

地域における文化財情報データの活用の試み —国立歴史民俗博物館の資料から—

後藤 真[†]

概要: 現在、国立歴史民俗博物館では「総合資料学の創成」というテーマにて基幹研究を行っている。そのプロジェクトの中で、1970年代前半の日本全国の文化財情報を集積したカードがあり、そのカードのデジタル化と活用を検討している。このカードは1970年代前半の(特に未指定の)文化財のスナップショットを示すものであり、この情報そのものが重要な意味をもつ「資料」足りうる存在である。また、大災害時の文化財のレスキューの間接的な情報源ともなりうるであろう。この資料を現在 Linked Data にて作成することを試みており、メタデータと表示形式を検討している。本発表は、その作成状況についての中間報告とするものである。

キーワード: 地域資料のデジタル化 Linked Data 研究資料情報基盤

To explore applications of Cultural property data in the area -Using the cards of The National Museum of Japanese History (Rekihaku)

MAKOTO GOTO^{†1}

Abstract: The National Museum of Japanese History (Rekihaku) conducts the core research project “Constructing Integrated Studies of Cultural and Research Resources”. One of this project is to explore digitization and applications of cards building up information of Japanese cultural properties in early 1970’s. These cards are snapshots of cultural properties (especially undesignated ones) in early 1970’s, and the information has an important implication for “cultural and research resources”. It can be used indirectly to aid rescues of cultural properties in times of major disaster. Rekihaku attempts to apply Linked Data to these cards, and examines their metadata and display formats. This paper presents the interim report about our current results.

Keywords: digitization of Regional Resource, Linked Data,

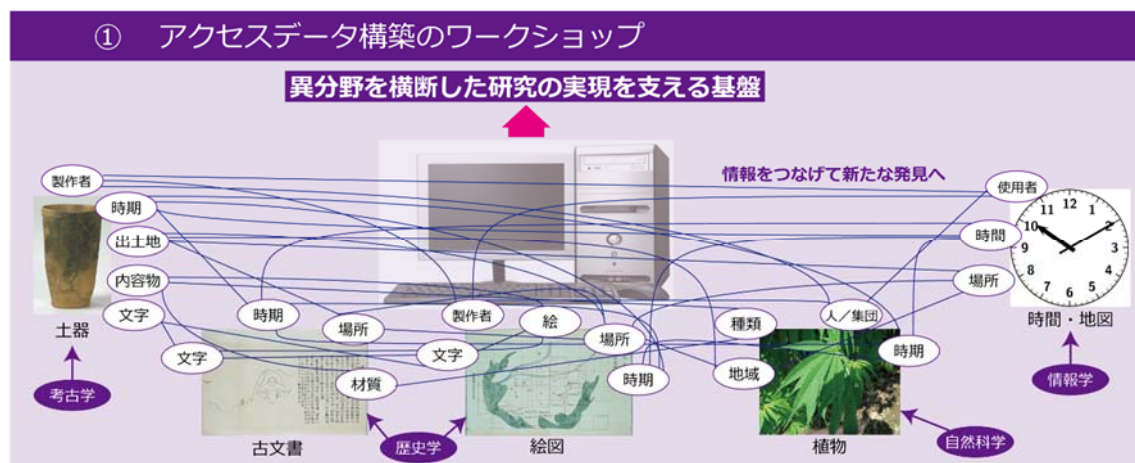
1. はじめに

現在、国立歴史民俗博物館では「総合資料学の創成」という基幹研究プロジェクトを実施している。そのプロジェクトの一環として、Linked Data による情報基盤の構築を行っている。その中でも、本報告では、1970年代前半に実施された、文化財調査によって作成された調査カード、通称

