

# れきすけ：歴史ビッグデータで知識と経験を共有する 異分野間協働プラットフォーム

市野 美夏 (ROIS-DS 人文学オープンデータ共同利用センター/ 国立情報学研究所)

増田 耕一 (東京都立大学 都市環境学部)

北本 朝展 (ROIS-DS 人文学オープンデータ共同利用センター/ 国立情報学研究所)

「れきすけ」とは、史料を利用するさまざまな分野の研究者をはじめとする史料に関わる人々が、協働して史料の知識や経験を共有するプラットフォームである。複数の情報提供者によるデータ構築、複雑な史料のデータ構造、多分野で利用するための時空間情報など、史料の情報共有における課題と、研究データである側面における課題の双方を解決するため、情報を複数のカードに分けた、データベースとそのためのユーザインターフェースを構築した。

## REKISKE: An inter-disciplinary platform for sharing knowledge and experience in Historical Big Data initiative

Mika Ichino (ROIS-DS-CODH / NII)

Koiti Masuda (Faculty of Urban Environmental Science, Tokyo Metropolitan University)

Asanobu Kitamoto (ROIS-DS-CODH / NII)

REKISKE is a platform for researchers in various fields as well as other people who are involved in maintenance and applications of historical documents to collaborate and share their knowledge and experience with each other. We develop a new data schema for REKISKE, which is composed of several cards divided into from each record of information. It is intended not only to share information of historical situation records with spatio-temporal information, but also to manage data from multiple information providers for multidisciplinary studies.

### 1. はじめに

地震、天文、気象などのほか、経済、社会など多様な分野で、歴史資料（以下、史料）を利用した研究が行われている。それらの研究では、史料探査や史料整理といった史料の処理過程で共通の作業があり、同じ史料を利用することもある。市野ほか [1]は、データ構造化などの処理過程だけではなく、史料に関するさまざまな点で、分野を超えた協働を提案した。そして、史料に記された地震や天候など、記録者の周囲の状況に関する情報を歴史的状況記録とよび、歴史的状況記録データベース構想を揚げた。

史料などの歴史的な情報を利用する多分野の研究者が、研究上のさまざまな過程で協力および協働することで、負担軽減に加え、研究の質も向上する。そして、市野ほか [1]から発展し、2018年3月にはCODHセミナー [2]が開催され、その後のさまざまな機会 [3]を通して、「歴史ビッグデータ」という枠組みが構築された。図1に示すように「歴史ビッグデータ」の枠組みは、情報基盤という機械による多分野のつながりと、研究それ自体およびそれを進める人々のつながりを

含めたものと考えている。周辺を囲む研究分野には、地震、気象などの自然科学から、地理、経済、社会科学、さらに哲学など、多岐におよぶ。



図1 歴史ビッグデータの概念図

画像の典拠：\*1 [4], \*2 [5]

Figure 1 Concept of Historical Big Data

「歴史ビッグデータ」は、史料のデータ整備などの情報技術開発を目指す固有のプロジェクトではない。人の活動とそれを支える情報基盤全体を指し、史料を利用する研究を協働して進めるイニシアティブである。史料を利用する分野同士の協働による研究推進、新しい研究分野の創出を目指している。すでに、史料のデータ構造化プロジ

ェクトや、さまざまな分野横断的な共同研究プロジェクトが始まっている [6], [7], [8].

## 2. 動機と目的

史料を利用し研究する中で、研究成果だけではなく、史料自体に関するさまざまな情報、いわゆる知識と経験も得る。これらはデータでも研究成果でもないが、情報としては重要である。たとえば、「〇〇日記には毎日の天候があり、研究に利用しているが、地震の記録もある」という情報である。それは、図書館の書誌情報には含まれていないが、書誌情報と合わせると、ある分野の研究には役立つ。

それだけではなく、史料を利用した研究では、その学問分野独自のデータ処理とは別に、史料の処理に関する知識や技術も必要となる。しかし、両方の知識や技術を十分に備えた者のみが研究を進めると、史料の利用者は減り、図書館、資料館など史料を所蔵する側にもデメリットとなる。

各分野で培った史料に関する情報、いわゆる経験や知識を相互に共有することは、史料を利用する研究者だけではなく、利用を促進したい所蔵者にとっても有益である。しかし、これまで、このような知識や経験といった情報の共有は、それを持っている者と、必要な者が偶然出会うという運任せであった。

