

# 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

国立情報学研究所／総合研究大学院大学  
北本 朝展 (KITAMOTO Asanobu)

<http://agora.ex.nii.ac.jp/~kitamoto/>

共同研究者：  
西村 陽子（花園大学）

# 本研究の目標

1. デジタル史料批判とエビデンスネットワーク（EN）の概念を明確化する。
2. ENを成長させる研究基盤「Digital Criticism Platform（DCP）」を構築する。
3. Semantic Web技術の一つであるRDFを用いてENの管理および検索を実装する。
4. 以上により、「エビデンスベース人文学」への道を拓く。

# デジタル史料批判とは？

# 人文学におけるCriticism

人文学の多くの分野において「Criticism / Critique」は基本的な概念である。

## データ史料批判 (Data Criticism)

精細だが正確ではない空間画像史料（地図・写真等）を批判的に読み解く方法。

## デジタル史料批判 (Digital Criticism)

デジタル技術を活用した史料批判の一般論。

# エビデンスベース人文学

**自然科学でも人文科学でも**、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推論過程が重要。



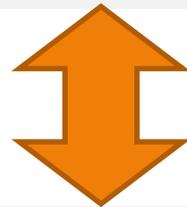
**人文科学では**、断片的な証拠を積み上げるボトムアップ的なアプローチの重要性が増加。