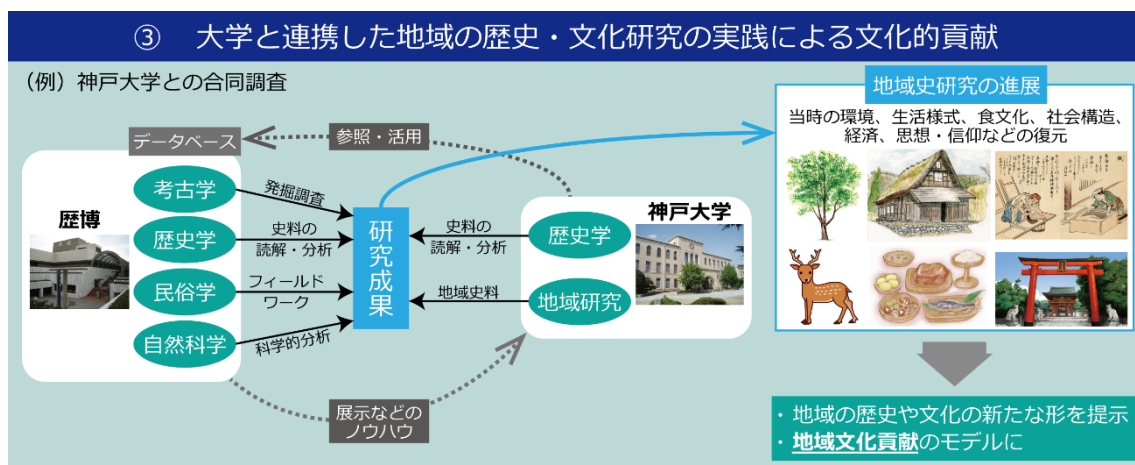
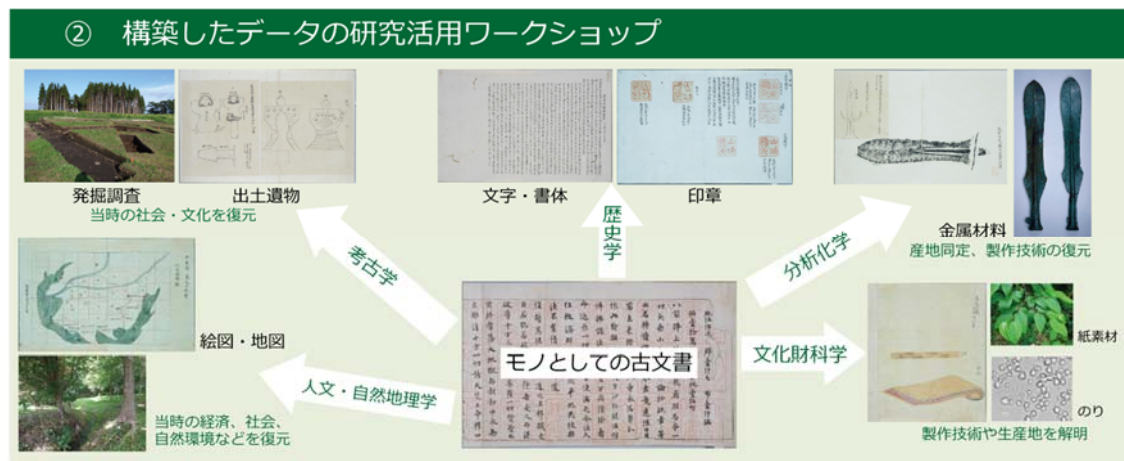
「歴民カード」のデータベース化について報告するものである。

2. 総合資料学の概要と情報基盤

「総合資料学」とは、主に大学等研究機関及び歴史系博物館が持つ資料を多様な形で分析・研究するための学問である。歴史系博物館といえども、いわゆる歴史学・考古学



^{†1} 国立歴史民俗博物館
National Museum of Japanese History, Japan



だけの資料ではなく、多種多様な資料を抱えている。古文書や、考古遺物は無論のこと、自然科学にかかわる植物・昆虫標本など、また技術史にかかわる資料などを抱えていることも珍しくない。無論、大学等の研究機関であれば、研究者の研究の過程で生まれてきたノートなどの中間生成物もあるであろう。これらは、人間の営みを示す基礎資料であり、過去を解き明かすための重要な資料であることは言うまでもない。これらの多様な資源を用いることで、歴史研究を今まで以上に分野横断的なものとし、日本歴史の解明という歴博のミッションの貢献に資することを狙っている。

また、単一の資料であっても、扱う研究分野により見方が異なる。このような「見方」の多様性について検討し、一つの資料を分野横断的に検討することも、一つの眼目としている。

たとえば古文書であっても、その見方によって、多様な側面を持っていることが浮かび上がる。歴史学の営みに即してみれば、古文書はその内容もさることながら、誰から誰へ伝えられたのかという内容伝達の方向性、どのような文脈の中で出されたのかという、時間的な流れの中での位置づけといった、物質としての文書を越えた、当時の複雑な時間的・空間的・社会的文脈の中に位置づけられることになるであろう。

