

令和元年5月27日現在

機関番号：12601

研究種目：基盤研究(A) (一般)

研究期間：2014～2018

課題番号：26240049

研究課題名(和文) 歴史知識情報のオープンデータ化にむけたスキームと情報利活用手法の再構築

研究課題名(英文) Mechanisms for providing open access to historical knowledge and information, and the restructuring of methods to permit effective use of historical information

研究代表者

久留島 典子 (KURUSHIMA, Noriko)

東京大学・史料編纂所・教授

研究者番号：70143534

交付決定額(研究期間全体)：(直接経費) 31,500,000円

研究成果の概要(和文)：本研究は、さまざまな目的のもとに設計・構築されたデータベースが、大学・研究機関をめぐる環境の変化から維持・更新に行き詰まりつつある現状を踏まえ、情報構造の転換を図ることで、貴重な学術資源としての収載コンテンツを普遍的に利活用しうる環境整備を推進した。具体的には、東京大学史料編纂所歴史情報処理システム(略称SHIPS)を事例に、属性に応じた知識レポジトリを構築してデータベース群からコンテンツを集約し、さらにオープンデータ形式に転換することで汎用的な利用環境を整備した。

研究成果の学術的意義や社会的意義

学術基盤として活用されているデータベースが、さまざまな事情で維持・更新が困難となる状況はさまざまな分野で発生している。本研究は、歴史系データベースを主対象として、収載されている貴重なコンテンツを失うことなく維持・発展させる方法論を提案した。あわせてオープンデータへの転換を図ることで、データベースや機関の枠組みを越えた汎用的利活用の実現を見通した。

研究成果の概要(英文)：Databases designed and built for various purposes are becoming stuck in maintenance and updating due to changes in the environment surrounding universities and research institutes. Therefore, in this research, by changing the data structure, we have created an environment that can universally utilize the content, which is a valuable academic resource. Specifically, taking the case of the historical information processing system of Shiryo Hensan-jo (the Historiographical Institute, the University of Tokyo) as an example, construct a knowledge repository corresponding to the type of data, collect contents from the database group, and convert it to an open data format. So, we prepared an environment that can be used for general purpose.

研究分野：日本中世社会経済史

キーワード：歴史知識情報 データレポジトリ 知識ベース API オープンデータ 機関間連携

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

2000年以降の電子的ネットワークの拡大に応じて、大学・研究機関では従前の研究蓄積を、データベースとして整備し、その活用を促進した。結果として歴史学研究においても、そうした環境は基本的なインフラとして認識され、広汎に利用されるに至っている。しかし開発に従事した研究者の引退や、研究環境の変化などにより、その維持・更新に支障が生じるケースが散見されるようになった。データベースとデータが一体化していることにより、有用なデータも失われてしまう可能性が予見されたことから本研究を企図したところである。

2. 研究の目的

貴重な学術資源としてのデータを保持・発展させてゆくために、情報システムの総合的な構造転換を図り、個別データベースの消長・再編に左右されない、普遍的な知の蓄積・公開のスキームを構築することを意図した。加えて機関や組織の別をこえた汎用的な知識ベースを設計し、オープンデータなどの手法を用いることで、歴史情報の共有化を推進することを目指した。

3. 研究の方法

研究代表者の所属する東京大学史料編纂所では、1980年代より歴史情報システム(SHIPS)構築を始め、現在およそ30種のデータベースを擁している。本研究においては、このデータベース群を素材として、以下の4つの視点から研究目的の達成につとめた。

(1)DB間に共通するデータ群の析出とDB固有のデータおよび機能の明確化

SHIPS配下のデータベース群に収載されたデータを精査し、共通の属性を有するデータ群を析出することで、データベースと一体となり固有に保持されているものと、分離して汎用的に運用すべきものの弁別を行い、後者をいかに知識ベースに集約してゆくべきか分析を進めた。

(2)共有データ生成に関する標準仕様の策定と知識ベースの整備

知識ベースへと誘導すべきデータ群について、データベースから抽出して汎用的に利活用を可能とするための標準化仕様を策定するとともに、知識ベースの設計および開発を推進した。

