

科学研究費助成事業（基盤研究（S））公表用資料
〔令和2（2020）年度 中間評価用〕

平成30年度採択分
令和2年3月31日現在

木簡等の研究資源オープンデータ化を通じた
参加誘発型研究スキーム確立による知の展開

Development of Integrated Knowledge through Establishment of
an Interactive Research Scheme based on the Open-Data of
Research Resources for Wooden Tablets and Related Topics

課題番号：18H05221

馬場 基（BABA, HAJIME）

独立行政法人国立文化財機構奈良文化財研究所・都城発掘調査部・史料研究室長



研究の概要（4行以内）

本研究は、A) IIF 準拠による歴史的な文字情報化の標準仕様を提案・奈文研所有の木簡研究資源をオープンデータ化・研究資源共有の国際的潮流を惹起、B) 参加誘発型スキームの確立を通じた国・分野を超えた多様な研究資源の蓄積を加速、C) 質量ともに拡充した研究資源にビッグデータ解析手法等様々な分析を施し、従来の枠組みを超越した研究の展開を目指す。

研究分野：史料研究、人文情報学、史学、日本語学、美術史

キーワード：木簡、オープンデータ化、文字文化、日本史、IIF、交流史

1. 研究開始当初の背景

木簡は、可能性に満ちた資料である。

近年では、歴史学のみならず、国語学・国文学等多様な研究分野で重要視されている。研究手法も多様化・精緻化し、共同研究も深化してきている。この中で、日本の文字文化の世界史的特性も明らかになりつつある。

この研究状況の背景には、データベースの拡充など木簡研究資源の蓄積がある。研究の加速・新展開には、研究資源の「量の拡大」「質の多様化」が必須である。だが、単独の研究機関の提供力には限界がある。

一方、世界的には「IIF」（相互運用性に向けた資源化ルール）に準拠した研究資源化と、そのオープンデータ化が潮流となりつつあった。IIF に準拠しつつ、必要に応じた検索機能や、共通検索を可能にする、より目的に特化したルール・手法があれば、さらに研究の展開に大きな意義を持つと期待された。

2. 研究の目的

木簡研究資源を質的・量的に拡充し、研究の新段階を切り開くことを最終目標とする。
○連携や共有による研究資源の充実

対等で開かれた研究資源連携・研究資源のオープンデータ化・参加誘発型スキームの確立を実現して、研究資源の充実、木簡に関する知の解放・共有を加速し、東アジアや世界での歴史的な文字の研究をリードする。

○多分野協力による新研究の展開

拡充した研究資源を、情報学的手法を導入しつつ広く活用し、新研究を展開する。特に、日本列島での文字文化獲得過程の具体的な様相の解明にむけた、多様な研究を展開する。

3. 研究の方法

本研究は、以下の A) ～C) の実現によって、目的の達成を期する。

A) 木簡研究資源の「量の拡大」

a 歴史的な文字の研究資源化のための IIF 準拠標準規格の策定

b 奈文研所有の既存研究資源の IIF 化によるオープンデータ化

c IIF 化を通じた国内外機関の参加・共同による連携体制の構築

B) 木簡研究資源の「質の多様化」

a 参加誘発型スキームに向けたツール開発

b 資源化過程の「見える化」による経験知の資源化

c 歴史的な文字に関する注釈データの充実

C) 文字資料研究を牽引する新研究の展開

a ビッグデータ解析手法等の分析手法を導入しての検討

b 分野・国を超越した研究の深化・公表

4. これまでの成果

順調に進捗し、成果を上げている。特に連携検索システム公開・国際学会開催などは、当初計画よりも早期に実現できた。

A) 木簡研究資源の「量の拡大」

・「IIIF に基づく歴史的な文字研究資源情報と公開の指針」の宣言と「オープンデータに関する仕様」の策定

標記宣言及び仕様を、連携機関（奈良文化財研究所・東京大学史料編纂所・国文学研究資料館・国立国語研究所・京都大学人文科学研究所・台湾中央研究院歴史語言研究所）で策定、世界に先駆けて公表した（R2 年 3 月）。