**エビデンスベース人文学では**、様々なエビデンスを蓄積・活用・共有する研究基盤が必要。

# デジタル史料批判

**デジタル史料批判**：エビデンスの論理的なモデル化を目的とし、「史料批判」という歴史学の基本概念の観点から、個別の技術的な方法論を評価・統合するもの。



**計量人文学**：史資料をデータとみなし、出現頻度などを数値的に計算するもの。定量的ではあるが、批判的な視点はやや薄い。

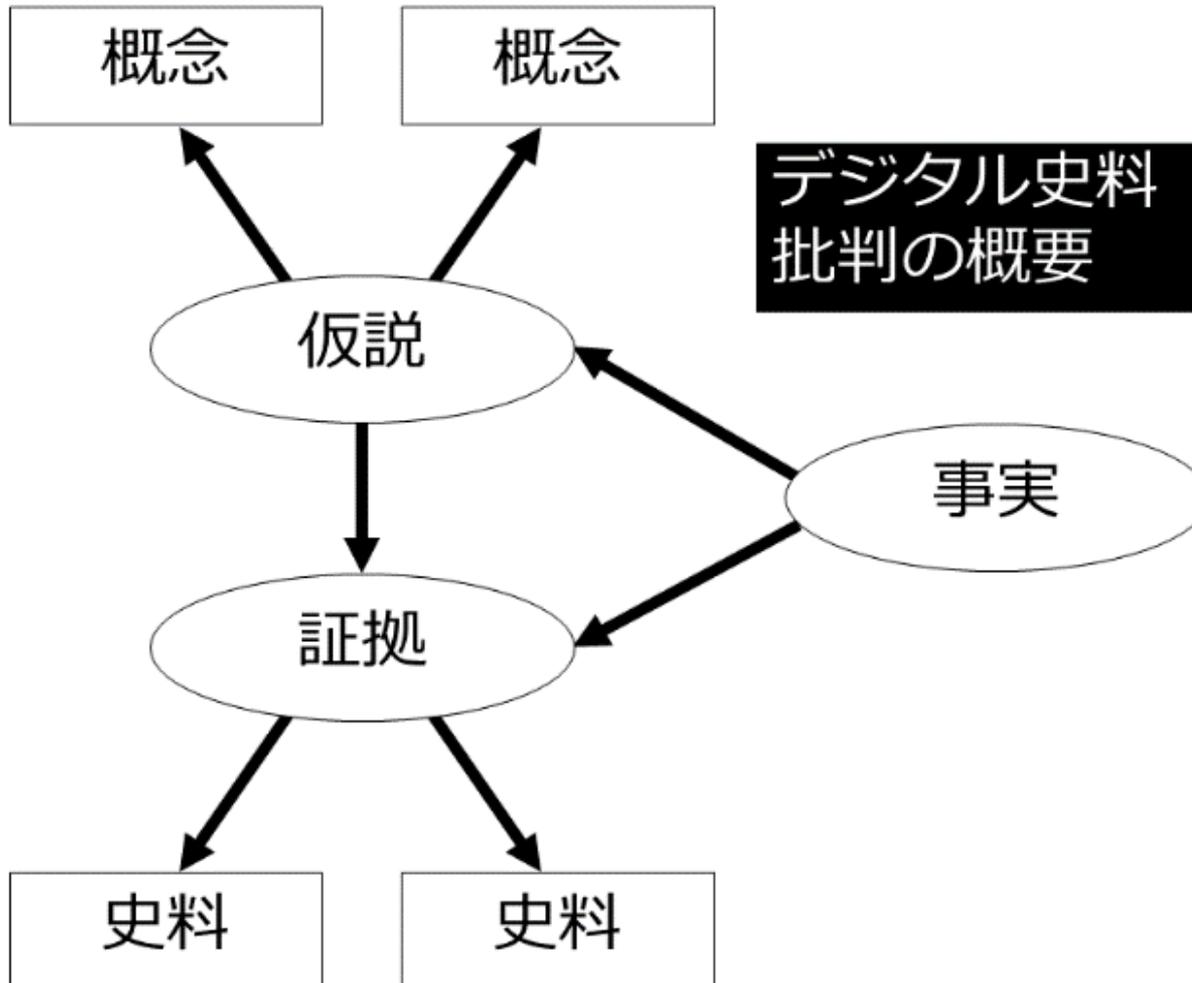
# デジタル史料批判プラットフォーム フォーラム（DCP）

# デジタル史料批判の構成要素

証拠	史料（データ）の間を指す。
仮説	歴史概念の間を指す。
事実	今後の研究の基礎となるような確たる関係を指す。
確信度	証拠や仮説の確かさを数字やカテゴリで表現する。

Question: これらの構成要素の間に成立する関係をどうモデル化するか？

# エビデンスネットワーク (EN)



- 証拠
- 仮説
- 事実
- 3種類の概念 (+ 確信度属性) を結ぶ有向グラフ。

# ENを用いた研究方法

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 (EN)** を**成長**させる。
2. エビデンスネットワークを吟味し、**矛盾する証拠や仮説**を見つけ出す。
3. 改善の余地あるエビデンスを再検討・洗練し、**証拠や仮説の確信度を増加**させる。
4. **研究の基礎となる事実**を**確定**させる。

# DCPの3つのコンポーネント



- **データリポジトリ**は、エビデンスネットワークを蓄積・管理する役割。
- **証拠ツール**と**探求ツール**は、エビデンスネットワークを成長させる役割。

# DCPのコンポーネント

# データリポジトリ

1. 史料や概念に**識別子**と**メタデータ**を付与。
2. エビデンスネットワークを**意味的**に構築。
3. 証拠ツール・探求ツールと連携し、証拠・仮説・事実に**識別子**と**メタデータ**を付与して登録。

豊富な機能を有し、長年の実績があり、Linked Data拡張でSPARQL検索にも対応できる、**DSpace**を基盤システムに採用した。

# 証拠ツール

- 1. マッピング** 地図を対象とした空間的な照合（ウェブアプリ）
- 2. フォトフィット** 写真を対象とした視覚的な照合（ウェブアプリ）
- 3. メモリーハンティング（メモハン）** 写真と実世界を対象とした空間的かつ視覚的な照合（モバイルアプリ）