(3)知識ベースとDB群を媒介するAPI機能の確立

汎用的なデータを集積した知識ベースに対して、個別のデータベースから参照を可能にするためのAPIを設計・構築し、高度な検索を可能とする環境の整備を目指した。

(4)データレベルにおける機関間連携

設計・構築した知識ベースを、他の機関からの応答も可能となる形へと汎用化することで、従前のデータベース連携と比して、より高度で柔軟な相互運用の実現を目指した。加えて知識ベースのオープンアクセス化を検討することで、研究者個人やグループが自由にデータを利活用しうる環境の構築を目指した。

4. 研究成果

(1)SHIPSが擁するデータベース群から、人物情報を集約し、利活用を可能とする人物情報レポジトリを設計・構築し、19種のデータベースから総計40万件以上の関連データを蓄積した。その結果、前近代における人物情報を総覧しうる環境が整いつつある。

(2)人物情報レポジトリをSHIPS配下のデータベースから汎用的に参照できるAPIを設計・構築し、実際の運用を開始した。同APIは現在史料編纂所が公開を準備している「新花押データベース」に実装され、花押を記した人物やその周辺情報を参照する機能を担っている。また同レポジトリを直接検索するためのインターフェイスも開発し、史料編纂所内で試験運用が始まっている。このインターフェイスには、Creative Commonsライセンスに対応したライセンスマークを表示する機能を与えるとともに、HTMLやCSVに加えてJSONやTurtleといった機械可読な形式での出力を可能とする環境を整え、機械学習、AI、ビッグデータ処理といった利用にも対応している。

(3)人物データを強化するために、史料編纂所がかつて刊行した史料集を中心に、人名索引データをデジタル化ののちレポジトリに収載した。本研究を通じて資源化した具体的な史料は『御堂関白記』『殿暦』『後深心院関白記』『後法成寺関白記』『建内記』『上井覚兼日記』『碧山日録』『二水記』などである。これらデータについてはSHIPSより公開する「中世記録人名索引データベース」から一般に利用することが可能となっている。

(4)SHIPSデータベース群から地理情報を集約した、地理情報レポジトリについては、所内データベースから参照するAPIのみならず、外部からの参照を可能とするAPIの構築を行った。同APIを活用することで、2017年度より国立歴史民俗博物館の「荘園データベース」との連携を実現した。これによって知識ベースを機関間で高度に利活用するスキームを提示した。

(5)知識ベースとなるレポジトリ収載データについては、オープンデータの標準的形式であるRDFへの転換を推進するとともに、データ利用条件に関するライセンスを一括付与するための機能を開発し、検索機能との連動性を担保した。

5. 主な発表論文等

[雑誌論文](計40件)