・「史的な文字データベース連携検索システム」（実証試験版）の公開

上記指針による、世界初となる IIIF 準拠の本格的な連携検索システムの実証試験版（日本語・英語・繁体中国語・簡体中国語・韓国語）を公開し（R2 年 3 月）、奈文研・東大史料編纂所・国文研の連携検索を実現した。

・木簡研究資源のオープンデータ化

「木簡庫」DB にダウンロード機能を付与し、テキストデータをオープンデータ化した（H30 年度）。また、上記連携検索システムによって、IIIF 化した木簡画像データのオープンデータ化を実現した。

・データの拡充

既存の「木簡庫」データベースの木簡文字画像（約 10 万文字）を IIIF 化し（R1 年度）、新規に約 1 万 5 千文字を研究資源化した。現在、合計 11 万 5 千文字（5 カ年の目標 20 万文字中）を研究資源化。

B) 木簡研究資源の「質の多様化」

・木簡文字の観察記録シートの作成

木簡文字観察記録シートを約 5 万件作成した（5 カ年の目標 14 万件中）。対象を東アジア各地の木簡に広げて作業を行い、今後の研究進展に向けた基礎データを拡充した。

C) 文字資料研究を牽引する新研究の展開

・学会・研究会の実施

国際学会を 2 回共催（台湾・中国）し、IIIF 連携の共同研究や、木簡の国際共同研究および成果報告を行った。国内で開催された学会（国際・国内とも）でも報告した他、研究チーム内での研究会を開催した。

5. 今後の計画

1 連携検索システムの強化 (A)

連携検索システムに、文字画像検索システム「MOJIZO」を実装（R2 年度）。また、ポータルサイト設計仕様書・API 等を公開（日本語・英語）（R2 年度）し、「開かれた」環境を提供。あわせて「木簡庫」DB も、「全国遺跡報告書総覧」DB 等とのリンクによって情報集約力を強化する。

2 参加誘発環境の開発と提供 (B)

気づき書き込みシステムおよび SNS データ蓄積システムの開発（（R2・3 年度）し、多様な知の収集を目指す。

3 研究資源化の加速 (B)

東大資料編纂所のシステムを改良・導入し、木簡整理作業デジタル化を推進（R2・3 年度）。また、木簡文字の情報資源化・観察記録シート作成を継続的に実施。

4 深層学習を利用した研究の展開 (C)

文字画像検索システム「MOJIZO」に深層学習を導入（R3・4 年度）。この他、木目除去による文字を鮮明化（R2・3 年）、オフライン運筆情報復元（R2・4 年度）の研究を実施。

5 東アジア漢字文化に関する共同研究 (C)

日中韓簡牘総合研究集会の共催・参加等により、台湾・中研院に加えて中国社会科学院との共同研究を実施（全期間）。また、韓国国立文化財研究所等韓国との共同研究に向けて調整する。質量ともに強化した研究資源を利用して、「ローカルな漢字文化」に迫る観点からの研究を推進する。

6. これまでの発表論文等（受賞等も含む）

1、馬場基「木簡学から見た日本語—文字について—」、『日本語学』11 月号、査読無、pp.16-29、2019 年

2、山田太造「オープンな歴史的な文字データを横断的に検索していく」、『第 31 回東洋学へのコンピュータ予稿集』31、査読有、pp.119-133、2019 年

3、笹原宏之「Yicún Wénzì(extant characters) in Japanese lexicon: Exploring characters of historically Chinese origin with evidence from Six dynasties, Sui and Tang dynasties, China(創于六朝、隋唐時期存迹于日本的“佚存文字”）」、『中国文字』2019 年第三期、査読有、p153-173、2019 年

4、末代誠仁・リー トゥアン ナム・グエン コング カー・中川正樹・山本和明「階層化された情報システムのためのくずし字解読機能の試作」、『日本情報考古学会講演論文集』Vol. 22(2019)、査読有、pp.11-15、2019 年

5、馬場基「奈良文化財研究所の ICT への取り組み」、『日本歴史』848、査読無、pp.9-13、2019 年

7. ホームページ等

<https://mojiportal.nabunken.go.jp/ja/>

<https://mokkanko.nabunken.go.jp/ja/>