# マッピング (Mapping)



複数の地図の照合を支援するウェブアプリ。  
位置ずれが大きな地図を重ね合わせて地点  
の対応関係を探索する。

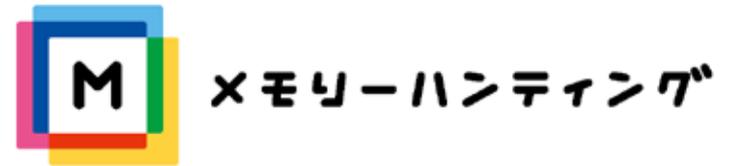
<http://dsr.nii.ac.jp/digital-maps/mapping/>

# フォトフィット (Photofit)



複数の写真の照合を支援するウェブアプリ。  
異なる時期に撮影された写真を重ね合わせて  
物体の対応関係を探索する。

# メモリーハンティング (MemoryHunt)



画像史料と実世界の照合を支援するモバイルアプリ。画像史料が紐づく実世界の位置と方向を探索する。

<http://dsr.nii.ac.jp/memory-hunting/>

# 証拠ツールの一般化

- **より多くの史料タイプに関する証拠**：ある史料の出現文字列と別の史料の出現文字列が、同じものを指すという関係。
- **異なる史料タイプにまたがる証拠**：「高い塔がある」というテキストと高い塔の写真が、同じものを指すという関係。
- **エビデンスのフォーマットを共通化**することで、統一的な扱いを可能とする。

# 探求ツール

- 証拠ツールを用いて成長させたエビデンスネットワークを対象に、そこから導ける**仮説**や**事実**を登録するツール。
- Semantic Web技術の一つである**RDF** (Resource Description Framework) を用いてエビデンスネットワークを構築。
- **SPARQL**を用いた意味検索を活用（例えば確信度が高いエビデンスのみを抽出）。