そこで、我々は史料の利活用における、さまざまな作業過程で得られた知識や経験を統合し、相互利用するためのプラットフォームとして「れきすけ」 [9]を構想した。「れきすけ」は、史料を巡る知識と経験を共有し、このプラットフォームの上で、人々が協働し、「歴史ビッグデータ」を促進していくことを目指す。本研究は、「れきすけ」プロトタイプ [10]とバージョン 1 [11]の結果を検証し、データスキーマの再構築、情報の取得方法と活用のためのユーザインターフェースを開発した。

## 3. れきすけの要件

「れきすけ」が目指すプラットフォームに必要な機能とは何か。「れきすけ」の対象は、史料そのものやそのデータではなく、知識や経験である。知識や経験の多くは人の中に存在する情報である。それはどのような情報であり、どのように共有するか。まず、対象となる具体的な情報の項目を考え、そのためのスキーマや情報の取得方法を検討する必要がある。さらに、利用者や情報を増やし、史料の活用を促す戦略も重要である。

八王子市で、1720年から現在まで書き継がれている石川日記 [12]を例に「れきすけ」の要件を整理する。石川日記には、天気、地震をはじめ、農業に関する記録や米価や作物の値段など、さまざま

な情報が記載されていて、近世史、農業経営史のほか、気候学をはじめとする地球惑星科学分野でも利用されている。

### 3. 1. 歴史的状況記録

たとえば、気候学者や地震学者は、石川日記に天気の記録があるのか、地震の記録があるのかを知りたい。つまり、研究対象の情報が必要である。

「歴史ビッグデータ」では、研究対象のデータを歴史的記録とし、歴史的状況記録、歴史的状態記録、歴史的行動記録と分類している [11]。どの史料に歴史的記録があるかという情報と、その記録の詳しい情報があると、史料を研究に利用する上で、大きな助けとなる。しかし、一般的な書誌情報に、歴史的記録に関する情報はほとんど含まれていない。

一方、所蔵者にとって史料の利用数が増えることは、継続して史料を保全していくために重要である。史料に歴史的記録に関する情報を付加することで、歴史的だけではなく、多様な分野での史料利用につながる。

本研究では、歴史的記録の範囲は広いので、その全てを扱う前に、地震や天気といった歴史的状況記録を対象として、具体的な情報の項目を検討した。

### 3. 2. 記録期間と記録場所

地球惑星科学研究で典型的に見られるように、歴史的状況記録は、時空間情報とあわせて利用されることが多い。つまり、記録期間と記録場所の情報が必要である。また、複数の史料から歴史的状況記録をデータ化し解析するため、あわせて利用する時空間情報は、数値として処理したい。

そして、多くの史料は旧暦や歴史地名が使われている。新暦で扱いたい人、緯度経度として扱いたい人もいる。新暦と旧暦、地名と緯度経度の両方が利用できることも重要である。そのようなシステムは用語習慣が違う人々の間、たとえば、自然科学分野と歴史分野での時空間に関する相互理解の手助けとなる。

さらに、たとえば天保の飢饉の頃の史料を探したい場合、書誌情報を「天保」などのキーワードによって検索するだけでは、年号や記録時期が明示されていない史料を取り逃がす。時間情報をあわせて収録しておけば、「天保の飢饉の前後の期間（1825年～1845年）」、「天保の飢饉の特定の期間（1835年～1837年）」などの、時期を指定した検索によって、研究目的にあう史料を探すことができる。しかし、十分に史料調査が行われるまでは、年号や年までという、時間分解能の粗い曖昧な時間情報しか得られない場合も多い。曖昧な時空間情報の扱いも課題である。

### 3. 3. 史料の状態

石川日記は、八王子市教育委員会から翻刻された出版物が発行されている。さらに、天気を抽出したデータが、歴史天候データベース [13]に収録されている。この情報があれば、八王子市の資料館に通って翻刻するところから始めなくても、デジタルデータを利用できる。つまり、対象の史料が、翻刻されているか、デジタルテキストがあるか、解析データとして構造化されているか、といった、データ処理、データの構造化がどの程度進んでいるかという情報はとても有用である。加えて、紙かデジタルかといった媒体情報、WEB 公開されているといった公開情報があるとなお有益である。