久留島典子、益田氏系図再考 史料編纂所寄託益田氏関係系図の紹介と考察、東京大学史料編纂所研究紀要、第29号、2019年、pp.10-31、査読無

- 山家浩樹、島津氏と禅宗寺院、黎明館調査研究報告、31、2019年、pp. 65-74、査読無
- 久留島典子、『家ノ家族』の社会的機能をどう考えさせるか？ 前近代日本史教科書記述の検討から、ジェンダー史学、第14号、2018年、pp. 91-99、査読無
- 遠藤基郎、Todaiji's Estate in Its Documentary Record: Perspectives on Obe Estate. (translated by Janet R. Goodwin)、Land, Power, and the Sacred. Janet R. Goodwin and Joan R. Piggott (ed.) Honolulu: University of Hawai'i Press、2018年、pp. 81-106、査読有
- Mamoru Shibayama, Susumu Morimoto, Akiko Tashiro, Akihiro Kameda, Taizo Yamada, Shoichiro Hara、Building an Ontology-Oriented Archaeological Knowledge-Base "ArcOnBase" in Mainland Southeast Asia、2018 Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC)、8579462、2018年、pp. 1-6、査読有
- 馬場基、奈良文化財研究所のICTへの取り組み、日本歴史、848、2018年、pp.9-13、査読無
- 久留島典子、益田家文書研究の可能性、しまねの古代文化、25号、2018年、pp. 100-122、査読無
- 久留島典子、危機に瀕する学術誌—商業化・電子化・オープン化に伴う諸課題、学術の動向、22巻9号、2017年、pp. 54-59、査読無
- 林讓、花押・筆跡データの網羅的収集と汎用的利用をめざして、情報処理学会電子図書館「研究報告人文科学とコンピュータ(CH)」、2017-CH-115、2017年、pp.1-4、査読無
- 山田太造、畑山周平、小瀬玄士、遠藤珠紀、井上聡、久留島典子、前近代日本史史料における人物関係とその時空間変化：天正期古記録『上井覚兼日記』を例に、じんもんこん 2017 論文集、2017、2017年、pp. 61-68、査読有
- Taizo Yamada、Detection of topics from newspaper and its analysis of temporal variations in regions、proc. of PNC2017、2017、2017年、pp. 44-49、査読有
DOI: 10.23919/PNC.2017.8203520
- Taizo Yamada、Satoshi Inoue、Collecting the Name of a Historical Person from Related Historical Material、proc. of DH2017、2017、2017年、pp. 807-808、査読有
- 馬場基、歴史的な文字に関する経験知・暗黙知の蓄積と資源化の試み、研究報告人文科学とコンピュータ、2017-CH-115(9)、2017年、pp. 1-4、査読無
- 久留島典子、女性に関する史料は少ないのか、歴史学研究、952、2016年、pp.1-12、査読無
- 久留島典子、日本中世における武家の養子・婚姻戦略、歴博、198、2016年、pp.11-14、査読無
- 山口英男、史料のデジタルアーカイブと歴史情報処理、人工知能、31巻6号、2016年、pp.787-793、査読有
- 馬場基、奈良文化財研究所における情報技術を活用した史料の利活用の促進、第5回 東亜細亜 史料研究編纂機関 国際学術会議 東京大学伊藤国際学術研究センター会議 アジア歴史資料の編纂と研究資源化、2016、2016年、pp.159-183、査読無
- 山田太造、遠藤珠紀、荒木裕行、井上聡、久留島典子、前近代日本史史料から人名を集める、じんもんこん 2016 論文集、2016、2016年、pp.159-164、査読有
- Taizo Yamada、Satoshi Inoue、A Management of Personal Name with Alternate Name and its Searching for Japanese Historical Study、Proceedings of Digital Humanities 2016、2016、2016年、pp.716-717、査読有
- 井上聡、『中院一品記』の史料学的再検討、大和文華、130、2016年、pp.17-32、査読有
- ②① 遠藤珠紀、誠仁親王の死去と「陽光院御登遐記」の成立、古文書研究、82、2016年、pp.139-141、査読無
- ②② 久留島典子、中世後期の武家と婚姻・養子関係 西国を素材に一、黎明館調査研究報告、28集、2016年、pp.95-104、査読無
- ②③ 山口英男、正倉院文書の書類学、日本史研究、643、2016年、pp.1-20、査読無
- ②④ 山田太造、井上聡、前近代日本史史料に関わる人名情報の収集・蓄積に関する考察、研究報告人文科学とコンピュータ、2016-CH-109、2016年、pp.1-4、査読無
- ②⑤ 山田太造 他、地域情報学の読み解き—；発見のツールとしての時空間表示とテキスト分析、地域研究、16、2016年、pp.267-291、査読有
- ②⑥ 近藤成一、イェール大学の所蔵する日本関連資料について、イェール大学所蔵日本関連資料研究と目録、2015、2016年、pp.3-14、査読無
- ②⑦ 遠藤珠紀、豊臣秀次の叙従一位について、戦国史研究、71、2016年、pp.31-32、査読無
- ②⑧ 木村直樹、「華夷変態」から蘭学へ 蘭学受容の歴史的前提、適塾、48、2015年、pp.40-51、査読有
- ②⑨ 山田太造、地域研究史資料を対象とした時空間的特徴の抽出と場面の構造化、第14回情報科学技術フォーラム講演論文集、14、2015年、pp.409-410、査読無
- ③⑩ Taizo YAMADA, Satoshi Inoue、Detection of People Relationship Using Topic Model from Diaries in Medieval Period of Japan、Proceedings of DH2015、2015、2015年、査読有
- ③⑪ 山田太造、地域研究資料と対象とした時空間情報に着目したデータの構造化、人文科学と

