

自己評価報告書

平成23年 3月31日現在

機関番号：84604

研究種目：基盤研究（S）

研究期間：2008～2012

課題番号：20222002

研究課題名（和文）木簡など出土文字資料積読支援システムの高次化と総合的研究拠点データベースの構築

研究課題名（英文）The improvement of the support system to decode excavated historical documents and the development of the comprehensive research center database

研究代表者

渡辺 晃宏（WATANABE AKIHIRO）

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

研究者番号：30212319

研究分野：人文学

科研費の分科・細目：史学・日本史

キーワード：古代史、木簡、出土文字資料、データベース、漢字、文字認識、OCR

1. 研究計画の概要

推論機能を有する木簡の文字認識システム「Mokkanshop」と木簡の文字画像データベース「木簡字典」を中核とし、多様な知識データベースを有機的に結合させた出土文字資料研究拠点データベースを構築する。そして、木簡情報の知の蓄積が解読を可能にし、それが新しい知としてさらなる解読を生み出すという、「知のスパイラル(循環)」を実現する。

2. 研究の進捗状況

研究は、木簡など出土文字資料積読支援システムの高次化、木簡など出土文字資料総合的研究拠点データベースの構築、の二つの大きな柱からなる。これまでに、については下記(1)～(2)を、については下記(3)～(8)を実現した。

(1)Mokkanshop の高次化 使用するデータ（字書）の改善の他、欠損部分の情報を利用者が付加するシステムを開発し、認識率が向上した。これを受け、2010年10月にMokkanshop 新版を公開した。また、Mokkanshop から文字画像データベース「木簡字典」の参照を可能にし、研究拠点データベースとしての連携を整えた。さらに、木簡画像の視認性を向上させるために、墨痕と背景を分離する画像処理技術の研究を行った。墨書土器・漆紙文書など木簡以外の出土文字資料の文字認識や、正倉院文書や奈良朝写経の文字の字書として活用も検討中である。

(2)積読支援システムの同時運用方式の開発 Microsoft SharedView を用い、東京大学史料編纂所との間で赤外線観察装置の画像共有実験を行った。若干の時間差が発生するものの、問題なく使用できることがわかった

が、音声共有については課題を残した。

(3)アノテーションツールの開発 削屑の効率的整理・観察用に、削屑のデジタル画像に管理情報や積読成果を書き込めるアノテーションツールを開発し実用の目途を得た。

(4)文化財撮影用レンズの開発 赤外線と可視光の撮影が同一工程内でフィルタの切り替えだけでできるよう特殊仕様レンズを開発し、撮影の大幅効率化を実現した。

(5)文字画像データの蓄積とXMLの導入 木簡の文字画像データベース「木簡字典」の充実のため、文字画像切り出しシステム Mojiga を改良し、時代・地位を越えたデータの蓄積に努めた。2010年度までに約28,000点の画像を蓄積し、累積数は約48,000点、木簡点数で約3,000点に達した。累計文字種も約1,500種で、木簡に登場する文字をほぼ網羅した。XMLの導入で意味による検索や他の情報とのリンクが可能になった。

(6)共通入力ツールの開発 「木簡字典」のデータ充実の効率化のため、木簡データベースと木簡字典の共通入力ツールを開発し、木簡字典のメタデータ付与方法を大幅に改善した。

(7)東京大学史料編纂所「くずし字字典データベース」との連携検索の実現 「木簡字典」と東京大学史料編纂所の「くずし字字典データベース」との共通検索システムを開発・公開し、機関相互の画期的な連携が実現した。

(8)出土遺構年代観データベースと木簡人名データベース 出土遺構年代観データベースと、東京大学史料編纂所教授田島公氏を代表者とする学術創成研究の分担者として「木簡人名データベース」を作成した。後者には出土遺構年代観データベースの出土地

点情報を盛り込み木簡字典とリンクさせた。

3. 現在までの達成度

おおむね順調に進展している。

本研究の究極目標は、木簡字典と Mokkanshop を中核とした出土文字資料研究拠点データベースを構築し、私たちが半世紀にわたって培ってきた木簡の整理・解読・保管のノウハウを形に残し、社会に還元するとともに、それを私たち自身の研究工具として活用し、新しい知見を再びノウハウに追加していく、いわば「知のスパイラル」を構築し軌道に乗せることである。上記の研究成果はほぼ当初の予定通りのものであり、研究は順調に進捗していると判断する。