史料の整理には、多くの時間と手間がかかり、先人の成果を活かし、より処理が進んだデータにアクセスできると大きな時間の節約になる。それだけではなく、データ処理の到達地点がわかると、道半ばで中断している作業を、別の人が継続し、データ構造化を進められる。さらに、異なる作業によるダブルチェックにもなり、品質向上にもつながる。史料のデータ構造化補助システム、たとえば、「みんなで翻刻」 [14]との情報共有も有用である。実際に、気候学者が非常に詳しい天候記録を含む日記史料を画像として収集した。個人の日記の翻刻は難易度が高く、処理が滞っていたが、「みんなで翻刻」が引き継ぎ、翻刻が進んだ。

### 3. 4. 史料の利用

戸谷 [15]は、石川日記を利用し、農業経営史の本を執筆した。それはさまざまな経済史の文献に引用された。そのほか、石川日記を利用したさまざまな分野での研究成果が、出版物、論文、学会などで発表されている。同じ史料を利用した成果の情報は有用である。

同じ分野の研究者は、同じ史料を用いた研究へアクセスすることで、史料に関する問題点や利用における課題を知ることができ、さらに先の研究へ進める。ほかの分野の研究者や史料の所蔵者は、新しい利用方法を知る機会となる。同じ史料を利用している場合、異なる分野や目的であっても、問題解決の緒や新たな発見につながる。

石川日記を用いた気候変動と農業に関する研究を進めていても、1949年の農業経営史の著書にはなかなか気づけなかった。しかし、近世史の研究者と話をすると、すぐに著書の話がでてきた。同じ分野に比べて、異なる分野の成果等へのアクセスは困難であり、史料の利用に関する情報は重要である。

歴史的状況記録の情報の付与は、史料の利用向上につながる。そして、史料がどのように利用されたかの具体的な記録を収録することは、史料所

蔵者やデータを整理した研究者の業績の評価につながる。

### 3. 5. 情報共有のあり方

3.1 節から 3.4 節までにあげた史料利用に関する知識と経験は、各々の利用者の中に存在する。「れきすけ」が提供する情報は、「れきすけ」を公開する機関が所蔵しているデータや史料自体ではない。史料が、さまざまな研究に利用された、または利用される可能性が高いと考えられたことに関する情報である。提供者は、特定されたデータライブラリアン、データキュレーター、データアーキビストではなく、不特定の「れきすけ」利用者である。

たとえば、石川日記は、近世史、経済史から、天文学、地震学、気候学まで、非常に広い分野で利用されている。そして、研究者に限らず、学生、一般市民にも利用されている。つまり、一つの史料に対して、複数の多様な利用者が提供する情報を受け入れる機能が求められる。

史料の利用には多くの作業過程がある一方で、学際的な研究で利用される機会が増えている。そのため、データ提供者とデータ利用者間でさまざまな問題が起きている。「れきすけ」を利用した協働による知識や経験の共有により、このような学際研究の課題解決に貢献することも期待される。

## 4. データスキーマの再構築

3章の要件を満たすことを目指し、システム構築を進めてきた。プロトタイプ [10]では対象となる情報の項目を検討し、つづくバージョン1 [11]では、「れきすけ」とその情報の活用を検討した。しかし、3.5 節であげた複数の多様な利用者が提供する情報を受け入れる機能の構築には至っていない。そこで、本研究では、データスキーマから見直すことにした。

### 4. 1. 先行システム

プロトタイプ [10]では、一般的な書誌情報 [16]に、3.1 節であげた歴史的状況記録に関する情報、3.2 節の史料の記録場所、3.3 節の史料の処理過程を加えたデータベースを構築した。結果として、3.2 節の時空間情報に課題が残った。

そこで、バージョン1 [11]では、時間情報の扱いに注目し、HuTime [17] [18]を用いて新暦と旧暦両方を扱えるようにした。史料をその記録期間で並べて表示できるようになり、地点情報を利用した地図表示機能と合わせ、時空間情報を利用した史料の選定が可能となった。「れきすけ」で扱う情報の基本的な項目も整理され、試験的な公開も行った [19], [20]。

## 4. 2. カード型データ構造

これまでの結果から、3.5 節の課題である複数の利用者からの情報提供を可能にするため、FRBR (Functional Requirements for Bibliographic Records) モデル [21]を参考にして、1レコードで持っていた一つの史料の情報を、複数のカードに分割し、それぞれのカードを独立させたデータスキーマを考えた。FRBR モデルの Work (著作カード)、Item (資料カード) に、「れきすけ」独自の、事項カード、所蔵カード、参考カード、地名カードの4種類のカードを加え、6種類のカードを用いて、情報を記述することにした(図2)。そして、これらのカードを相互にリンクし、独立しているカードを相互参照し、一つの史料の情報を再現する。

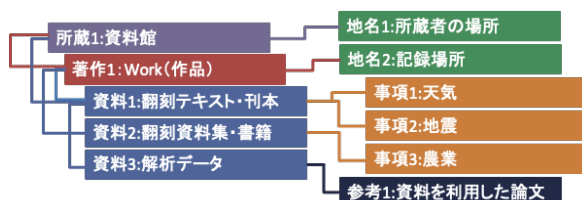


図2 カード型データ構造  
Figure 2 Schema of card-type data model.

