

平成 27 年 6 月 5 日現在

機関番号：12601

研究種目：基盤研究(A)

研究期間：2011～2014

課題番号：23240031

研究課題名(和文) ボーンデジタル画像管理システムの確立に基づく歴史史料情報の高度化と構造転換の研究

研究課題名(英文) Research concerning the enhancement and structural changes of historical materials data with the establishment of born-digital management system

研究代表者

山家 浩樹 (YAMBE, Koki)

東京大学・史料編纂所・教授

研究者番号：60191467

交付決定額(研究期間全体)：(直接経費) 33,900,000円

研究成果の概要(和文)：「ボーンデジタル進捗状況管理システム」を構築して、無秩序に生成されがちなデジタル撮影画像(ボーンデジタル)を、組織として一貫して管理・運用するシステムを確立し、歴史史料のデジタル画像を共有する基盤を整えた。さらに、標準化された仕様に適合しないデジタル画像を、メタデータとともに管理する一例として、ガラス乾板など古写真を取り上げ、「ガラス乾板情報管理ツール」を開発して、ガラス乾板の研究資源化および保存にむけた研究を行なった。あわせて、具体例をもとに、デジタル画像を主たるレコードとするデータベースの構造転換に向けた研究を推進した。

研究成果の概要(英文)：In this study, we built a system to manage and operate digital photographed images of Japanese historical materials consistently as an organization. In addition, as an example to manage the digital image which did not adapt to standardized specifications, we took up glass photographic plate, and developed a tool for management of digital images and metadata of them. With this tool, we researched the preservation of them. And based on the concrete example, we promoted a study for the structural conversion of the database which assumed a digital image a main record.

研究分野：日本史 歴史情報学

キーワード：日本史 図書館情報学 デジタル画像管理 古写真 ガラス乾板 くずし字

1. 研究開始当初の背景

(1) フィルムからデジタルへという写真技術の激変のもと、デジタルデータでは、画像形式や規格が一定でないため、多様な画像データがそのまま無秩序に蓄積される。この問題は、組織的調査を担っている歴史系研究機関・史料保存機関で急速に顕在化しつつあり、研究資源として組織的に蓄積されるべき画像データが、雲散霧消する危機を迎えていた。そこで、本研究においては、デジタル撮影画像(ポーンデジタル)の生成から管理・運用・維持を一貫して担保する安定的なシステムの確立を目指した。

(2) 史料編纂所のみならず歴史系研究機関において構築・公開してきたデータベースは、史料1点単位でレコードを立て、各種のメタ情報やフルテキスト、デジタル画像などを付加する構造を持っている。しかしデジタル撮影画像が飛躍的に増大するなか、撮影画像を主たるレコードとする構造に変換し、画像ファイルにメタデータを付加するほうが利便性は高い。そこで、デジタル画像ファイルを主レコードとするデータベース構造の転換・修正を課題とした。

(3) 史料編纂所においては、古文書や古記録といった一般的な形態の史料群については、デジタル様式の標準を確立しつつある。しかし、それから逸脱する物理的に取り扱うことの難しい史料群については、仕様の統一的運用がなされていない。また各種研究や業務のなかで標準仕様から逸脱して生成されてくるデータ群があり、研究や業務を裏支えるデータとして、維持・活用を目指してゆく必要がある。そこで、標準的なデータ仕様から逸脱するデジタル画像の生成・管理・運用スキームの確立を課題とした。

2. 研究の目的

(1) ポーンデジタル画像を歴史学界共有の研究資源として有効に活用し、研究成果を社会に還元するために、ポーンデジタル画像管理システムを、史料編纂所歴史情報処理システムの上に構築し、組織的な運用実践を試みる。

(2) デジタル画像の劇的な増加に対応したデータベース構造の転換を念頭に、デジタル画像データを主コードにすえ、それぞれに史料学的なメタ情報を付して分類や編成を行うことを、具体的なデータベースをもとに研究する。

(3) デジタル画像生成の標準的スキームが確立していない多様な史料群について、データ生成・管理・運用に即した仕様を検討し、そのうえにデータの管理・運用のありかたを検証する。

3. 研究の方法

(1) ポーンデジタル画像の管理システム研究では、具体事例として歴史史料調査とデジタル撮影を重ねながら、システムの安定的な運用に必要な条件を分析し、その経験の上にシステム開発を実現する。開発ののち、調査・撮影で検証を続け、更新を行う。また、こうした経験を学界に公開・発信することで、標準化にむけた取り組みを進める。