- コンピュータ研究会報告、2015-CH-105、2015年、pp.1-6、査読無
- ③② 有川正俊、まちあるきとオーディオツアー、日本地図学会 学会誌「地図」、53(1)、2015年、pp.89、査読無
- ③③ 遠藤珠紀、徳川家康前半生の叙位任官、日本歴史、803、2015年、pp.65-73、査読有
- ③④ 金子拓、天正四年興福寺別当職相論をめぐる史料、禁裏・公家文庫研究、5、2015年、pp.197-213、査読無
- ③⑤ 久留島典子、明治二年の「軍忠状」、思想、1082、2014年、pp.2-5、査読無
- ③⑥ 久留島典子、中世文書の作成から伝来まで、歴博、184、2014年、pp.10-14、査読無
- ③⑦ 山口英男、史料編纂所と正倉院文書調査、国立歴史民俗博物館研究報告、192、2014年、pp.33-39、査読有
- ③⑧ 清野陽一、山田太造、高田智和、古瀬蔵、人文科学データベースからの人名一覧表示システムの構築、人文科学とコンピュータ研究会報告、2014-CH-103、2014年、pp.1-6、査読無
- ③⑨ Taizo YAMADA, Satoshi INOUE, A Text Encoding Support System for Pre-modern Japanese Historical Materials, Proceedings of Digital Humanities, 2014、2014年、pp.558-559、査読有
- ④⑩ 馬場基、古代木簡研究における情報の活用と今後の課題、人文科学とコンピュータ研究会報告、2014-CH-102(12)、2014年、pp.1-3、査読無

[学会発表](計20件)

Taizo YAMADA, Satoshi INOUE, A Common Base of Knowledge for Japanese Historical Materials and its Application, 2018 Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC)、2018年

Mamoru Shibayama, Akiko Tashiro, Toshikatsu Ito, Masako Marui, Yuni Sato, Hiroshi Sugiyama, Exploring Ancient East-West Corridor of Mainland Southeast Asia "Poster Competition, PNC Annual Conference and Joint Meetings 2018、2018年

Ruochen Si, Masatoshi Arikawa and Ryosuke Shibasaki, A preliminary interface design for polyline-based georeferencing, International Conference on Spatial Analysis and Modeling、2018年

Taizo Yamada, Flow and Utilization of Japanese Historical Data in the Historiographical Institute, International Symposium "DIGITAL HUMANITIES AND DATABASES"、2018年

久留島典子、『家/家族』の社会的機能をどう考えさせるか? - 前近代日本史から -、ジェンダー史学会、2017年

Taizo Yamada, Japanese History Research by Historiographical Institute the University of Tokyo and its Contribution, 14th International Conference on Digital Preservation (iPRES2017)、2017年

久留島典子、日本における『知』の蓄積と共有 - 日本史史料の所蔵のあり方から考える、東京大学附属図書館アジア研究図書館上廣倫理財団寄付研究部門主催シンポジウム、2017年

久留島典子、学術資産とデジタルアーカイブズ、日本デジタル・ヒューマニティーズ学会 JADH2016 プレシンポジウム、2016年

有川正俊、佐藤史弥、イラスト地図鑑賞時の注目点シミュレーション手法、日本地球惑星科学連合2016大会、2016年

遠藤基郎、史料編纂所閲覧室での東南院文書の高精細画像の提供、正倉院文書研究会第35回定期研究会、2016年

山田太造、前近代日本史史料から人名を集める、人文科学とコンピュータシンポジウム「じんもんこん2016」、2016年

山田太造、Text structure of Japanese history historical materials and effort for applying TEI in Historiographical Institute of the University of Tokyo, The 1st International Workshop on Models of Japanese Texts and TEI、2016年

山田太造、地域研究史資料に対するテキストマイニング適用の試み、H-GIS 研究会、2016年

山口英男、正倉院文書マルチ支援システム SHOMUS の概要、第34回正倉院文書研究会、2015年

山田太造、史料編纂所における人名・地名に関するデータについて、H-GIS 研究会、2015年

遠藤基郎、実運用となった Hi-CAT Plus - 新しい発信方式の意義と課題 -、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、2015年

荒木裕行、人物史データベースと近世幕府政治史研究、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、2015年

