化」人文科学とコンピュータシンポジウム論文集, pp.29-36, 2011年12月10日

【審査付き】 吉村衛, 木村文則, 前田亮「古文テキスト解析のための文字Nグラムの出現確率を利用した単語分割」人文科学とコンピュータシンポジウム論文集, 2011-8, pp.261-268, 2011年12月11日

【審査付き】 久山岳夫, Biligsaikhan Batjargal, 木村文則, 前田亮「浮世絵を対象とした異種データベースの多言語統合アクセス手法の提案」人文科学とコンピュータシンポジウム論文集, 2011-8, pp.275-280, 2011年12月11日

#### 〈口頭発表〉

【審査付き】 Biligsaikhan Batjargal, Fuminori Kimura and Akira Maeda, 'Accessing Multiple Japanese Humanities Databases Using English Queries', *Osaka Symposium on Digital Humanities 2011*, Osaka University (Osaka, Japan), 13 September 2011

《講演》

Akira Maeda, 'Integrated Information Access and Analysis of Japanese Humanities Databases',  
*Seminar talk for New Directions in Digital Humanities*, Department of Digital Humanities, King's  
College London (London, U.K.), 15 December 2011