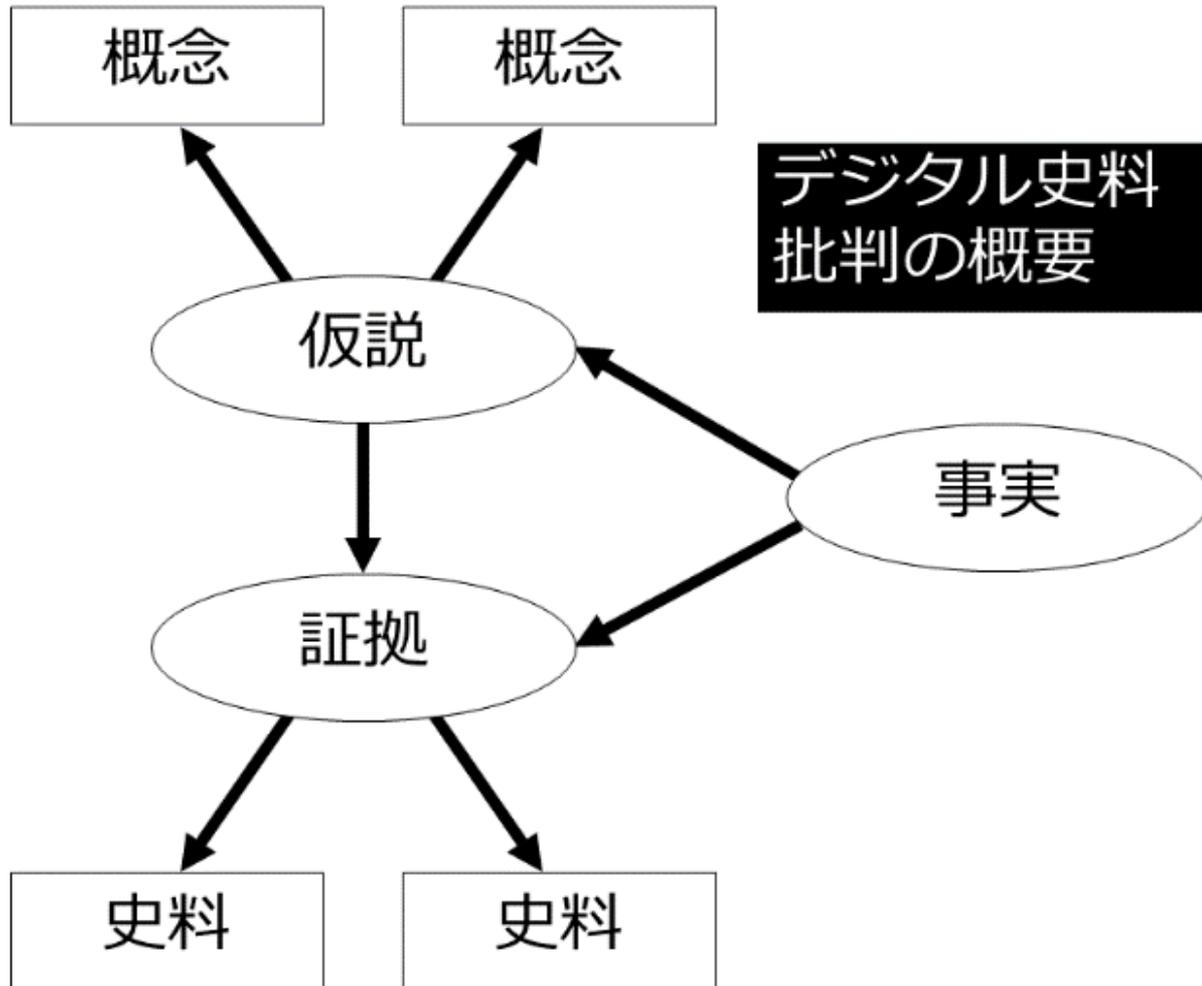
# コンポーネントの接続



- **証拠ツール**は史料等の識別子を用いて、データリポジトリからデータを読み出す。
- 照合結果をメタデータに記述してエビデンスとし、**データリポジトリ**に登録する。

# エビデンスネットワーク

# エビデンスネットワーク



- 証拠
- 仮説
- 事実
- 四角はDCPに与えるノード、楕円はDCPで作るノード。

# ノードの種類

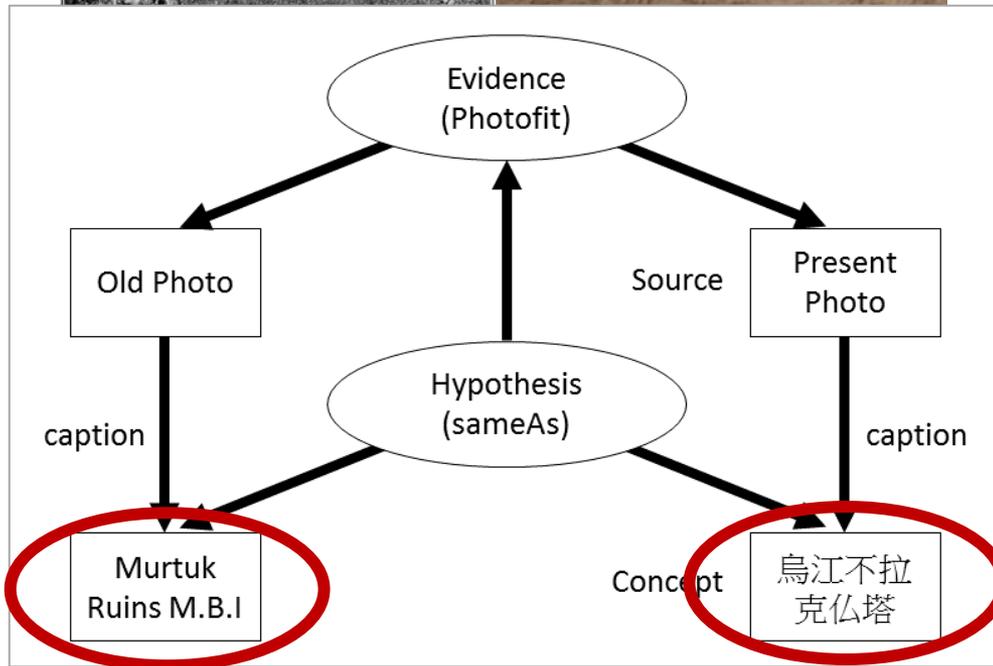
ソースノード	デジタル史料に対応するノードであり、識別子を用いて特定できる。
インスタンスノード	歴史を記述する概念単位に対応するノードであり、識別子を用いて特定できる。
エビデンスノード	証拠および仮説に対応するノードであり、それぞれ確信度という属性を用いて確からしさを表現する。
事実ノード	確実に成立する関係性や、研究者のコンセンサスが得られる関係性を選び出して登録する。

