

Digital Criticism Platform for Evidence-based Digital Humanities with Applications to Historical Studies of Silk Road

National Institute of Informatics / SOKENDAI

Asanobu KITAMOTO

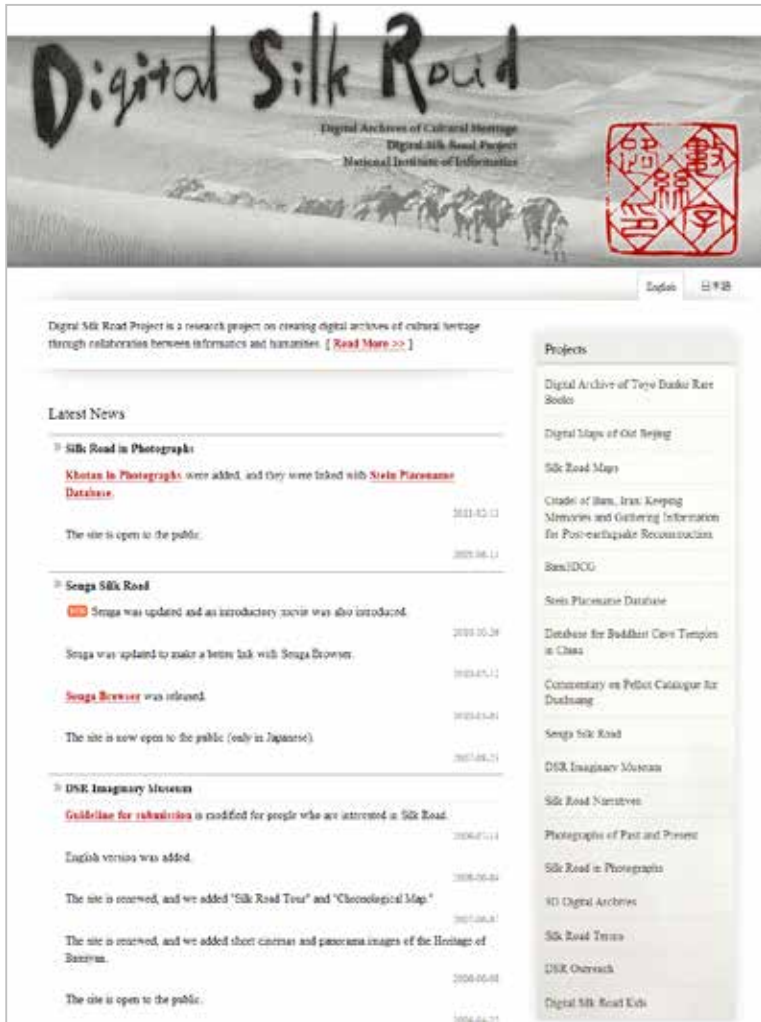
<http://agora.ex.nii.ac.jp/~kitamoto/>

Collaborator:

Yoko Nishimura (Toyo University)

Digital Silk Road

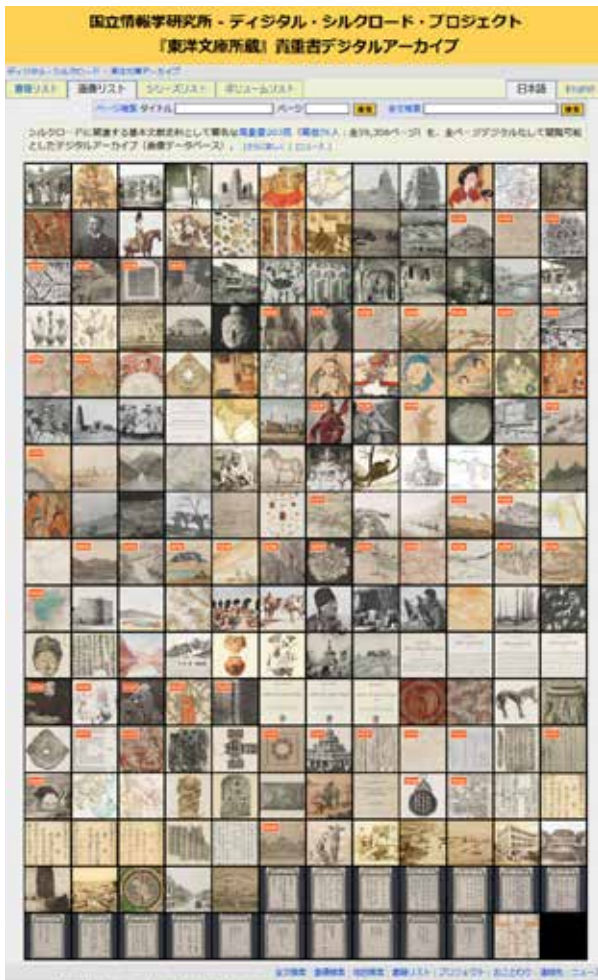
<http://dsr.nii.ac.jp/>



- Started in 2001, and still ongoing.
- Digital humanities project based on collaboration between informatics and humanities.
- Results are publicly available on the Web.

