

## 科学研究費助成事業 研究成果報告書

平成 28 年 6 月 17 日現在

機関番号：62615

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26540178

研究課題名(和文) データ史料批判：非文字史料の情報学的解析に基づくシルクロード像の再構築

研究課題名(英文) Data Criticism: Reconstruction of the Concept of Silk Road Based on Informatics-based Approaches to Non-textual Sources

研究代表者

北本 朝展 (Kitamoto, Asanobu)

国立情報学研究所・コンテンツ科学研究系・准教授

研究者番号：00300707

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究は新しい史料批判の方法論であるデータ史料批判(デジタル史料批判)を提案し、これを主に非文字史料に適用するための情報プラットフォームDCPを構築した。写真や地図を照合するための各種ツールをDCPに統合することで、照合というエビデンスを結合したエビデンスネットワークを構築することができた。さらにエビデンスネットワークをSPARQL言語を用いて意味的に検索し、各種史料に出現するシルクロード遺跡の関係を信頼度に基づき結合する新しい方法を提案した。ついで、データ史料批判の方法を実際のシルクロード遺跡の探索に適用することで、シルクロード遺跡の全体像を把握する見通しを得ることができた。

研究成果の概要(英文)：We proposed the concept of data criticism (digital criticism), a new approach to historical source criticism, and constructed an information platform called DCP to apply the approach mainly to non-textual sources. We integrated several tools for matching photographs and maps into DCP to construct an evidence network connecting evidences obtained from matching. Moreover, we implemented a semantic search on the evidence network using SPARQL language to propose a new methodology for linking Silk Road ruins across historical sources based on the reliability of relationship. Moreover, we applied the data criticism approach to show the first step toward understanding overall relationship among Silk Road ruins.

研究分野：人文情報学

キーワード：デジタル・ヒューマニティーズ デジタル史料批判 遺跡データベース 歴史情報 セマンティックウェブ 非文字史料 データレポジトリ シルクロード

### 1. 研究開始当初の背景

史料批判、あるいはより一般に source criticism と呼ばれる研究プロセスは、多種多様な史料を評価し、読み、解釈するという人文科学研究の基礎をなす。歴史研究では、文字史料を研究者が読むという伝統的な場面での史料批判の方法論は確立しているが、地図や写真、画像といった非伝統的史料を「読む」場合の方法論は確立していない。

史料を批判的に読む態度を表す言葉「source criticism」、あるいは一般的な意味での criticism や critique は、歴史学、考古学、美学、宗教学、ジャーナリズム、法律学、メディア研究に至る、人文科学の広大な分野における基本的概念である。筆者らがこれまで提案してきた「データ史料批判」はこうした方向の研究の一つと位置づけることができる。

その一例にシルクロードの遺跡照合に関する我々の研究がある。そこで鍵となったのは、地図や写真等の空間画像史料の正確性に対する「疑い」であった。空間画像史料は精細 (precise) であるがゆえに正確 (accurate) であるとの誤解を招きやすく、その信頼性を判定する方法も確立していなかった。そこで我々は、地図や写真を「批判的に読む」方法を確立することで地図の誤差を定量的に把握する方法を確立し、誤差を考慮すればシルクロードの遺跡が照合できることを示した。

### 2. 研究の目的

こうした研究成果に基づき、本研究は「データ史料批判」およびそのアップグレード版となる概念である「デジタル史料批判」について研究を進める。

「デジタル史料批判」という名称は、史料批判に有用なデジタル技術を共通の言葉で議論する場を作りたいという意図がある。従来は個別の技術的な方法論 (アルゴリズム) として扱われてきたものを、史料批判という歴史学の共通の俎上に載せることで、概念を共有して議論を深化させるとするのが意図である。歴史学は従来からエビデンスに基づく推論が人文科学の中でも特に重視されてきた分野であるからこそ、デジタル技術を用いた史料批判のモデル化によって、研究の新たな可能性が開けることへの期待が大きい。

以前の概念と比べると、「データ」という史料の種類よりも「デジタル」という方法の違いに注目を移した点が異なる。そしてデジタル技術の利点を活かす観点から、ボトムアップ的な推論過程を記録するエビデンスネットワークというモデルを考案し、このモデルを核とした史料批判の体系化を進めるとの構想に至った。