資料実体は、もとの作品のうちの部分について、翻刻・出版・解析用データ作成などが行われた結果であることが多い。そのような構造の情報を共有するため、作品そのものについての情報は「著作カード」とし、実体については Expression, Manifestation, Item の3つを分類せず、対象の粒度に関係なく、「資料カード」とした。そして、3.1 節の歴史的状況記録の情報は、情報提供者ごとに大きく異なりうるので、「事項カード」として独立させた。著作カード、資料カード、事項カードにわけることによって、一つの史料に対して、複数の提供者からの情報を持つようにした。

所蔵場所、史料を利用した論文などの成果や情報の元となった資料、記録場所は、複数の史料で同じ情報を持つことがある。そこで、所蔵カード、参考カード、地名カードとして独立させた。それぞれ一つのカードに情報を集約し、複数の史料へリンクするようにした。

前述したように、提供される情報は、提供者ごとに大きく異なりうるため、項目を細かく立てることは難しい。そこで、項目を細かく立てず、「記事」というマークダウンを利用したテキスト項目を用意し、提供者それぞれが多様な情報を自由に記述できるようにした。カード固有の項目も削減され、利用者にもわかりやすくなると思った。

## 5. カード

6種のカードについて、どのような情報を共有するのか、その役割と利点などを紹介する。

### 5. 1. 著作カード

著作カードは作品そのものについての情報を持つ。作者のほか、3.2 節の時間情報と空間情報も扱う。

作者は史料を書いた者で、日記などでは代々の当主や藩の役人など、複数のメンバーから構成される場合もある。石川日記の作者は石川家の代々当主である。作者欄に複数の名をあげることもできるが、「記事」を利用して、詳細に記述することもできる。

時間情報は、記録期間(記録開始と記録終了)として持つことができる。西暦、和暦での入力が可能であり、入力時の表記を残した上で、HuTime API で変換した、西暦、和暦、検索および年表表示のための時間データも持つ。たとえば、石川日記は現在も記録が継続されているが、終了日時が決まっていない情報も利用できる。空間情報は、GeoLOD と連携させた地名カードを参照する。

### 5. 2. 資料カード

資料カードはそれぞれの実体についての情報を持つ。基本的な書誌情報のほか、3.3 節の史料の状態についての情報を、資料タイプと資料タグで持つ。

資料タイプは、史料のデータ構造化の段階の情報で、原史料(写本)、画像、翻刻テキスト、半構造化テキスト、構造化テキストまたはデータ、解析用データ、解析結果データなどである。媒体の情報は多様であるため、資料タグとあわせて、「記事」も利用して記述する。

資料シリーズなども資料タグを利用することで、資料カード同士の紐づけができる。

記録期間の情報も持つことができる。著作カードは作品全体が対象であるが、資料カードの対象の記録期間である。石川日記を例にする。

- ・ 著作カードの記録期間は享保5年～\* (\*は終了日時なしを示す記号)である。
- ・ 八王子市から出版されている翻刻テキスト 1巻～15巻 [12]の記録期間は1720年～1940年、タイプは翻刻テキスト、タグは刊本である。
- ・ 戸谷 [15] (4章 pp.157-349)も農業経営の史料として石川日記から収量と作付面積を翻刻している。記録期間は1720年～1889年、タイプは翻刻テキスト、タグは刊本である。
- ・ 歴史天候データベース [13]にも石川日記は収録されているが、記録期間は1720年～1885年、タイプは解析用データ、タグはテーブルである。補足として、スペース区切りテキスト

トファイルと「記事」に記す。

このように、一つの原史料から派生した出版物やデータの情報は、それぞれ個別のカードで持ち、相互にリンクし参照する。

資料カードの対象には、非公開の資料もあれば、一般公開されているデータなどもある。「記事」に「データ期間は一部非公開、直接交渉すれば利用できる」などの情報を記すと、より利用者の助けとなる。資料カードの情報は、一般の書誌情報とは異なり、「歴史ビッグデータ」へ貢献を考慮し構成されている。