一方で、「古文書そのもの」に焦点をあててみると、例えば文書の書かれ方、文体などについて分析を加えることもあるであろう。その場合には、歴史学だけではなく、その国の語学（史）を外して検討することは困難である。

さらに要素を分解し、一文字単位になった場合には、語学研究の中でも、文字研究の分野の要素が求められる。当時、どのように文字が書かれたかというのは、当時の文化に加えて、社会的な文書の位置づけなどの検討に重要な示唆を与えることになるであろう。

より細かい情報を見ていけば、(日本であれば) 墨・紙などの情報にまでたどり着く。これら墨・紙の材質を分析するためには、技術史的な視点を欠くことはできない。また、紙などの材料である植物の利用などに踏み込んでいけば、当時の自然環境を検討するための重要な材料ともなる。これらについては、自然史との協業によって明らかになる部分が多い。

このように、古文書一つをとって見ても、検討すべき要素は多種多様である。これらの諸要素にすべての検討を、一つのディシプリンの中で行うことが困難なのは自明である。複数の研究分野の協業によって検討することで、はじめて可能となる。そこで、総合資料学の場を活かした多様な研究を試みるものである。

無論、これらの問題意識については、例えば日本のある

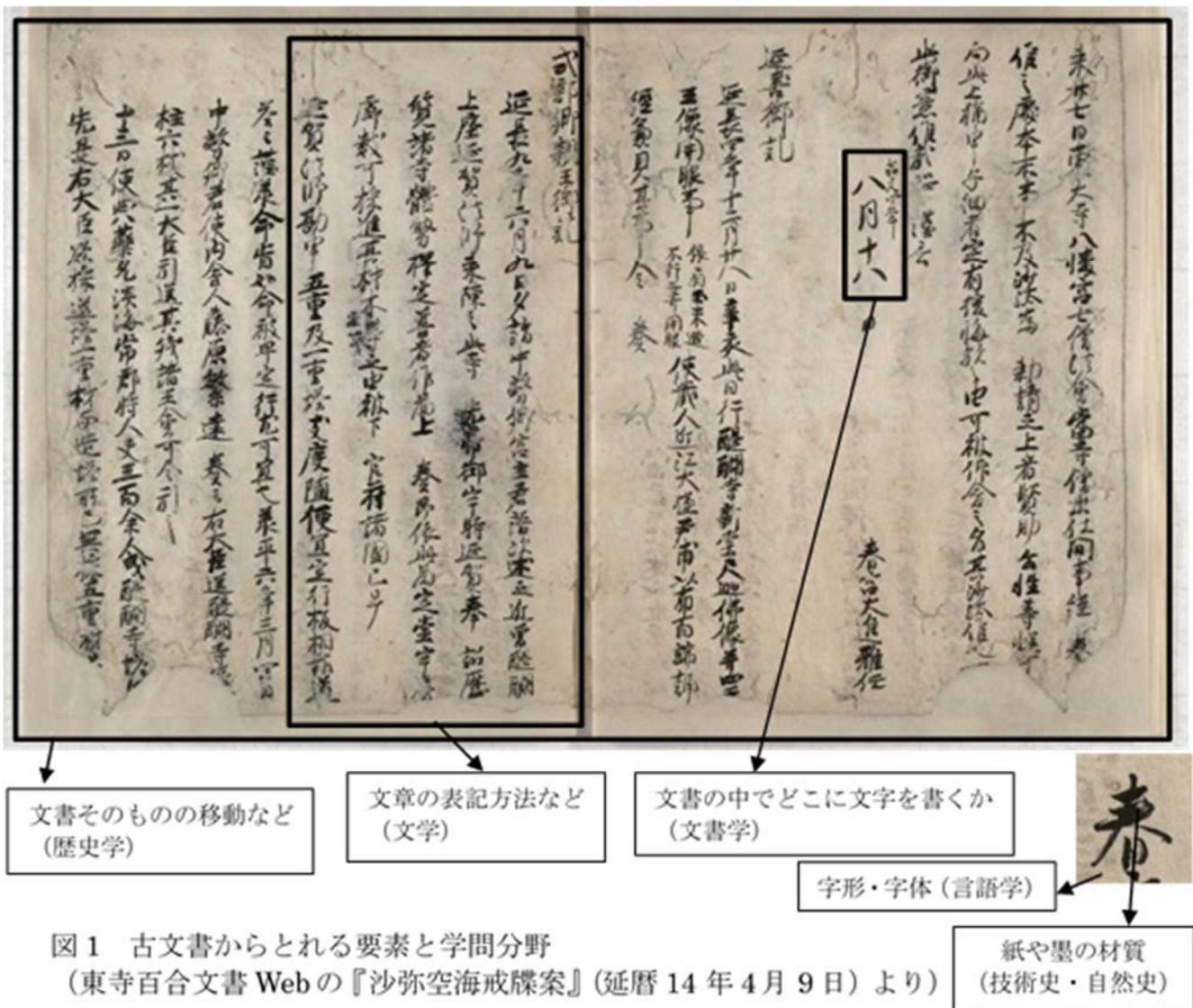


図1 古文書からとれる要素と学問分野
 (東寺百合文書 Web の『沙弥空海戒牒案』(延暦14年4月9日)より)

地域を研究した「地域学」など、各大学等で先行した事例がある [1]。また、(学問分野としての) 地域研究などの文脈においては、すでに蓄積が多数ある。これらの「地域学」・地域研究の枠組みに学びつつも、地域といった空間的切り口だけではなく、より多くの切り口で、これらの研究を行うことを目指している。

また、歴博であっても考古学などでも、自然科学的手法を多く取り入れている。これらの手法の取り入れをより発展させることも目指している。

上記のように、多様な切り口で、かつ資料を総合的に多様にみることで、新たな歴史像を作り出そうという「場そのもの」を「総合資料学」と位置付けている。

3. 総合資料学のための情報基盤

これらの多様な情報を、活用するためには、より効果的な情報発見モデルが必要となる。研究分野の要請により、資料に求める情報が異なるということは、資料目録に必要な情報が異なるということでもある。