# ケーススタディ

- 過去の研究例（下記）を対象に、もしENがあったらどのような推論過程が可能であったか、という再現実験をおこなう。
- どの証拠を使って仮説をたて、事実に至ったか、といった研究プロセスを明確化することの効果を検証する。

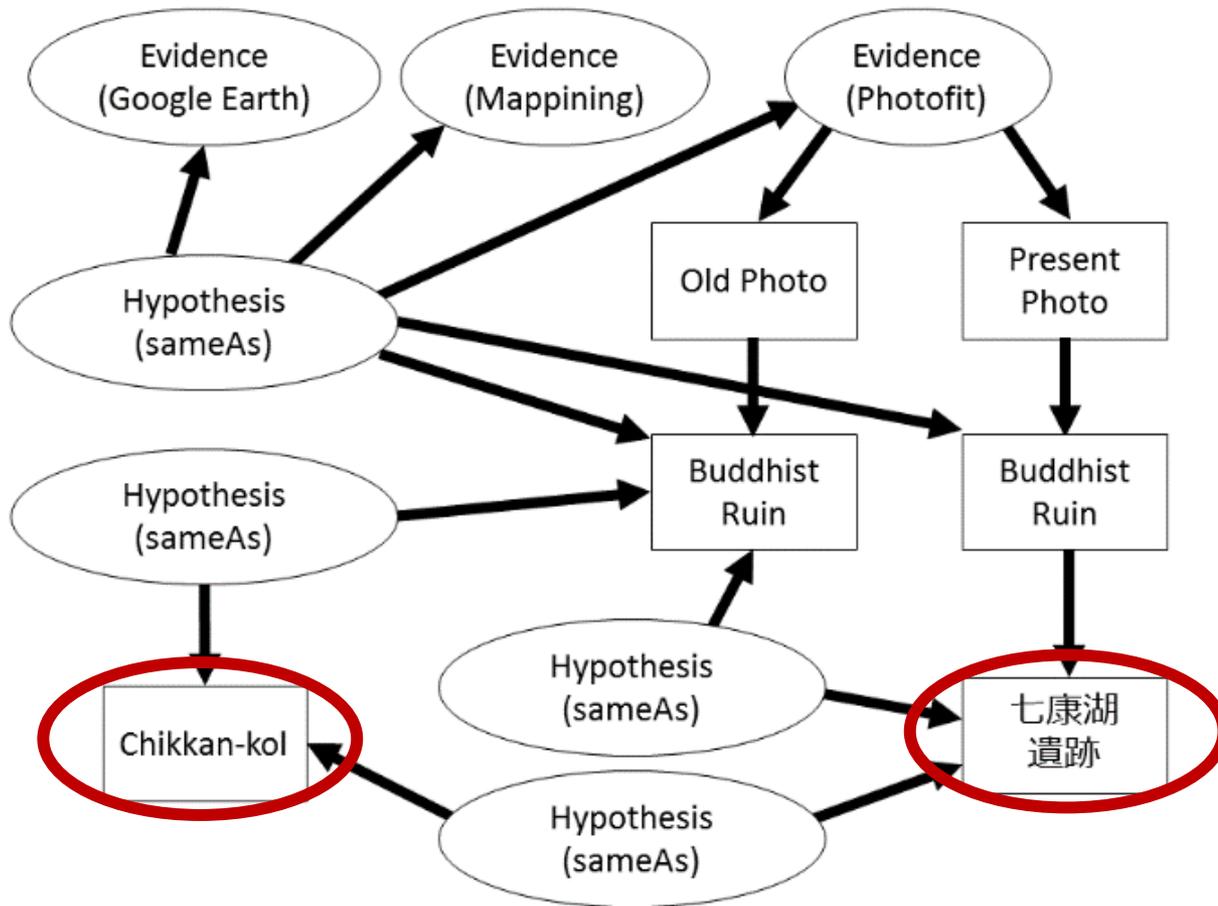
西村 陽子, 北本 朝展, "スタイン地図とGoogle Earthを用いた名寄せと場寄せに基づくシルクロード探検隊調査遺跡の解明", 人文科学とコンピュータシンポジウム じんもんこん2010, pp. 255-262, 2010.