(2) デジタル画像を主たるレコードとするデータベースの構造転換研究では、既存データベースの基本構造をすべて転換することは人的・費用的に困難と言わざるを得ないため、具体的なデータベースを対象に、デジタル画像を主コードとなし、画像データに各種メタデータを付与し、蓄積する機能を研究する。そのうえで、画像コンテンツを格納したストレージと既存データベースを連携する方法論を探る。

(3) 標準的なデータ仕様に適合しないデジタル画像の生成・管理・運用スキームの研究では、

特殊形状史料のデジタル化仕様の策定と実践：研究の準備がもっとも整っている、ガラス乾板及び古写真を主たる対象とし、史料編纂所所蔵のガラス乾板およびその焼付などからなる台紙付写真を対象に、デジタル化仕様を検討し、デジタル画像とメタデータとの結合を図って、その標準化を目指す。

研究上生成される二次的デジタル画像への対応：くずし字画像、すなわち字形・字体を分析するため文字や語量単位でトリミングした画像を主たる対象とし、他の研究グループと連動しながら、史料実体に応じたモデルを提示する。

業務上生成されるデジタル画像への対応：史料修復や保管、紙質の分析などに伴って生成されるデータがどのようなバリエーションで発生しているかを把握し、管理システムの一環に位置づけられるよう仕様研究を進める。

4. 研究成果

本科研の構成要素のうち、(1)ポーンデジタル画像管理システム研究 および(3)のうちガラス乾板や古写真を対象とした研究において、とりわけ大きな成果が得られた。以下、担当者は特記すべき場合のみ掲げる。事務局は、おもに連携研究者井上聡・川本慎自が担った。

(1) ポーンデジタル画像の管理システム研究

歴史史料調査とデジタル写真撮影の実例を踏まえながら、データの生成から管理にいたる一連の流れを精密化し、そのうえに「ポーンデジタル進捗状況管理システム」を構築した。史料編纂所で生成されるデジタル撮影

画像群につき、撮影仕様から統一したうえで、「採訪コード」という特定指標を基礎に一貫して把握することを可能とし、生成前からサーバへの格納、さらには公開までの過程を一元的に管理するシステムとして開発した。このシステムにより、教員・技術職員・図書職員といった多様な立場の人間が複雑に関わりながら生成されていくデジタル画像を、組織として明確に管理することが可能となった。さらに、史料調査・写真撮影を継続してこのシステムを運用し、ポーンデジタル画像管理システムとしての機能や操作性を検証して、可能な範囲でシステム改良を行うとともに、改善点を整理した。

デジタル画像を管理するシステムについては、公開について学界からの強い要請があり、学会発表などで紹介に努めてきたが、具体例として提示することとし、マニュアル的な詳しい紹介を東京大学史料編纂所研究報告 2014 - 8 『組織によるデジタル撮影画像の生成・管理』にまとめ、公開して共有化した。



ポーンデジタル進捗状況管理システム画面

(2) デジタル画像を主たるレコードとするデータベースの構造転換研究

具体的なデータベースとして、「翻刻支援システム」を取り上げた。同システムは、デジタル史料画像をディレクトリから指定してクライアント PC に呼び出し、画像データに翻刻テキストほかのアノテーションを付与する形をとっており、付与するメタデータの深化、とくに既存目録データとの連関が課題となっていた。研究分担者山田太造を中心に、他の研究グループと連動しながら、画像表示環境の改善を手始めに、入力環境や出力環境の機能向上を行い、また画像ディレクトリごとの細目録と既存の書誌目録データとの相互連関にむけた検証を継続して、より合理的なメタ体系の確立を試行した。同時に機能検証をかねて、データ蓄積も進めた。

(3) 標準的なデータ仕様に適合しないデジタル画像の生成・管理・運用スキームの研究 特殊形状史料のデジタル化仕様の策定

と実践では、古写真を素材とした研究に注力した。史料編纂所の所蔵するガラス乾板を対象に、撮影仕様を検討したうえで、状態調査と同時にデジタル撮影を実施し、研究期間を通じて 1540 点の調書作成とデジタル画像化を実現した（全体の 7 分 1 ほどにあたる）。