### 5. 3. 事項カード

事項カードは、3.1 節の歴史的状況記録に関する情報や、さまざまな知識や経験から得た情報を持つ。「れきすけ」の特徴を示す重要なカードである。単独でカードの作成は可能であるが、歴史的状況記録に関する情報を付与したい著作カードや資料カードと関連させることが望ましい。

「れきすけ」の根幹となる歴史的状況記録の種類は、この史料がどのような研究に有用かという情報になる。これを「事項タイプ」とし、地震、天気、農業、水害、天文、飢饉、火山、病気、火災、旅行のほか、自由な情報を登録できる。主な対象は歴史的状況記録であるが、歴史的記録の中での分類は明確ではないため、歴史的状態記録、歴史的行動記録に関する情報を含めてもよい。

事項カードに記載する情報は、分野の多様性を象徴し、情報の広がりや予測が難しいため、「記事」を利用することで対応する。たとえば、別所万右衛門記録 [22]の事項タイプは、地震、天気、水害のほか、穀物価格、昆虫、植物季節など多様な情報が並んでいる。穀物価格は米、大豆、大麦、小麦など、昆虫は蚊、蚩など、植物季節は桜などといった、より詳細な内容は「記事」を利用することで対応できる。

事項タグを利用すると、複数の史料を研究目的などで紐づけすることができる。たとえば、天保の飢饉、天明の飢饉といった社会的な災害、浅間噴火、安政江戸地震といった自然災害などのタグをつけ、研究対象の史料や情報を紐づける。

記録期間を持つことができる。史料の記録期間は長い、対象の歴史的状況記録の記録期間は短いこともある。著作カード、資料カードと異なる記録期間を持つことは、研究上有用である。

### 5. 4. 所蔵カード

所蔵カードは、史料の所蔵者である機関、団体、個人の情報を持つ。所蔵場所を独立させることで、複数の史料を所蔵する所蔵者を一元管理できる。所蔵者に関する複数の提供者からの情報が蓄積され、史料収集や調査に役立つ情報が充実する。

多くの資料館がその地域で記録された多数の史料を所蔵している。一つの所蔵カードから、これら複数の史料にリンクすることで、複数の史料

を相互に活用することができる。歴史的状況記録の内容を比較でき、信頼性の向上につながり、研究上での利点は大きい。

### 5. 5. 地名カード

地名カードは、3.2 節の空間情報を記す。複数の史料が同じ場所で記録されることも、史料の中に複数の場所での記録があることもある。地名カードを利用することで、複数の地名カードから史料へ、複数の史料から地名カードへ紐づけることができる。

さらに、GeoLOD [23] [24]を利用し、地名の識別子を他のアプリケーションとも共有できる。すでに作成済みの地名リスト（たとえば歴史地名データ）を利用でき、地点情報を「れきすけ」で管理する必要もない。さらに、地名がない場合は、新たな地名を GeoLOD [23]に追加し、GeoLOD [23]のデータ蓄積にも貢献できるという、相互の利点もある。

また、GeoLOD とは別に都道府県の情報を持つことによって、空間範囲からの史料探査も可能にした。

### 5. 6. 参考カード

参考カードは、史料を利用した研究成果、または、カードに登録した情報の典拠を記す。著作カードや資料カードとのリンクにより、その史料の利用実績を表現する。これは3.4 節で述べたデータ提供者へのフィードバックともなる。

多くの研究で利用されている史料がある一方、一つの研究で複数の史料を利用する場合もある。一つの参考カードから関連する複数の史料へリンクする、または複数の参考カードから一つの史料へリンクすることで、この関係を表現できる。

典拠についても1枚の参考カードとして、複数カードから関連させて利用できる。たとえば、石川日記の事項カードに天気の連続性に関する情報がある。この情報は、石川日記の天気を利用した市野ほか [25]の情報に基づいている。市野ほか [25]の参考カードを作成し、石川日記の事項カードにリンクすると、情報の所在が明らかとなり信頼性が向上する。

### 5. 7. 関連

カード間にはその関連を示すリレーションを付与する。バージョン1では1レコードを構成していたカード間には関連を持たせる。たとえば、事項カードには、歴史的状況記録に関する情報が書かれているが、史料に関する、作者、記録期間、史料の所蔵先などの情報は、事項カードと関連をもった著作カード、資料カード、所蔵カードにある。同じ著作カードや資料カードに、複数の情報提供者からの事項カードや参考カードが関連づ