こうした批判的な読みを支えるのがエビデンスという概念であり、「エビデンススペース」という科学の一般原理である。科学研究の基本は、証拠と仮説の積み上げで形成される「エビデンスの網の目」から事実や理論を見出すという推論過程にあり、そのレベルでは自然

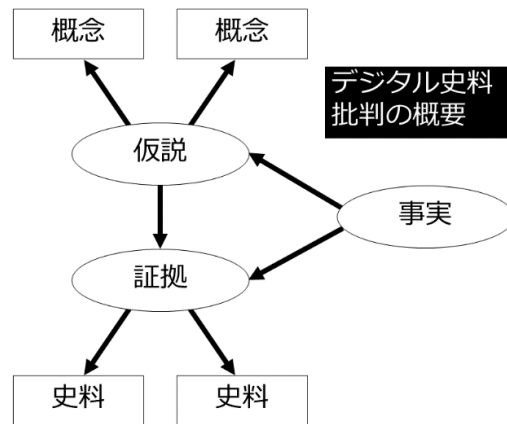


図1 エビデンスネットワークの模式図

科学と人文科学に大きな違いはない。しかし、自然現象の客観的な観測の分析から始まる自然科学とは異なり、人文科学では人間が作り出す文字・非文字の記録の分析から始まるという点に根本的な違いがある。そのため、自然科学では数式一つで幅広い分野をカバーするトップダウン的なアプローチが有効な場合もあるが、人文科学では記録を丹念に解釈してエビデンスの網の目に埋め込んでいくボトムアップ的なアプローチが不可欠となる。すなわちエビデンスベース人文学を実現するには、ボトムアップにエビデンスを積み重ねていける研究基盤を構築する。

なお誤解を避けるため、人文情報学で典型的な計量的アプローチに基づく人文科学と、本論文が提唱するデジタル史料批判との違いを強調しておきたい。計量人文科学は、史資料をデータとみなして史資料 (群) から導かれる数字を分析することを目的とするが、デジタル史料批判は、史資料を解釈した結果をエビデンスとして論理的にモデル化することを目的とする。例えば、テキスト中の語の出現頻度や特徴的な語などを数値的に計算するのが計量人文科学、一方史資料に含まれる誤りや偏りを検証したエビデンスを論理的に結合していくのがデジタル史料批判である。

### 3. 研究の方法

史料や概念に関する証拠や仮説を連結したエビデンスネットワークという理論的なモデルを構築し、これを Semantic Web 技術を用いて実装するというのが本論文の目標である。そして「デジタル・シルクロード」プロジェクトにおいて長年取り組んできたシルクロード遺跡データベースなどをケーススタディとして、デジタル史料批判を深化させ一般化することを基本的な方針とする。

本研究では、エビデンスベースの史料批判について、証拠、仮説、事実という3つの概念を定義し、さらに確信度という属性を導入する。

- **証拠** 史料 (データ) の間の関係を指す。例えば史料 A の一部 B と史料 C の一部 D が意味的に関連している等の知見を

証拠と呼ぶ。

- **仮説** 歴史概念の間の関係を指す。例えば遺跡概念 E と遺跡概念 F には関係がある等の知見を仮説と呼ぶ。
- **事実** 今後の研究の基礎となるような確たる関係を指す。複数の証拠と仮説をエビデンスとして明示的に登録するが、研究者の考え方によってその判断は異なる場合もある。
- **確信度** 人文科学では客観的な観測に基づく確かな証拠や仮説を得ることが難しく、証拠や仮説には史資料に起因する制約が残る場合もある。そこで証拠や仮説に確信度という属性を設け、証拠や仮説の確からしさを数字やカテゴリで表現する。

図1は、証拠、仮説、事実をノードとするエビデンスネットワークの模式図を示す。そして以下のような研究の推論過程を通して、エビデンスネットワークの構築が進むと想定する。

- (1) 様々な史料に対する証拠と様々な概念に対する仮説を結合したエビデンスネットワークを成長させる。
- (2) エビデンスネットワーク上で矛盾する証拠や仮説を見つけ出す。
- (3) 内容の再検討を通して証拠や仮説のエビデンスを洗練させ、確信度を増加させる。
- (4) 今後の研究の基礎となる事実を確定させる。