Toyo Bunko Archive of Rare Books

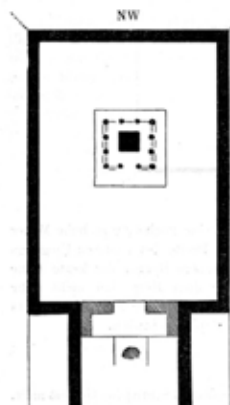
<http://dsr.nii.ac.jp/toyobunko/>



- Digitization of **237** books with **70,898** pages.
- Important references in the research community.
- Caption and TOC are manually transcribed.
- Full text search is based on OCR without correction.

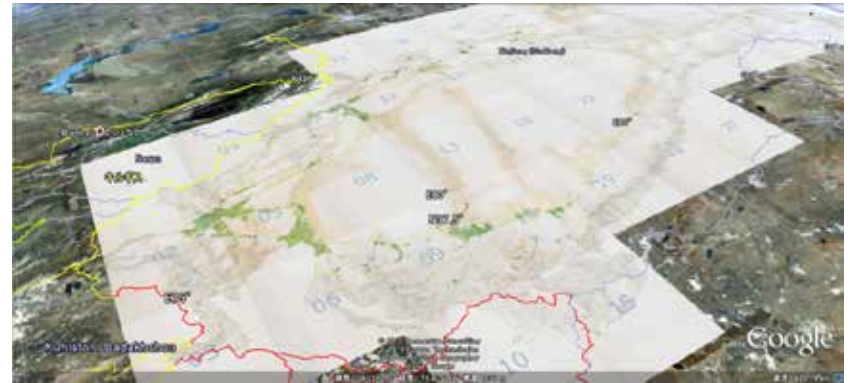
Dealing with Heterogeneous Data

Text



die obere sich wie eine in eine niedrigere 3,10 m tiefe Plattform eingepasste Bank darstellt (auf der Skizze schraffiert) und die Mitte offen liest. Vor dieser großen Unterstufe liegt der Rest eines mächtigen Sockels, in welchem ein tiefes Loch sich zeigt: hier hat also wohl eine große Statue oder eine Fahne gestanden. 12 m nach innen zu vom S-Rand der Plattform des Hauptbaues, 5,50 m von den Seitenmauern und 7 m vor der Rückmauer, erhebt sich eine niedrige, 8 m ins Geviert betragende Stufe, auf deren Mitte ein jetzt zerstörter, 2 m großer, viereckiger Sockel steht; um diesen Sockel geht ein Gang herum, vorne und an den Seiten je 1,50 m breit, hinten aber nur 90 cm breit. Dieser Umgang ist nach außen von einer Mauer umgeben, welche durch zwölf kleine Stäbe in kleine Abteile geteilt ist, von denen der mittlere der Frontseite den Eingang bildet. Auf der Rückseite ist dies aus zwei Eck- und zwei Mittelstützen bestehende System sehr zerstört. Vor den sechs Interkolumnien der Seiten und den zwei Interkolumnien neben dem Eingang sind je noch Sockel für Statuen erhalten: nach mancherlei dekoratives Bei-