4. 今後の研究の推進方策

2011 年度に、木簡字典フラッシュ版の公開の実現、墨書土器字典データベースの開発、遠隔地システムの本格的な運用の開始を、2012 年度に、木簡字典に外国語検索ページを用意し、海外との関係に備える、知識データベース群の充実、研究拠点データベースの本格的な運用、印刷版『日本古代木簡字典』の買い手などを実現し、「知のスパイラル」を確立したい。

また、今後、木簡字典・Mokkanshop を介した利用者との接点として、奈良文化財研究所のホームページで公開中の「木簡ひろば」の機能充実を図っていきたい。

5. 代表的な研究成果

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 30 件)

渡辺晃宏、今後の平城京研究と木簡研究、上代文学、査読有、第 105 号、p.p.22-40、2010 年

井上聡・馬場基、文字字形総合データベース作成の試み 電子くずし字典データベースと木簡画像データベース・木簡字典の連携について、『人間文化研究情報資源共有化研究会報告集』、査読無、1 号、p.p.99-112、2010 年

Jun TAKAKURA, Akihito KITADAI, Masaki NAKAGAWA, Hajime BABA and Akihiro WATANABE, Techniques to Enhance Images for Mokkan Interpretation (口頭発表あり) Proc. 12th International Conference on Frontiers in Handwriting Recognition, 査読有、Vol.1, No.1, p.p.358-362、2010 年

渡辺晃宏、日本古代の習書木簡と下級官人の漢字教育、高田時雄編『漢字文化三千年』(臨川書店刊) p.p.91-112、2009 年

末代誠仁・齋藤恵・戸根康隆・石川正敏・中川正樹・馬場基・渡辺晃宏、古代木簡解

読支援のための文字パターン検索、情報処理学会論文誌、査読有、Vol.50 (4)。p.p.1444-1455、2009 年

〔学会発表〕(計 20 件)

Sherini Somayeh、古代木簡解読支援システムにおける字体検索の高性能化、情報処理学会人文科学とコンピュータシンポジウム 2010、2010 年 12 月 11 日、東京工業大学大岡山キャンパス(東京)

末代誠仁、木簡解読支援のための情報処理技術、木簡学会第 32 回研究集会、2010 年 12 月 4 日、奈良文化財研究所平城宮資料館講堂(奈良)

Jun TAKAKURA、Techniques to Enhance Images for Mokkan Interpretation、12th International Conference on Frontiers in Handwriting Recognition、2010 年 11 月 16 日、Kolkata, India

渡辺晃宏、日本古代の都城木簡と羅州木簡、韓国国立羅州文化財研究所開所 5 周年記念国際学術大会「6~7 世紀栄山江流域と百済」、2010 年 10 月 28 日、羅州、韓国

末代誠仁、奈良木簡解読支援のためのパターン認識、電子情報通信学会 PRMU 研究会技術報告[特別講演ならびにパネラーとして]、2009 年 2 月 19 日、東京大学(東京)

〔図書〕(計 2 件)

渡辺晃宏、平城京 1300 年全検証 奈良の都を木簡から読み解く、総頁 377 頁、柏書房刊、2010 年

馬場基、平城京に暮らす、総頁 237 頁、吉川弘文館刊

〔産業財産権〕

取得状況(計 1 件)

名称: Mokkanshop 第 9 類 木簡に記された文字を判読するためのコンピュータソフトウェア、インターネットを介してダウンロードされるコンピュータソフトウェア

発明者:

権利者: 独立行政法人国立文化財機構

種類: 商標登録

番号: 商標登録第 5307067 号

取得年月日: 2010 年 3 月 5 日

国内外の別: 国内

〔その他〕

<http://hiroba.nabunken.jp/>: 木簡ひろば

<http://jiten.nabunken.jp/>: 木簡字典

<http://r-jiten.nabunken.jp/>: 連携検索

<http://www.tuat.ac.jp/~koukai/gakuho/2010/504/topics2.html>: 東京農工大の HP