これらデータは、「ガラス乾板情報管理ツール」の開発を行って搭載した。これは、ガラス乾板のデジタル撮影画像と、史料学的情報、保存科学的情報を一体化して管理するツールであり、研究期間中に改良を重ねた。このツールに一元的に集約された多様な情報は、ガラス乾板を研究素材として有効活用し、かつ保存対策を講ずる上で貴重な情報の宝庫となる。調査の完了したガラス乾板は、状態に応じた保存方法を検討して、オリジナルな専用保存箱を作成した。

2012 年度に研究会「東京大学史料編纂所所蔵ガラス乾板を中心とする写真史料の調査・整理保存に関する研究」、2014 年度に研究集会「ガラス乾板の調査・保存・研究資源化に関する研究」・研究会「ガラス乾板の後加工技術について 膜面返しを中心に」（いずれも画像史料解析センターのプロジェクトと共催）を開催し、アメリカ・デラウェア大学文化財保存修復学科長デボラ・ヘス・ノリス教授をはじめとする、第一線の研究者との意見交換を積極的に行った。研究集会以降、希望者に開発したツールを配付し、成果を公開し共有するとともに、さらなる向上を目指した。本科研による取り組みは高い関心を呼び、これを土台にガラス乾板をめぐる研究は大きく飛躍すると見込まれる。上記研究報告には、ガラス乾板に係る成果の一端も併載し、他機関で参照されるに適したものとした。

また、上記プロジェクトとともに、ガラス乾板の焼付を主内容とする、史料編纂所所蔵の台紙付写真のうち未登録分を対象に、デジタル写真による複製、複製データの情報化作業を行ない、約 5000 枚につき編纂所既登録分との照合を完了した。また、ガラス乾板・台紙付写真の関係を解明する手掛かりとなる撮影台帳の電子データ化も行った。

研究分担者高橋則英・連携研究者谷昭佳・研究協力者高山さやかを中心に遂行した。

研究上生成される二次的デジタル画像では、くずし字画像を主対象とし、連携研究者井上聡を中心に、研究分担者高田智和・馬場基・末代誠仁の協力のもと、他の研究グループと連動しながら研究を進めた。既存の電子くずし字字典データベースを基盤とし、第一に、史料編纂所所蔵史料および正倉院文書を中心とした字形の登録を継続し、約 32000 件の登録を実現した。第二に、既存登録字形 22 万件余から抽出した代表的字形 25000 件をベースとして、OCR 技術を用いた文字判定の研究を展開し、解析ソフトの試行版を構築した。

業務上生成されるデジタル画像では、複雑化した現状の管理状況の整理を進める一方、管理システムの検討を行った。当初は、「翻刻支援システム」を史料管理に生かすという発想に取り組んだが、現状スタンドアロンで動く「ガラス乾板情報管理ツール」の機能を発展させて、多様なデジタル画像を総合的に管理・運用する方向性を検証した。本研究課題の中では実現は難しいと判断し、別の研究費への継承を目指して、2015年度から科学研究費基盤研究(A)「原史料メタ情報の生成・管理体系の確立および歴史知識情報との融合による研究高度化」(研究代表者山家浩樹、課題番号15H01722)の交付を受け、その柱の一つとして研究を継続することが可能となった。

5. 主な発表論文等

[雑誌論文](計58件)

- 1 井上聡、東京大学史料編纂所「電子くずし字字典データベース」の概要と展望、情報の科学と技術、査読有、65-4、2015、pp.176-180
- 2 谷昭佳、「ガラス乾板の調査・保存・研究資源化に関する研究」の内覧会報告、東京大学史料編纂所附属画像史料解析センター通信、査読無、68、2015、pp.5-7
- 3 竹内涼子、史料編纂所所蔵ガラス乾板の劣化と保存方法の考察、東京大学史料編纂所研究紀要、査読無、25、2015、pp.50-61
- 4 井上聡、デジタル画像を用いた歴史史料管理体系の再構築 東京大学史料編纂所の取り組み、日本写真学会誌、査読無、77-1、2014、pp.41-57
- 5 谷昭佳、写真史料としてのガラス乾板
4年間のプロジェクト成果と課題から、「ガラス乾板の調査・保存・研究資源化に関する研究」予稿集、査読無、2014、pp.5-21
- 6 井上聡・木下聡、日本史研究におけるガラス乾板の史的意義について、同上、pp.23-40
- 7 高山さやか、調書作成データ入力システムとガラス乾板画像のデジタル化について、同上、pp.59-69
- 8 竹内涼子、ガラス乾板の調書作成・整理から見てきた様々な劣化状況と保存方法について、同上、pp.41-57
- 9 山田太造、日本史史料を対象としたテキスト構造化と読解支援、東京大学史料編纂所[共同研究拠点と歴史情報]シンポジウム「資料情報の新たな発信」予稿集、査読無、2014、pp.20-24