けられる。カード同士を統合して閲覧することで、情報の充実も図れる上、3.5 節の課題解決にもつながる。

## 6. ユーザーインターフェース

3.5 節で述べたように、「れきすけ」のユーザは、多様である。そのため、よりわかりやすく情報を提供し、必要な史料を探せるような工夫が必要である。4 章と 5 章で述べたカード型データ構造を考え、カード化された情報を扱えるデータベースとユーザーインターフェースを構築した。

### 6. 1. ユーザ

「れきすけ」では、まず、Firebase [26]を利用したログインユーザと、ログインなしのユーザの利用が可能である。3.5 節にあるように、ユーザからの情報提供を受け付けるために、ログインユーザのうち「編集者」権限を持つ人が情報を提供できるようにした。編集者以外のログインユーザは、カードの作成はできないが、カードへのコメントは可能である。

プロジェクトなどのグループメンバーの共同作業のために、カード編集者のほかに編集パスワードを知るユーザもカードを編集可能とした。作業途中の情報は公開フラグで非公開にもできるが、閲覧パスワードを知るログインユーザは閲覧可能とした。

### 6. 2. 登録

#### 【各カードの編集】

カード別に編集画面を作成した。必須項目は少ないが、タイトルとカード著者は、すべてのカードで必須である。カードタイトルは史料のタイトルではなく、個別のカードのタイトルである。カード著者は、カードの情報の内容の提供者であり、内容の責任者でもある。カードの所有者はカードの編集作業をするログインユーザである。両者は同じことが多いが、たとえば、史料に詳しいカード著者からの情報を、カード所有者が代わりに編集するといった状況に対応した。

事項カードの事項タイプ（歴史的状況記録の情報）も必須である。登録される可能性の高い 10 項目がチェックボックスで表示されており、編集を助けるとともに、表記揺れを防止する。表示項目は登録されている頻度の高い値を優先する。

このようなユーザビリティを考慮した機能をほかの編集画面でも採用している。5.3 節の事項タグでは、編集者ごとによく使うタグは異なるため、編集者ごとに利用頻度の高いタグから表示するようにした。たとえば、歴史気候学分野の編集者は天気の詳細度は天気記録の信頼性と関係が

深いと考えている [27]。天気の詳細度の情報をできるだけカードに記し「天気の詳細度タグ」を付与することが多くなる。その結果「天気の詳細度タグ」が優先的に表示されるようになる。

各カードで明示されている項目に当てはまらないさまざまな情報は、マークダウン形式の「記事」を利用し自由に記述できる。たとえば、史料の公開サイトのリンクなども表現できる。

#### 【関連作成】

5.7 節で述べたように、カード間のリンクは複数のカードに分散している史料情報を統合し、閲覧するために欠かせない。そこで、関連させたい史料を選び、関連マークのボタンを押すとショッピングサイトのカートのように選択され（図 3）、登録するとカード間の関連が作成される。



図 3 関連作成画面

Figure 3 Screen of making relations to cards

#### 【時空間情報】

地名カードの作成では、GeoLOD を利用し、位置情報を得ることができるようにした。それにより、編集者が位置に関する情報、緯度、経度などを調べることなく登録できる。また、都道府県については、チェックボタンで登録できる。

5.1 節で述べたが、記録期間の登録は、西暦和暦に対応しており、入力した値から、HuTime API で変換された和暦、西暦が表示され、編集者が確認できるようになっている。

### 6. 3. 検索

必要な情報を得るためには、検索でヒットしたカードだけではなく、複数の関連するカードを検索結果とする必要がある。Solr [28]を利用することで、関連するカード一覧を取得するとともに、時間情報を用いた期間検索にも対応できた。

検索はキーワード検索とオプションの詳細検索がある（図 4）。キーワード検索は、検索ボックスへの入力で、OR の検索も可能である。

「れきすけ」特有の項目やその値は、慣れるまではわかりにくい。そこで、詳細検索の資料カードや事項カードのタイプやタグの絞り込みでは、登録されている値を表示し、チェックボックスで選択できるようにした。たとえば、事項タイプの

地震、天気の表示名から事項タイプの値が指す事柄がイメージできる。さらに、これらの値は、データベースでの登録数の高い値を優先的に表示し、ユーザビリティを高める工夫を施した。