さらに確信度が高い証拠や仮説だけを選択して史料批判を行うなど、動的な推論過程を実現できることもデジタル環境ならではの利点である。このような推論過程を支える研究基盤とツールを、多数のケーススタディを通して洗練させることが、デジタル史料批判の重要な課題である。

こうしたエビデンスネットワークを構築するために、本研究ではデジタル史料批判プラットフォーム DCP を構築する。これは図2に示す3つのコンポーネントから構成されるものとする。

1. **データリポジトリ** デジタル史料批判の素材となる史料(データ)や、仮説の



図2 デジタル史料批判プラットフォーム(DCP)の模式図。



図3 Murtuk Ruins M.B.I (左) 烏江不拉克仏塔 (右) を照合した例。

基礎となる概念、そして史料や概念に対するエビデンスを蓄積する。データはテキストに限らず画像や地図なども含み、これらに一意的識別子を与えた上でメタデータを付与し、証拠ツールや探求ツールからの利用を可能とする。

2. **証拠ツール** 史料(データ)を解釈した結果を証拠として登録するツールである。例えば2枚の写真に撮影されている遺構が同一のものであるという解釈や、2枚の地図に記載されている別名の遺跡が同一のものであるという解釈などを証拠として登録する。史料タイプの特性に合わせてツールをデザインし、解釈の結果だけでなく、状況や根拠もデータリポジトリに登録できるようにする。
3. **探求ツール** その時点で得られているエビデンスネットワークを分析し、そこから導ける仮説や事実を登録するツールである。エビデンスネットワーク上に積み上げた証拠や仮説の信頼度と、それらの整合性を考察しながら、エビデンスネットワークを洗練させていく。そのためには、エビデンスネットワークを自在に操作する機能が必要となる。

この3つのコンポーネントについて、それぞれツールを開発する。例えば証拠ツールについては、以下の3つのツールを開発した。

- (1) **マッピング** 地図を対象とした空間的な照合
  - (2) **フォトフィット** 写真を対象とした視覚的な照合
  - (3) **メモリーハンティング** 写真と実世界を対象とした空間的かつ視覚的な照合
- その詳細については別の論文に譲る。

#### 4. 研究成果

筆者らがすでに出版した論文の事例を対象として、エビデンスネットワークの構築方法を例示する。これらのエビデンスネットワークは、実際に研究が進行した過程に対応するというよりは、様々な試行錯誤を経て得られた研究成果を、エビデンスネットワークというツールを使って再現したものである。また本論文では史料批判の中身については議論せず、史料批判の結果として得られるエビデンスネットワークに焦点を合わせる。

まず、Murtuk Ruins と烏江不拉克仏塔の例を取り上げる。図3は Murtuk Ruins と烏江不拉克仏塔を撮影した写真を示す。ここでの出発点は、ある古写真中の「Murtuk Ruins」という名前の遺跡と、ある現代写真中の「烏江不拉克仏塔」という遺跡が同じものではないかという推定である。写真と遺跡名は写真のキャプションを通して結ばれているが、遺跡名には綴りの類似性がないため、名寄せだけで同一性を判定することは不可能である。しかし、古地図を使った場寄せを進める過程で、これらの遺跡の同一性に関する探究を進めることになった。そこでフォト



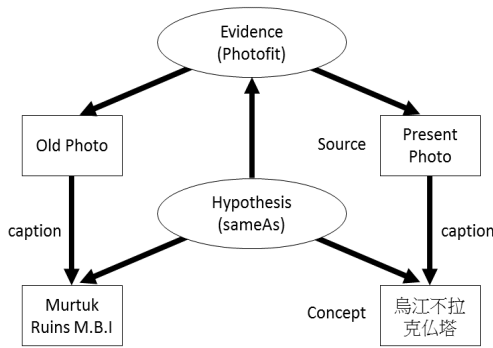


図 4 Murtuk Ruins M.B.I と烏江不拉克仏塔に関するエビデンスネットワーク。

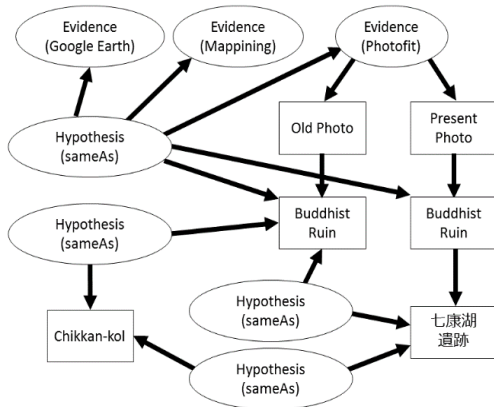


図 5 Chikkan-köl と七康湖遺跡に関するエビデンスネットワーク。

フィットを使って 2 枚の写真を比較すると、写真を平行移動・拡大縮小することで、写真中の遺跡形状の対応付けが可能であることが判明した。このように写真に対する史料批判を通して、2 つの遺跡の同一性に関する仮説を提示することができた。