さらに、一つの組織だけではなく、複数の組織の情報を、効果的に発見できなければ、多様な切り口の研究は困難である。歴史研究では、多くの場合には一つの資料群に閉じることではなく、複数の資料群を横断的に検討することが重要となる。しかし、実際には、大学や博物館が持つ目録は、機関独自の実務の中で作られており、機関ごとに異なるものであるとってよい。これらの多様な目録も、横断的に検索できるデータベースシステムこそが重要であろう。

本総合資料学で実現したシステムの全体像については、別稿に譲るとして、本稿では、ポイントのみ、簡単に触れておく。

1. 一つの資料が複数の目録を持つことを許容するモデルにすること
- 2 章で触れた古文書のように、一つの資料に関し、求められる情報は複数のものになる。しかし、これらの情報を、単純に一つのデータベースに入れることは、情報発見の観点から考えても、決して効率のよいものではない。そのため、資料をキーに、複数の目録を持つことができるモデルが求められる。

2. 分野横断を実現するために、セマンティックな検索や、芋づる式の情報発見を可能とすること。

歴史研究において、データベースを用いる場合、自らの専門分野においては、多くの場合、キーワードを含む文字列での検索は比較的容易である。場合によっては、検索によらない情報発見（資料群の傾向からあたりをつけて、総攬するなど）も可能である。一方で、共同研究を行う際の関連分野の情報等については、その入れるキーワードが適切であるかどうかを含め、検索が困難である場合が多い。そのため、可能であれば、自分の専門に近い情報を持った目録から、隣接の情報をリンク等で引き出すことで、新たな知見へといたることが可能になると思われる。

3. 複数の機関や異なるコレクションの目録を、横断的に検索し、かつ、統合的にならない検索モデルとして考えること

機関によって似ているが異なる目録を横断的に検索するためのモデルが求められる。本来、異なる目録を横断的に検索するためには、データベース作成の段階で、決まった項目へと統一することが望ましい。しかし、それを強いることで、それぞれの機関やコレクションが持つ目録分類の多様性が失われると同時に、実務的に統一が困難であることは、想像ができる。

また、統一した項目を多数持たせるだけのモデルは、研究のために多数の目録を用意するという、本来の趣旨に反したものになってしまう。そのため、統一せず、可能な限りゆるく似た項目の関係性をマッピングすることで、横断的な検索を可能とするしくみが求められる。