10 山田太造・井上聡、A Text Encoding Support System for Pre-modern Japanese Historical Materials、Proceedings of Digital Humanities 2014、査読有、2014、pp. 558-559

11 末代誠仁・白井啓一郎・馬場基・渡辺晃宏・井上聡・久留島典子・中川正樹、古文書字形検索サーバの設計と試作、日本情報考古学会講演論文集、査読有、2014-13、2014、pp.75-77

12 高田智和・小助川貞次、古典籍原本画像と翻字テキストの対照ビューアの作成と教育利用事例、国立国語研究所論集、査読有、8、2014、pp.129-140
<http://www.ninjal.ac.jp/publication/papers/08/pdf/NINJAL-Papers0808.pdf>

13 山家浩樹、史料デジタル収集の基準と方法、東京大学史料編纂所研究成果報告2013-2『史料デジタル収集の体系化に基づく歴史オントロジー構築の研究』、査読無、2013、pp.43-46

14 谷昭佳、黎明期歴史史料写真の研究資源化 ガラス乾板資料を中心にして、日本写真学会誌、査読無、76-2、2013、pp.182

15 谷昭佳、写真のある美術館・博物館・資料館 東京大学史料編纂所、日本写真学会誌、査読無、76-6、2013、pp.439-440

16 白井啓一郎・遠藤友樹・末代誠仁・井上聡・久留島典子・馬場基・渡辺晃宏・中川正樹、Character Shape Restoration of Binarized Historical Documents by Smoothing via Geodesic Morphology、Proc. 12th International Conference on Document Analysis and Recognition、査読有、2013、pp.1317-1321、DOI:10.1109

17 高山さやか・竹内涼子、史料編纂所所蔵ガラス乾板を中心とする写真史料の調査・整理保存に関する研究会について、東京大学史料編纂所画像史料解析センター通信、査読無、61、2013、pp.15-17

18 井上聡、「電子くずし字字典データベース」の課題と将来構想、情報処理学会研究報告・人文科学とコンピュータ研究会報告、査読無、2013-CH-97、No.9、2013、pp.1-5
https://ipsj.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=89426&item_no=1&page_id=13&block_id=8

19 馬場基・渡辺晃宏・井上幸・中川政樹・

末代誠仁、個別 DB の深化と連携確保をめぐって、情報処理学会研究報告・人文科学とコンピュータ研究会報告、査読無、2013-CH-97, No.10、2013、pp.1-4
https://ipsj.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=89427&item_no=1&page_id=13&block_id=8

20 鳥海早喜、写真展「ザ・ファミリー・オブ・マン われらみな人間家族」のガラス乾板に関する調査研究、芸術・メディア・コミュニケーション、査読有、10、2013、pp.1-15

21 山田太造（第一著者、他3名）、人文科学データベース統合検索のためのメタデータとその応用、情報処理学会・人文科学とコンピュータシンポジウム論文集、査読有、2012-7、2012、pp.71-78

22 林讓、東京大学史料編纂所における研究資源の利活用と保全について、人間文化研究情報資源共有化研究会報告集、査読無、3、2012、pp.49-53
<http://www.nihu.jp/sougou/kyoyuka/pdf/reference/07.pdf>

23 大内英範・山田太造・高橋典幸・綱川歩美・林讓・保谷徹・山家浩樹・横山伊徳、Hi-CAT Plus: デジタル史料の検索・閲覧システム、情報処理学会・人文科学とコンピュータシンポジウム論文集、査読有、2011-8、2011、pp.105 ~ 110

24 山田太造・井上聡・遠藤珠紀・久留島典子、日本史史料読解支援のための候補文字検索、同上、pp.43-50

25 山田太造・井上聡・遠藤珠紀・久留島典子、日本史史料における翻刻テキストの構造化支援手法、情報処理学会研究報告・人文科学とコンピュータ研究会報告、査読無、2011-CH-91, No.5、2011、pp.1-8
https://ipsj.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=75517&item_no=1&page_id=13&block_id=8