図 4 詳細検索画面  
Figure 4 Search screen

記録期間では開始と終了を指定する。6.2 節の登録と同様、和暦も西暦も利用できる。3.2 節で「天保」の検索例に挙げたが、この記録期間を利用すると記録期間の検索が可能である。また、都道府県の絞り込みで、空間範囲での検索もできる。

## 6. 4. 情報の表示

### 【一覧表示】

一覧表示では、検索条件にヒットするすべての種類のカードが表示される。絞り込みは詳細検索を利用する。詳しい情報が見たい場合は、カードを選択すると、カードの詳細情報が表示される。



図 5 年表表示  
Figure 5 Graph of recorded periods



図 6 地図表示  
Figure 6 Distribution of locations

### 【詳細表示】

対象となるカードの詳細だけではなく、関連づけられているカードの一覧も表示され、ほかのカードにある史料情報もあわせて閲覧できる。

### 【年表表示・地図表示】

市野ほか [11]は、時間情報を利用し、複数の史料の記録期間を視覚的に確認する方法を紹介した。そこで、時間情報から、史料の記録期間をグラフ表示できるようにした(図 5)。史料探査の段階で、史料の時空間情報が視覚的に確認できる機能は有用である。同様に、史料の所在地や記録地を視覚的に確認できるように、地名カードの情報を利用して地図表示(図 6)できるようにした。

年表表示すると、時間を言葉で認識する人と、時間軸で認識する人の対象の共有ができる。つまり、年表表示、地図表示などの可視化は、「言葉のギャップ」を解消する役割もあり、「歴史ビッグデータ」において重要な機能である。

## 7. 考察

多分野における多様な複数の情報提供者に対応するため、カード型データ構造を考え、カード化された情報を扱えるデータベースとユーザインタフェースを構築した。

カード化により、一つの史料に関する複数の人からの情報を個別に登録できるようにになった。情報の対象となる史料の粒度を決める必要もなくなった。また、多様な情報を 1 対多、多対多で関係づけることができ、情報が散在することなく、統合して閲覧できるようになった。加えて、更新や修正、追加も容易になった。

さらに、参考カードや所蔵カードにより、史料周辺の情報も充実できるようになると期待できる。それは、史料を利用するために、さまざまな段階で貢献する人々に光をあてるしくみになると期待している。

多様な提供者からの情報は、相反することもある。知識や経験という情報は、提供者の主観が入りやすく、間違いや曖昧さもある。情報の価値や学術的な意見が異なることもある。新しいスキーマにより、情報の取捨選択で争うこともなく、相互に情報を共有できるようになった。情報が蓄積されると、多様な利用者が価値ある情報を精査できるようになり、信頼性の向上にもつながる。

近世史の研究者である国文学研究資料館の岩橋清美氏に、専門家の視点から、「れきすけ」に登録された情報の確認および修正、追記を依頼した。その時の作業はオフラインであったが、「れきすけ」バージョン 2 により、岩橋氏から提供された情報は、既存のデータを変更することなく、登録できるようになった。より信頼の高い新たな情報が追加され、情報の信頼性も向上した。

## 8. おわりに

「れきすけ」は、自らの足で史料を探索し、翻刻し、孤独にこつこつデータを構築し、研究に利用してきた研究者や、史料関係者の苦勞から始まったアイディアである。技術的新規性を狙ったものではなく、発想の新規性により生まれたシステムである。史料の専門家と、史料を利用する多分野の研究者、双方の要求に応えるように、「れきすけ」のさらなる改善に努めていく。そして、専門分野を超えた協働による学術の革新を目指す。

## 謝辞

本研究は科研費(18H03794, 20K01152)、ROIS機構間連携文理融合プロジェクト、東北大学災害科学国際研究所共同研究の助成を受けました。これらのプロジェクトメンバーの方々のご協力に感謝します。