これらの諸要素を勘案して、データベースには **Linked Data** の発想に基づく **RDF** ベースのデータベースを採用することとした。

以下のような実装を目指している。

1. あるデータをキーとして、他の情報へとアクセスすることを可能にする。

例えば、古文書の目録にある材質から、同じ材質を持つ資料はほかにもどのようなものがあるか。ある古文書の人名から、他館にある同一人名の文書を探すなどを可能にする。

2. 検索結果から資料を発見し、その資料が持つ他の分野の目録を見つけ出す。

例えば、文化財情報として小判を発見した際に、その小判に関する、技術史的な目録へと、アクセスすることで、より多様な情報を得る

3. 一つの資源に対して、URI を一つふるという原則が踏襲される

これにより、一つの資料に対して、一つのアドレスを作ることが可能となる。そのため、資料の URL そのものを参考文献や注としてあげることが可能となり、研究のための資源として、より用いやすいものとなる。

4. 複数の目録でも詳細とシンプルな **Dumb-down** されたも

のを同時に持つことで、横断検索にも対応し、かつ統一的な縛りを設ける必要がなくなる。

これにより、機関独自やコレクション独自の項目による検索と、より汎用的な世界標準のような項目検索の両者を持つことができるデータとして持つことができる。データを呼び出すアプリケーション側の工夫で、必要に応じた項目名での検索を可能にできる。

このようなモデル設計をすることで、柔軟でかつ研究にも使えるデータベースとして作成することが可能であるとの見通しを得ている。システムの詳細な解説や展望、課題は別稿に譲るが、このようなシステムの実現によって、情報基盤の構築が可能となるのではないかと考えている。この全体像の中で、歴博が所蔵する歴史民俗調査カード、通称「歴民カード」をデータとして入れ、活用を目指すこととした。

4. 「歴民カード」の概要

歴博の館蔵資料と合わせて、**Linked Data** の機能検証と、資料の公開の方法を検討すべく、「歴民カード」をコレクションの一つとして **RDF** 化し、データベースとした。これにより、情報基盤としての有用性を検討するものである

「歴民カード」は、歴博開館（1983年）に先立つ、1972～1974年に文化庁によって実施された、全国の文化財調査とその状況を示したカード群である。歴史・考古・民俗の3部門にわかれ、全体としては「歴史民俗調査カード」と呼称されることから、通称「歴民カード」と呼ばれている。

「歴民カード」が作成された目的は、必ずしも資料上では明確に残っておらず、判然とはしない。ただし、歴史・考古・民俗という現在の歴博の枠組みと近似していることや、歴博開館後に、歴博に残されたという点からは、歴博の開館に際して何らかの調査が行われたのではないかという推測だけは成り立ちうる。

歴民カードは、文化庁から全国都道府県の文化財担当者に委託する形式となっており、調査者名がすべてのカードに書かれている。また、歴史・考古・民俗それぞれ少しずつ調査の項目が異なっており、それぞれの学問特性を活かした記述方式となっていることが特徴である。

カードの総枚数は65000枚を超える。そのうち、30%が歴史、20%が考古・50%が民俗という比率である。

おおむね、以下のような項目が作成されている。

共通

都道府県名／県名番号＋年番号＋通し番号

歴史

名称／所有者（住所・電話番号付き）／保管者（住所・電話番号付き）／発掘地／時代／作者／法量／品質／形状（ここに概要も記す）（賛・奥書・銘文）／伝来／写真拓本等／備考／指定の有無

考古

時代／遺跡（遺構）名／保管者／所有者／遺跡名／遺跡所在地／遺跡の概要／発掘調査者／参考文献／遺物（遺構）の原状／寸法／伴出遺物／備考／指定の有無

民俗

標準名／種別／地方名／採集地／採集年月日／旧所有者／採集経過／製作地／製作年代／製作者／採集の可能性／製作法・材料／使用地／使用年代／使用者／使用法／分布・由来／備考／保管者（住所・電話番号付き）

図2 歴民カードサンプル（表）

これらの資料情報が書かれている面の背面にあたる裏側には原則としてモノクロの写真が付されており（一部にカラーや図面などの例外もある）、資料の状況がわかるようになっている。