井上聡、地理情報蓄積システムの構築と SHIPS-DB による活用、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、2015年
久留島典子、東京大学史料編纂所における歴史資料の研究資源化 2006年 2014年、第4
回東アジア史料研究編纂機関国際学術会議、2014年
Taizo Yamada、Text mining for Historical Documents -RDF and Linked Open Data-、
Pre-Symposium of Kyoto University ASEAN Center (Bangkok Office) Opening Ceremony and
Commemorative Symposium、2014年

〔図書〕(計9件)

遠藤基郎 他、吉川弘文館、中世荘園村落の環境歴史学 東大寺領美濃国大井荘の研究、
2018年、288
馬場基、吉川弘文館、日本古代木簡論、2018年、360
山口英男 他、吉川弘文館、史料・史跡と古代社会、2018年、564
金子拓、河出書房新社、織田信長 不器用すぎた天下人、2017年、199
林讓 他、同成社、地域社会の文化と史料、2017年、458
久留島典子 他、お茶の水書房、歴史をひらく 女性史・ジェンダー史からみる東アジア
世界、2015年、252
有川正俊 他、古今書院、地理情報科学 GISスタンダード、2015年、201
久留島典子、長野ひろ子、長志珠絵、大月書店、歴史を読み替える ジェンダーから見た
日本史、2015年、278
久留島典子、遠藤基郎、近藤成一、本郷恵子、林讓、山口英男、伴瀬明美、金子拓、中央
公論新社、日本史の森をゆく、2015年、227

6. 研究組織

(1)研究分担者

研究分担者氏名：林 讓
ローマ字氏名：HAYASHI ,yuzuru
所属研究機関名：東京大学
部局名：史料編纂所
職名：教授
研究者番号(8桁)：00164971

研究分担者氏名：本郷 恵子
ローマ字氏名：HONGO,keiko
所属研究機関名：東京大学
部局名：史料編纂所
職名：教授
研究者番号(8桁)：00195637

研究分担者氏名：柴山 守
ローマ字氏名：SHIBAYAMA,mamoru
所属研究機関名：京都大学
部局名：国際戦略本部
職名：研究員
研究者番号(8桁)：10162645

研究分担者氏名：有川 正俊
ローマ字氏名：ARIKAWA,masatoshi
所属研究機関名：秋田大学
部局名：理工学研究科
職名：教授
研究者番号(8桁)：30202758

研究分担者氏名：山口 英男
ローマ字氏名：YAMAGUCHI,hideo
所属研究機関名：東京大学
部局名：史料編纂所
職名：教授
研究者番号(8桁)：40182456

研究分担者氏名：遠藤 基郎

ローマ字氏名：ENDO,motoo
所属研究機関名：東京大学
部局名：史料編纂所
職名：准教授
研究者番号（8桁）：40251475

研究分担者氏名：木村 直樹
ローマ字氏名：KIMURA,naoki
所属研究機関名：長崎大学
部局名：多文化社会学部
職名：教授
研究者番号（8桁）：40323662

研究分担者氏名：古瀬 蔵
ローマ字氏名：FURUSE,osamu
所属研究機関名：国文学研究資料館
部局名：研究部
職名：教授
研究者番号（8桁）：50462172

研究分担者氏名：山家 浩樹
ローマ字氏名：YANBE,koki
所属研究機関名：東京大学
部局名：史料編纂所
職名：教授
研究者番号（8桁）：60191467

研究分担者氏名：馬場 基
ローマ字氏名：BABA,hajime
所属研究機関名：奈良文化財研究所
部局名：都城発掘調査部
職名：主任研究員
研究者番号（8桁）：70332195

研究分担者氏名：山田 太造
ローマ字氏名：YAMADA,taizo
所属研究機関名：東京大学
部局名：史料編纂所
職名：助教
研究者番号（8桁）：70413937

研究分担者氏名：近藤 成一
ローマ字氏名：KONDO,shigekazu
所属研究機関名：放送大学
部局名：教養学部
職名：教授
研究者番号（8桁）：90153717

研究分担者氏名：小宮 木代良
ローマ字氏名：KOMIYA,kiyora
所属研究機関名：東京大学
部局名：史料編纂所
職名：教授
研究者番号（8桁）：90186809