

立命館大学アート・リサーチセンター  
 文部科学省 国際共同利用・共同研究拠点  
 「日本文化資源デジタル・アーカイブ国際研究拠点」  
 2020年度 国際共同研究成果報告書〔研究費配分型〕

2021年 5月 7日 提出

1. 研究課題名	
芳賀矢一『日本人名辞典』からの歴史人物情報の自動抽出および情報構造化 (英文課題名) Automatic Extracting Personal Information of Historical Persons from "Nihon Jinmei Jiten" and Creating Structured Data	
2. 研究代表者	
氏名 (ふりがな)	所属機関・職名
Fuminori Kimura	Onomichi City University
3. 研究分担者 (合計: 2名)	
氏名 (ふりがな)	所属機関・職名
Akira Maeda	College of Information Science and Engineering, Ritsumeikan University・Professor
Makoto Goto	National Museum of Japanese History・Associate Professor

4. 研究課題の概要 (300字程度) (申請書から変更がある場合は、変更点分かるように明記してください)
<p>Person information of historical persons are one of the most fundamental sources of information for the study of Japanese history. At present, the National Diet Library provides authentic biographical information, but it is limited to books and other sources, and does not cover a wide range of historical persons. Therefore, it is possible to provide basic information on a wide range of historical persons by creating biographical data using the "Dictionary of the Japanese Name" edited by Yaichi Haga (hereinafter referred to as the "Haga Biographical Dictionary"), which contains the names of about 5,000 persons and whose copyright has expired. Currently, this work is done manually by experts, but in this research, we aim to support this work by realizing the automation of extraction.</p> <p>In addition, since the most important condition for the information infrastructure is that the data be machine-readable, the information is recorded as Linked Data and stored in a database. In this way, the meaning of the stored information can be expressed and the information can be easily linked to information in other databases, which is expected to promote research on Japanese history and other subjects.</p>
5. 研究成果の概要 (この項は、本センターのホームページ・紀要等で公開することがあります)
<p>In this research, we extracted person information of historical persons by analyzing for the sentences about historical persons in "Haga Biographical Dictionary" written by Yaichi Haga. We used word2vec, a kind of neural network, to detect words that correspond to the attributes to be extracted and obtained the words as the person information. For several attributes ("title," "year," "geography," and "alias"), we achieved high extraction accuracy (precision 0.76-1.00).</p> <p>Although there is still room for improvement, the proposed method can be used to assist experts in the manual extraction of historical persons information.</p>

6. 研究業績 (日本語以外に英語名称もあるものは、できるだけ日英両言語でご記入ください)

- (1) 著書
- (2) 論文
- (3) 研究発表等
- (4) 主催したシンポジウム・研究会等
- (5) その他研究活動 (報道発表や講演会等)
- (6) 受賞学術賞
- (7) 科学研究費助成事業
- (8) 競争的資金等 (科研費を除く)
- (9) その他