〔学会発表〕(計20件)

1 谷昭佳、東京大学史料編纂所におけるデジタル画像の作成・管理についての取り組み、「デジタルアーカイブの「色」を考える」セミナー、2014年12月5日、慶応義塾大学メディアセンター（東京都港区）

2 谷昭佳、歴史史料写真の研究資源化 ガラス乾板を中心にして、第4回東アジア史料研究編纂機関国際学術会議、2014年11月12日、ソウル（韓国）

3 山田太造、東京大学史料編纂所のデータベースについて、H-GIS 研究会、2014年5月10日、京都大学地域研究統合情報センター（京都市左京区）

4 井上聡、デジタル画像を用いた歴史史料管理体系の再構築 東京大学史料編纂所の取り組み、日本写真学会・写真保存セミナー、2013年11月1日、東京都写真美術館ホール（東京都目黒区）

5 大内英範、原本所蔵者との複製デジタル史料共有システムについて、第44回デジタル図書館ワークショップ、2013年3月14日、九州大学附属図書館（福岡県東区）

6 谷昭佳、19世紀日本関係古写真の調査収集と日本認識素材の復元的研究、日本資料専門家欧州協会、2012年9月19日、ベルリン（ドイツ）

7 大内英範、史料編纂所の公開 DB とデジタルアーカイブ、日本資料専門家欧州協会、2011年9月8日、ニューカッスル（イギリス）

〔図書〕(計11件)

1 山家浩樹・井上聡・谷昭佳・高山さやか・竹内涼子、東京大学史料編纂所研究成果報告2014-8 『組織によるデジタル撮影画像の生成・管理』、山家浩樹、2015、総頁74

2 高田智和・山田太造（共著）漢字字体史研究、勉誠出版、2012、pp.79-93 173-190 246-258 379-395

6. 研究組織

(1) 研究代表者

山家 浩樹 (YAMBE, Koki)
東京大学・史料編纂所・教授
研究者番号：60191467

(2) 研究分担者

林 讓 (HAYASHI, Yuzuru)
東京大学・史料編纂所・教授
研究者番号：00164971

久留島 典子 (KURUSHIMA, Noriko)
東京大学・史料編纂所・教授
研究者番号：70143534

鴨川 達夫 (KAMOGAWA, Tatsuo)
東京大学・史料編纂所・教授
研究者番号：60214566

高橋 則英 (TAKAHASHI, Norihide)
日本大学・芸術学部・教授
研究者番号：10188039

高田 智和 (TAKADA, Tomokazu)
大学共同利用機関法人人間文化研究機構
国立国語研究所・理論・構造研究系・准教授
研究者番号：90415612

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・研究員
研究者番号：70332195

大内 英範 (OUCHI, Hidenori)
筑紫女学園大学・文学部・准教授
研究者番号：60462173
(平成24年度まで連携研究者)

未代 誠仁 (KITADAI, Akihito)
桜美林大学・総合科学系・講師
研究者番号：00401456
(平成25年度より)

高橋 敏子 (TAKAHASHI, Toshiko)
東京大学・史料編纂所・准教授
研究者番号：80151520

遠藤 基郎 (ENDO, Motoo)
東京大学・史料編纂所・准教授
研究者番号：40251475

山田 太造 (YAMADA, Taizo)
東京大学・史料編纂所・助教
研究者番号：70413937

(3) 連携研究者

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

小倉 慈司 (OGURA, Shigeji)
大学共同利用機関法人人間文化研究機構
国立歴史民俗博物館・歴史研究系・准教授
研究者番号：20581101

高橋 典幸 (TAKAHASHI, Noriyuki)
東京大学・大学院人文社会系研究科・准教授
研究者番号：10292799

井上 聡 (INOUE, Satoshi)
東京大学・史料編纂所・助教
研究者番号：20302656

谷 昭佳 (TANI, Akiyoshi)
東京大学・史料編纂所・技術専門職員

研究者番号：70532670

川本 慎自 (KAWAMOTO, Shinji)
東京大学・史料編纂所・助教
研究者番号：30323661

(4) 研究協力者

高山 さやか (TAKAYAMA, Sayaka)
東京大学・史料編纂所・技術職員