以上の推論過程をエビデンスネットワークとして表現したのが図 4 である。四角ノードはソースノード、楕円ノードはエビデンスノードであり、それらを有向グラフで結んだものがエビデンスネットワークとなる。まずは写真のキャプションを用いて、史料と遺跡をリンクする。次にフォトフィットを用いて写真の重ね合わせを行い、史料上で的一致という証拠をエビデンスノードとして登録する。そして、フォトフィットで得られた証拠を基に、2 つの遺跡が同一のもの (sameAs) であるという仮説をエビデンスノードとして登録する。これにより、この研究者が 2 枚の写真の照合という解釈から導いたエビデンスネットワークが完成した。今後の現地調査などでさらにエビデンスを蓄積できれば、この 2 つの遺跡の同一性を事実と認定し、今後の研究に引用していくことが可能になるかもしれない。

次に、Chikkan-köl と七康湖遺跡の例を、より複雑なエビデンスネットワークの例として、そして複数の史料を照合して証拠を積み上げ、それらの証拠を基にして遺跡の同一

PREFIX dcterms: <<http://purl.org/dc/terms/>>

PREFIX dcp: <<http://dsr.nii.ac.jp/dcp/#>>

```
SELECT ?instance_a ?Relation ?instance_b ?CF
```

```
WHERE {
```

```
?Hypothesis
```

```
dcp:certaintyFactor ?CF ;
```

```
dcp:subjectUri ?a ;
```

```
dcp:objectUri ?b ;
```

(a) SPARQL を用いたクエリ。確信度の高いエビデンス (CF>80) のみを検索した。

Instance_a	Relation	Instance_b	CF
"仏教寺院遺構 (hotscho)"	< <a href="http://schema.org/sameAs">http://schema.org/sameAs</a> >	"Chikkan-köl"	85
"仏教寺院遺構 (hotscho)"	< <a href="http://schema.org/sameAs">http://schema.org/sameAs</a> >	"仏教寺院遺構 (現代現地調査)"	90
"Chikkan-köl"	< <a href="http://schema.org/sameAs">http://schema.org/sameAs</a> >	"七康湖遺跡"	90

(b) SPARQL 検索結果。今後はこれをグラフとして可視化する計画である。

図 6 エビデンスネットワークの SPARQL 検索例。

性に関する仮説を立てる例として取り上げる。

出発点は、ある仏教寺院遺構と別の仏教寺院遺構とが同一のものではないかという直観である。そこで、マッピングを用いた地図上で「場寄せ」の証拠、フォトフィットを用いた写真照合の証拠、そして Google Earth を用いた景観照合の証拠などを次々と登録することで、エビデンスネットワークを成長させていく。一方これとは別に、仏教寺院遺跡と七康湖遺跡との同一性に関する証拠が得られており、さらにこの仏教寺院遺跡と Chikkan-köl として知られている場所の同一性に関する証拠が得られている。

そこでこれらがお互いに同一であるという仮説を立ててエビデンスネットワークに登録して図 5 のネットワークを得た。これにより、3 つの遺跡と地点が同一であるという仮説が得られ、エビデンスネットワークを再検討しても個別の証拠や仮説に問題はなさそうであると判断できたことから、七康湖遺跡と Chikkan-köl は同一の地点であるという、遺跡や地点に関する未知の関係性を仮説として提示することができた。

最後に、このエビデンスネットワークを対象とした SPARQL 検索の例を図 6 に示す。これは確信度が高いエビデンスのみを検索するというクエリであるが、これにより検討対象とするエビデンスを動的に選択できるため、エビデンスネットワークの分析や洗練化をより効率的に行えるようになると考える。