# Murtuk Ruinsと烏江不拉克仏塔



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」という遺跡は、同じものではないか？

# Chikkan-kölと七康湖遺跡



ある仏教寺院遺構と別の仏教寺院遺構とは、同じものではないか？

# SPARQL検索の活用

```
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 ;
  dcp:relation ?Relation .
  ?a dcterms:title ?instance_a .
  ?b dcterms:title ?instance_b .
  FILTER (?CF > 80)
}
```

instance_a	Relation	instance_b	CF
"仏教寺院遺跡 (Chotscho)"	<http://schema.org/sameAs>	"Chikkan-kol"	85
"仏教寺院遺構 (Chotscho)"	<http://schema.org/sameAs>	"仏教寺院遺構 (現代現地調査)"	90
"Chikkan-kol"	<http://schema.org/sameAs>	"七康湖遺跡"	90

- 確信度が高いエビデンスのみを検索するクエリを発行。
- ENの部分グラフを動的に選択し、限定したビューの枠内で再検討。
- ENを多角的に吟味することで、効率的に洗練化。

# DCPとオープンサイエンス

# DCPの目標

デジタル技術を梃子にして研究の方法論をバージョンアップし、歴史学の可能性を広げること。

- 証拠と仮説を積み上げて事実を導く研究の推論過程は、既存の方法と大差ない。
- 研究者にとって違和感のないワークフローを用いて、無理のない導入を進める。

# 研究のエビデンスの公開

## 利便性

- 研究成果のオープン化を通して、**再利用を促進**する。
- 誰かがエビデンスを再利用し、新たな知見を付け加えることで、**知識の網の目を成長**させる。

## 透明性

- 研究成果のオープン化を通して、**不正を検証可能**とする。
- 誰もがエビデンスを検証可能な形で公開することで、**研究成果の信頼性を向上**させる。

# オープンサイエンスに向けて

- 論文という形式から漏れる「小さなエビデンス」もデジタル出版が可能となる。
- 出版されたエビデンスに対する引用（サイテーション）を通して、エビデンスを公開する価値を高め、細粒度の研究成果を流通させる道を開く。
- オープンな史料と明示的なエビデンスに基づく開かれた人文学へ。

# 本研究の目標

1. デジタル史料批判とエビデンスネットワーク（EN）の概念を明確化する。
2. ENを成長させる研究基盤「Digital Criticism Platform（DCP）」を構築する。
3. Semantic Web技術の一つであるRDFを用いてENの管理および検索を実装する。
4. 以上により、「エビデンスベース人文学」への道を拓く。

# 関連情報

デジタル史料批判プラットフォームの構築には、ソフトウェアエンジニアの池崎友博氏の協力を得た。本研究は、JSPS科研費26540178「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」の助成を受けた。またメモハンの構築の一部には、京都大学地域研究統合情報センター共同研究の助成を受けた。

ディジタル・シルクロード <http://dsr.nii.ac.jp/> では、マッピングとメモリーハンティングを中心に、その他の成果を公開しています。