歴民カードの特徴は大略以下のようにまとめられる。

1. 1970年代前半の全国の文化財の情報を示すスナップショットである。

総枚数が65000枚を超える文化財情報が記録されており、一部には指定文化財があり、かつ、いわゆる「民具」のような個人蔵の資料までまとまって記録されている。博物館や資料館などのデータだけではなく、各地域の文化財を調査しているという点では、貴重な全国調査の例である。

ただし、対象となっている文化財については、都道府県による精粗があり、統一的な情報を示しているものではない。

2. 地理情報がすべてのカードに存在する
3. 時間情報も多く含まれるが、正確性が低い。



図3 歴民カードサンプル（裏）

4. 資料によっては、現在残されていないものも存在していると考えられる。

このような性質の資料であることにかんがみると、一定の歴史的価値を持つと同時に、日本全国の文化財の基礎的情報になりうると考えられる。また、歴博が長く所蔵していた資料群であり、それらを公開することは、総合資料学がもつ「大学・博物館資料をより活かした存在にする」というミッションにもかなうものである。そのため、これらのカードをデジタル化し、Linked Dataで公開することにした。

5. 歴民カードの Linked Data 適用

そこで、歴民カードに RDF によるデータを作成することを試みている。基本的には Fuseki による SPARQL エンドポイントの構築と、SPARQL Timeliner によるデータの構築とプロトタイプ構築を実施した。

実際のプロトタイプアプリケーションでは、館蔵資料情報に即した検索画面とセマンティックな検索の実現、および、コレクションごとの資料情報の提供モデルと、緯度経度情報によるアプリケーションの可能性について、検討を行った。





図5 検索結果の一覧

地名情報・材質名・時代名などをクリックすると、絞り込み検索となる

図4 検索画面

検索については、館蔵資料や他のコレクションと横断的に検索できるモデルとした。これは総合資料学が様々な異なるデータベースを同時に検索できるものとして、作られていることとも関係する。また、地域の博物館資料と、歴民カードの情報を同時に閲覧することによるメリットも想定されている。

検索結果には、詳細画面へのリンクだけではなく、多数のリンクが存在している。これは、検索結果から芋づるで

情報を得ようとする場合を想定したものである。

図5については、例えば、材質名等からリンクを張るように設計している。この場合は、検索結果のうち、紙を材質とするもののみを絞り込んで表示させることを想定したものである。絵画等でも、歴民カードから別の館のものを芋づる式に探すことを想定している。無論、紙などの材質では一般的過ぎて使いにくいであろうが、ある決まった材質の情報などが充実してくることで、より高度な研究のためのデータ提供を可能とできるであろう。

当然ながら、上記検索結果は、すべて http のアドレスと



図6 詳細画面

材質名・時代名などをクリックすると、同じ情報をもつ別の項目にアクセスできる。また、博物館名から、博物館の情報にアクセスでき(図5)、そこからさらに館の資料が検索可能となる。

して表示されるため「検索結果の保存」は、アドレスの文字列のみで可能となる(6)。そのため、「研究者の検索結果」などを、リンク集として、自分のサイト等で表示することも可能となる。

図6は詳細画面である。詳細画面からは、絞り込みではなく、興味のあるデータでの再検索を行うような動作を仮に設定している。これは、この資料詳細情報が、一つのURIを持つことを想定しているためであり、Googleなどでこの資料情報に直接アクセスすることも考えられる。その場合には、ここから、情報基盤の検索を行いたい場合の想定と、異なる「芋づる」のモデルを示すためでもある。

また、この資料詳細情報と、機関リポジトリなどの情報を結び付けることで、資料と研究とを結びつけることが可能となるリンクを作成することもできるようになる。研究成果と資源の結びつけを、資料のレベルで可能とできるモデルである。

さらに、この資料詳細画面からは、コレクションごとのリンクも設定されている。このリンクにより、横断検索で資料、もしくはGoogleなどの検索エンジンから資料にアクセスした場合にも、そこから「この資料を持っている博物館はどのようなものか」や「この資料のコレクションは、ほかにどのような資料があるか」などを検索することを可能としている。