Map



Photograph

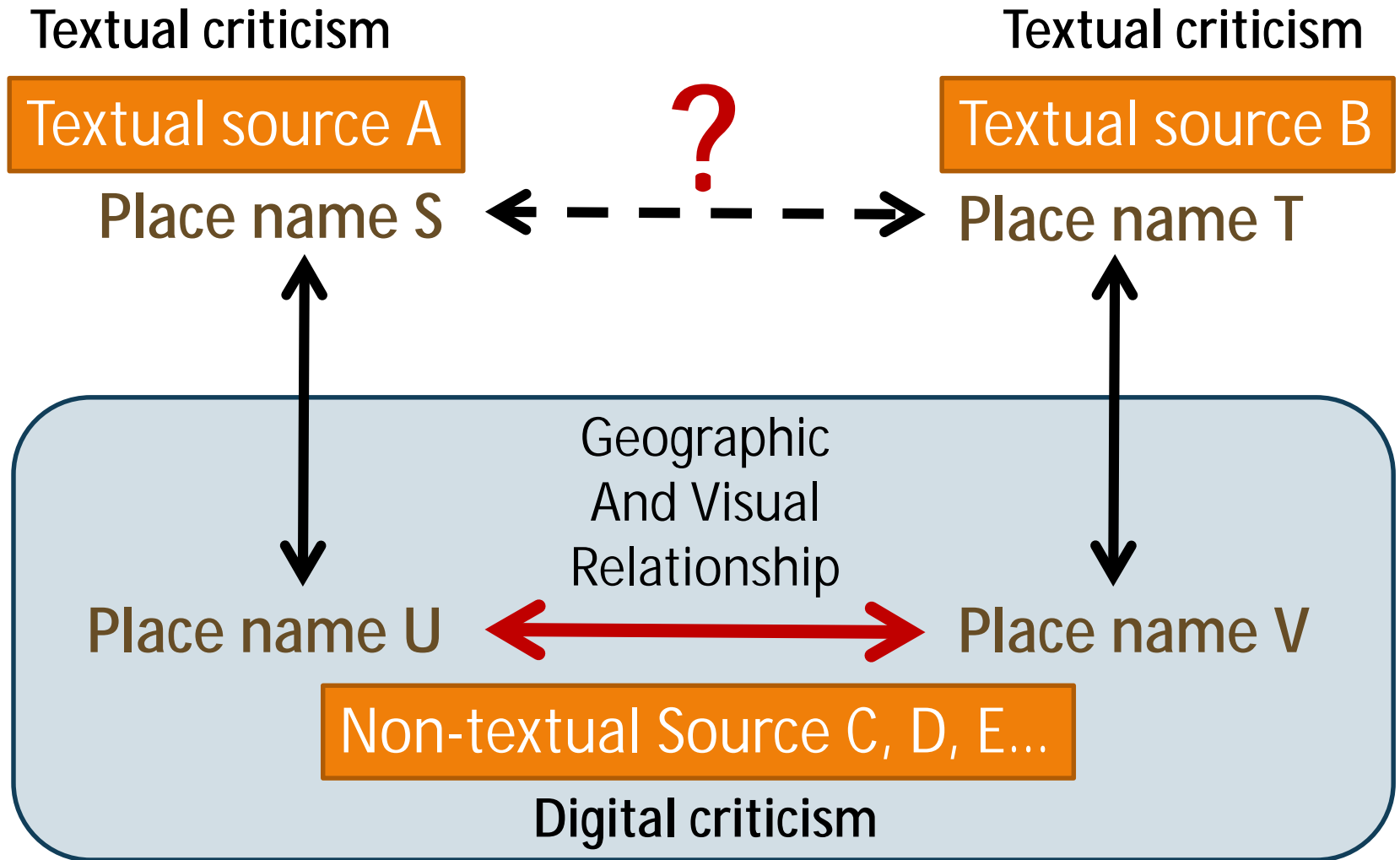


Y
1904(Le Coq, 1913, Tafel. 70, I)

Gazetteer

Abāb-langar, habit., 14. B. 3.	Aechhik-bulak (of Turfān), spring, 28. B. 4.
Ābād (of Ak-su), market-town, 12. A. 3.	Aechhik-bulak (of Yai-döbe), spring, 4. C. 4.
Ābād (of Kara-yulghun), vill., 12. B. 1.	Aechhik-daryā, river, 21. A. 2.
Ābād (of Karghalik), vill., 5. C. 4.	Aechhik-dawān, pass, 9. B. 3.
Ābād (of Kāshgar), vill., 5. A. 2.	Aechhik-jilga (of Duwa), valley, 9. B. 3.
Ābād (of Turfān), vill., 28. C. 3.	Aechhik-jilga (of Kara-tāsh), valley, 2. D. 3.
Ābād (of Yārkand), vill., 5. C. 2.	Aechhik-jilga (of Khotan), valley, 9. C. 3.
Ābād-jilga, valley, 12. B. 2.	Aechhik-jilga (of Sampula), valley, 14. A. 3.
Abdal, vill., 30. B. 2.	Aechhik-jilga (of Tawak-kēl), loc., 14. A. 1.
Abdalkash-mazār, shrine, 14. C. 3.	Aechhik-kōl, lake, 15. D. 1.
Abdul-ghafūr-langar, loc., 10. C. 1.	Aechhik-kuduk (of Kapa), well, 23. A. 1.
Abdul-rahmān-jilga, valley, 9. A. 4.	Aechhik-kuduk (of Kuruk-tāgh), well, 28. C. 4.
Abshak-bēl, Pass, 2. B. 1.	Aechhik-kuduk (of Marāl-bāshi), well, 5. D. 2.
Ach-tāgh, hill and vill., 7. C. 2.	Aechhik-otan, loc., 7. C. 2.
Acha-dong (of Chizghān), hill, 19. C. 3.	Aechhik-su, loc., 31. A. 4.
Acha-dong (of Yārkand R.), loc., 7. D. 4.	Aechhik-tügemen, loc., 5. D. 2.
Acha-kuduk, loc., 7. D. 4.	Achi-tāgh, hill, 32. B. 1.
Acha-shipang, loc., 22. D. 4.	Achik-aghzi, loc., 9. D. 3.
Achak-aghzi, loc., 5. A. 4.	Achma (of Hanguva), vill., 14. A. 2.
Achal (of Ak-su), vill., 12. A. 3.	
Achal (on Charchak R.), loc., 21. C. 2.	