なお図 4 や図 5 で示したエビデンスネットワークは、研究の再現として構築した仮想的

なものであり、実際の研究をこのツールで推進した例はまだ存在しない。今後は多数のケーススタディにおいて DCP を実利用し、研究者のフィードバックに基づき DCP を改良する計画である。

#### 5. 主な発表論文等

(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4件)

西村 陽子, 北本 朝展, "絲綢之路遺址之重新定位与遺址数据库之建立", 陝西師範大学学報(哲学社会科学版), 査読無, Vol.45, No.2, 2016, 76-86

西村 陽子, 北本 朝展, "高昌故城調査の統合による探検隊調査遺構の同定—地図史料批判に基づく都市遺跡・高昌の復原—", 高田時雄教授退職記念学術論文集, 査読有, 2014, 181-196

西村 陽子, 北本 朝展, "地図史料批判に基づくシルクロード都市遺跡・高昌故城の遺構同定", 人文科学とコンピュータシンポジウム じんもんこん 2014, 査読有, 2014, 43-50

西村 陽子, Erika Forte, 北本 朝展, 張 勇, "古代城市遺址高昌的遺構比定: 基于地圖史料批判的絲綢之路探險隊考察報告整合", 西域文史, 査読有, Vol. 9, 2014, 153-197

[学会発表](計7件)

北本 朝展, 西村 陽子, "Digital Criticism Platform: エビデンススペースの解釈を支援するデジタル史料批判プラットフォーム", 人文科学とコンピュータシンポジウム じんもんこん 2015, 2015年12月20日, 同志社大学京田辺校地(京都府京田辺市)

Asanobu KITAMOTO, "Memory Hunt: A Mobile App with an Active Viewfinder for Crowdsourced Annotation through the Re-experience of the Photographer" (in English), Fifth Annual Conference of the Japanese Association for Digital Humanities (JADH2015) (国際学会), 2015年9月2日, 京都大学人文科学研究所(京都府京都市)

北本 朝展, "古写真の撮影者になりきるアプリ メモリーハンティング(メモハン)", ジオメディアサミット大阪 2015 (招待講演), 2015年09月4日, グランフロント大阪(大阪府大阪市北区)

西村 陽子, 北本 朝展, "絲綢之路遺址の重

新定位與遺址数据库之建立", (絲綢之路研究與歷史地理資訊系統建設 國際學術研討會招待講演), 2015年08月22日, 陝西師範大学(中華人民共和國陝西省西安市)

Yoko NISHIMURA, Erika FORTE, "A new method for re-identifying ancient excavated structures on the Silk Road: the case of Qočo", Medieval Pre-Islamic Architecture in Qočo: Architectural, Archaeological, Art Historical and Scientific Evaluation of a Collection of Wooden Architectural Elements in the Museum für Asiatische Kunst(AKu) (招待講演)(国際学会) - a project supported by the Gerda Henkel Stiftung, 2015年5月28日, Museum für Asiatische Kunst(AKu) (ベルリン国立博物館アジア美術館)(ドイツ連邦共和国ベルリン市)

Asanobu KITAMOTO, Yoko NISHIMURA, "Data Criticism: General Framework for the Quantitative Interpretation of Non-Textual Sources"(in English), 2014年7月10日, Digital Humanities 2014, (スイス連邦ローザンヌ市)

Asanobu KITAMOTO, Yoko NISHIMURA, "Geohumanities and the Digital Silk Road" (in English), GeoHumanities SIG Meeting, 2014年7月8日, (スイス連邦ローザンヌ市)

[図書](計1件)

西村 陽子, Erika Forte, 北本 朝展, 張 勇, 文物出版社, "高昌故城遺址諸遺跡の比定—基于地圖史料批判的絲綢之路探險隊考察報告整合", 高昌故城及其周辺地区の考古工作報告: 一九〇二~一九〇三年冬季, 新疆考古研究所・吐魯番学研究院(編), 2015, 197-245

[その他]

ホームページ  
<http://dsr.nii.ac.jp/digital-maps/>  
<http://dsr.nii.ac.jp/>

#### 6. 研究組織

##### (1) 研究代表者

北本 朝展 (KITAMOTO, Asanobu)  
国立情報学研究所・コンテンツ科学研究系・准教授

研究者番号: 00300707

##### (2) 研究分担者

西村 陽子 (NISHIMURA, Yoko)  
花園大学・文学部・講師  
研究者番号: 70455195