

立命館大学アート・リサーチセンター

文部科学省 共同利用・共同研究拠点「日本文化資源デジタル・アーカイブ研究拠点」
2017年度 共同研究成果報告書〔研究費配分型〕

2018年 4月 21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抽出および情報構造化 (英文標記: 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) きむら ふみのり	所属機関・職名
木村 文則	尾道市立大学 経済情報学部・講師
3. 研究分担者 (合計: 3名) ※アート・リサーチセンター所属者は、「ARC 所属教員欄」に○印を付してください	
氏名(ふりがな)	所属機関・職名
前田 亮(まえだ あきら)	立命館大学 情報理工学部・教授
後藤 真(ごとう まこと)	国立歴史民俗博物館・准教授
SONG Yuting	立命館大学 情報理工学研究科・D2

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>日本の歴史研究を行うにあたり、最も基盤となる情報の一つとして人名があげられる。現在国立国会図書館が人名典拠情報を提供しているが、書籍など限られた人名になっており、歴史上の人物を広く網羅できているわけではない。そこで、約 50000 人の人名を収録し、かつ著作権が切れている芳賀矢一編『日本人名辞典』を活用した人名データを作成することで、広く歴史的人名の基礎状況を提供することができる。</p> <p>このデータは、機械可読であることが情報基盤の最大の条件であるので、Linked Data として情報を記録し、データベースへと格納する。これにより、格納された情報の意味を表現できるとともに、他のデータベースの情報とのリンクが容易となり、日本史をはじめとする諸研究の促進が期待される。</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>芳賀矢一編『日本人名辞典』に記述されている人物の解説文を解析することにより、その人物の人物情報を取得した。本研究においては、抽出すべき属性ごとにパターンマッチングを行い、該当パターンとなる文字列を抽出することで、その人物の情報を取得した。取得した情報の種類は、「別名」「親」「地理情報」「仕えた人」「時代」「死没年月日」などである。本手法の抽出精度を検証するため、専門家が作成した正解データ(1,257 人分)と比較し、正解率を測定した。今回は完全一致した場合のみを正解とし、測定した結果、属性により4割から7割程度の再現率が得られた。</p> <p>本手法は改善の余地が残ってはいるが、専門家が人手により人物情報の抽出作業を行う際に、提案手法により支援を行うことができるだけの成果が得られた。</p>

6. 研究業績

(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