また、この詳細情報から、別の研究などで用いたメタデ



図7 SPARQL Timelinerによる資料情報のプロット

ータへのリンクを張ることで、基盤用目録から、より詳しい研究目録への接続も可能となるであろう。

図7では、実際のアプリケーションのモデルを想定したものを作成している。

「歴史カード」の緯度経度情報を仮に入れたものによって、地図等による表示を可能とした。「歴史カード」には、前述のとおり、必ず地理情報がある。そのため、空間的な把握を行いやすいというメリットがある。当時の文化財の全体の状況を把握するためにも、有益なものであると考えられる。



図8 歴史カードのうち、「歴史」カテゴリの地理情報をプロットしたもの

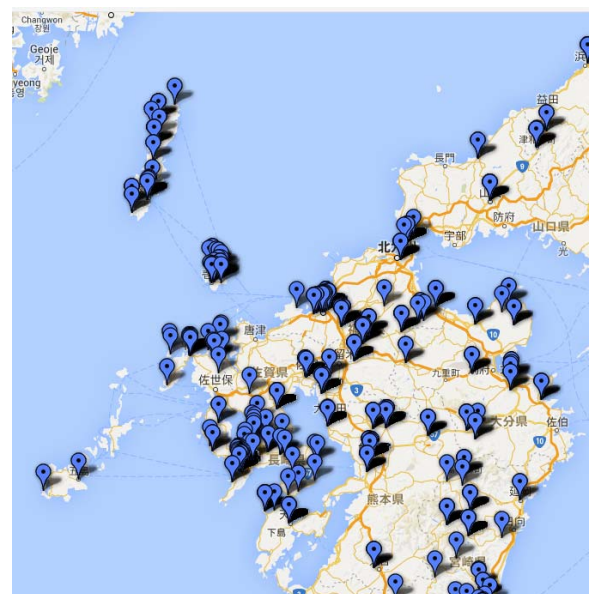


図9 図8のうち、九州北部

6. プロトタイプの課題

歴史カードの全体像をRDFで効果的に表現するための工夫は行えているが、以下の課題をクリアしなければならない。

1. 何と何をリンクさせることが望ましいか。

データが増えるにしたがって、多くのリンクを作成することで、よりリッチなコンテンツになることは想定できる。しかし、複雑なリンクは、ユーザを混乱させることになりかねず、どのようにリンクの基準を選定するのかは課題となる。また、歴史カードは、緯度経度以外にどのような情報とつながるのが最も望ましいのか、などの資料の特性を

活かしたリンクを検討することが肝要となる。

2. 用語そのものの揺れを吸収するためのオントロジ

RDF はデータベースの要素を合わせるための項目を横断的に処理することは、比較的得意だが、項目の中の用語の揺れなどを処理することはできない。この揺れを処理するためには、ある程度オントロジなどでの処理が必要となってくるであろう。このオントロジモデルの構築は、将来的な課題であると同時に、オントロジ事態の構築が、総合資料学から「メタ資料学」への可能性を胚胎するものであると考えている。

7. 地域資料の文化財情報を記録するということ—まとめにかえて—

本プロトタイプは、RDF モデルの有用性と、リンクによる検索が、どのように有益かを検証することを眼目としている。このモデルの中で、いかにして、より複雑でニーズにそったデータと、基盤を両立させうるかが、この総合資料学の情報ネットワークを成功させうるかの、カギとなるであろう。本システムの主眼は、あくまでも研究のための情報基盤だが、一方で、文化財のバックアップとしても重大な意味を持つことが分かる。

とりわけ、最近、列島で起こる大災害のようなものについては、起こってからではなく、起こる前のリスト化が重要であると近年指摘されつつある。そのため、各地で動きが進みつつある。そのような動きに対応する際の材料としても重要な意味があるのではないかと考える。

もちろん、40年前の文化財の状況が現在に直接つながるわけではない。その後の動き等も十分に考慮しながら、データの公開方法を考えていく必要があるであろう。本データ群の持つポテンシャルを、地域の文化財のあり方に還元していくことで、今後の地域の歴史や文化の見直しにつながるようなデータ構築を、今後とも目指していくものである。

謝辞 本システムは、人間文化研究機構基幹研究プロジェクト「総合資料学の創成と日本歴史文化に関する研究資源の共同利用基盤構築」による研究の一部であり、多様な目録ネットワーク構築と融合型研究の基盤の一部として構築されている。

参考文献

- [1] 『地域学創出のための医文融合型研究（略称・地域学創出プロジェクト）2011年度報告書』佐賀大学地域学歴史文化研究センター編、2012年3月