## 参考文献

- [1] 市野美夏, 増田耕一, 北本朝展, 平野淳平, 庄建治朗. 人文情報学の素材としての歴史気候学の経験, *じんもんこん 2017 論文集*, 2017, p.139-146.
- [2] “人文学オープンデータ共同利用センター: 第6回 CODH セミナー「歴史ビッグデータ～過去の記録の統合解析に向けた古文書データ化の挑戦～」”<http://codh.rois.ac.jp/seminar/historical-big-data-20180312/> (参照 2020-09-03).
- [3] “歴史ビッグデータ研究会” <http://codh.rois.ac.jp/historical-big-data/sig/> (参照 2020-09-03).
- [4] “早稲田大学図書館古典籍総合データベース” <https://www.wul.waseda.ac.jp/kotenseki/> (参照 2020-11-09).
- [5] “国立国会図書館デジタルコレクション” <https://dl.ndl.go.jp> (参照 2020-11-09).
- [6] 市野美夏. 史料が語るいにしへの気候～歴史ビッグデータ過去の記録の統合解析を目指して～, *情報・システム研究機構シンポジウム 2018*, 2018.
- [7] 川口洋. 天明 - 天保期の東北地方における気候と人口 -歴史気候学と人口学との対話-, *日本人口学会第71回大会企画パネル 1*, 2019.
- [8] Ichino, M., Masuda, K., Mikami, T., and Takatsuki, Y. Reconstruction of solar radiation based on historical weather records in Japan - Climatic condition and market economy in the famine of 1830s -, *EGU General Assembly 2020*, Online, 4-8 May 2020, EGU2020-12300, <https://doi.org/10.5194/egusphere-egu2020-12300>, 2020.
- [9] “れきすけ” <https://rksk.ex.nii.ac.jp>. (参照 2020-11-06).
- [10] Ichino, M., Hirano, J., Masuda, K., Kitamoto., A and Den. H. The Metadata Hub for Interdisciplinary Knowledge Sharing of Historical Situation Records, *Proceedings of the 8th Conference of Japanese Association for Digital Humanities*, 2018, p.106-108.
- [11] 市野美夏, 橋本幸恵, 平野淳平, 増田耕一, 北本朝展. 目撃情報の収集による歴史的状

況記録パスファインダーの構築, *じんもんこん 2018 論文集*, 2018, p.343-350.

- [12] 八王子市郷土資料館(編). 石川日記, 八王子市, 八王子市教育委員会, 1979-.
- [13] 吉村稔. 古日記天候記録のデータベース化とその意義, *歴史地理学*, vol.267, No.55-5, pp.53-68, 2013.
- [14] “みんなで翻刻” <https://honkoku.org/> (参照 2020-09-07) .
- [15] 戸谷敏之. 近世農業経営史論, 日本評論社, 1949, 529p.
- [16] 日本図書館情報学会用語辞典編集委員会(編), *図書館情報学用語辞典 第4版*, 丸善出版, 2013, 284p.
- [17] “HuTime” <http://www.hutime.jp/index.html/> (参照 2020-09-03) .
- [18] 関野樹. 時間名による時間参照基盤の構築 -Linked Data を用いた期間の記述と リソース化. *じんもんこん 2019 論文集*, 2019, p. 267-272.
- [19] 市野美夏, 橋本幸恵, 増田耕一, 北本朝展, 平野淳平. れきすけ: 多分野連携による歴史資料に関する情報共有システム, *日本地球惑星科学連合大会*, 2019.
- [20] Ichino, M., Masuda, K., Kitamoto, A., Hirano, J. and Hashimoto, Y., *Sharing Information of Historical Documents through Interdisciplinary Collaboration*, *The 27th IUGG General Assembly*, 2019.
- [21] 和中幹雄, 古川肇, 永田治樹(訳). 書誌レコードの機能要件 IFLA 書誌レコード機能要件研究グループ最終報告, *日本図書館協会*, 2004, 121p.
- [22] 佐藤大介. 別所万右衛門記録, *東北大学東北アジア研究センター*, 2010, 73298p.
- [23] “GeoLOD” <https://geolod.ex.nii.ac.jp>. (参照 2020-11-06) .
- [24] 北本朝展, 村田健史. 歴史的行政区域データセット β 版をはじめとする地名情報基盤の構築と歴史ビッグデータへの活用, *情報処理学会技術報告*, 2020, Vol. 2020-CH-124, No.1, p1-8.
- [25] 市野美夏, 三上岳彦, 増田耕一. 歴史天候記録から推定した日本における19世紀前半の日射量変動, *地学雑誌*, 2018, vol.127, No4, 543-552, doi:10.5026/jgeography.127.543.
- [26] “Firebase” <https://firebase.google.com>. (参照 2020-11-06) .
- [27] 庄建治朗, 鎌谷かおる, 富永晃宏. 日記天気記録と気象観測データの照合による梅雨期長期変動の検討, *水文・水資源学会誌*, 2017, vol.30, 294-306.
- [28] “Solr” <https://lucene.apache.org/solr/> (参照 2020-11-09) .