Linking Entities Across Sources



Digital Criticism

- **DH 2014: Data Criticism:** General Framework for the Quantitative Interpretation of Non-Textual Sources
- **DH 2016: Digital Criticism** Platform for Evidence-based Digital Humanities with Applications to Historical Studies of Silk Road
- **Digital criticism** or **digital critique** focuses on the methodology rather than the target.

Digital is more than Quantitative

Quantitative Humanities

- Quantitative analysis
- Statistical tools
- Convert sources to numbers

Digital Criticism

- Critical interpretation
- Semantic tools
- Quantitative and qualitative insights

Evidence-based Humanities

Natural science and humanities share the same reasoning process of finding facts from the network of evidences and hypotheses.

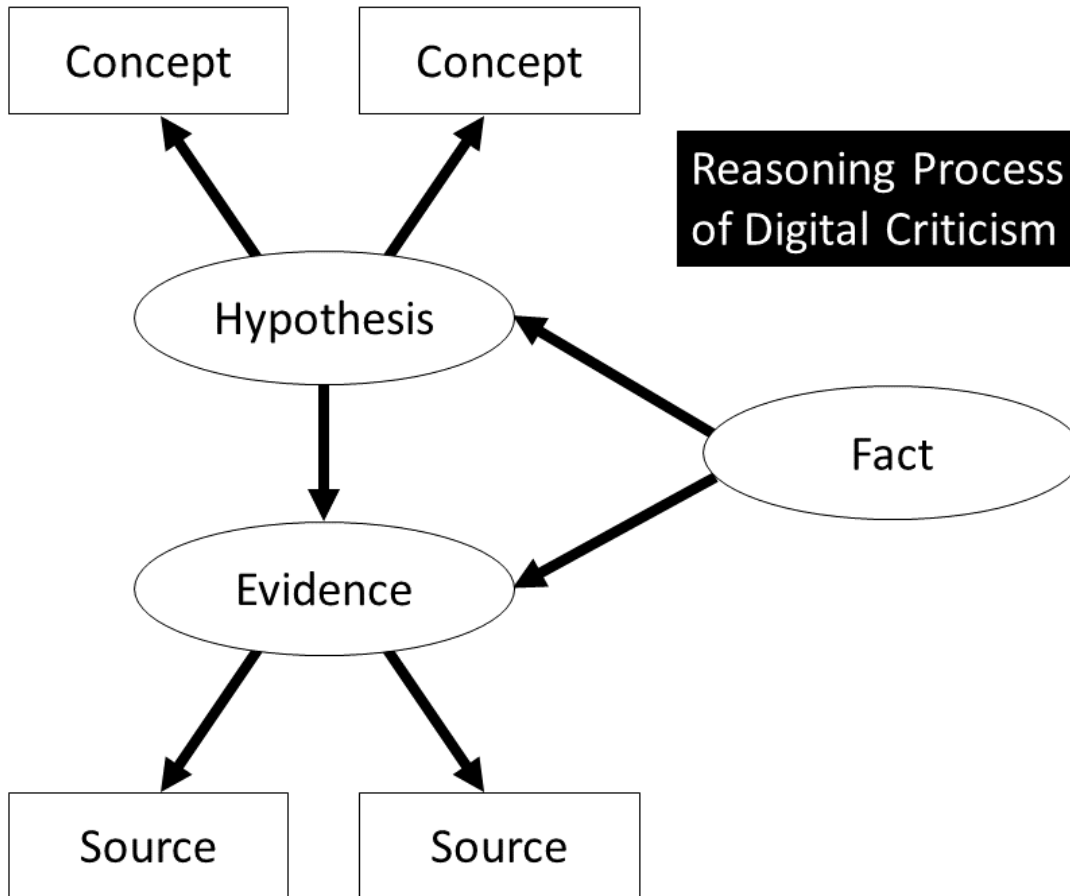


But, humanities prefers a bottom-up approach of accumulating fragmented evidences.



Hence, evidence-based humanities requires its own research infrastructure for evidences.

Evidence Network (EN)



- Evidence
- Hypothesis
- Fact
- Reliability attribute
- Concept
- Source

Basic Elements

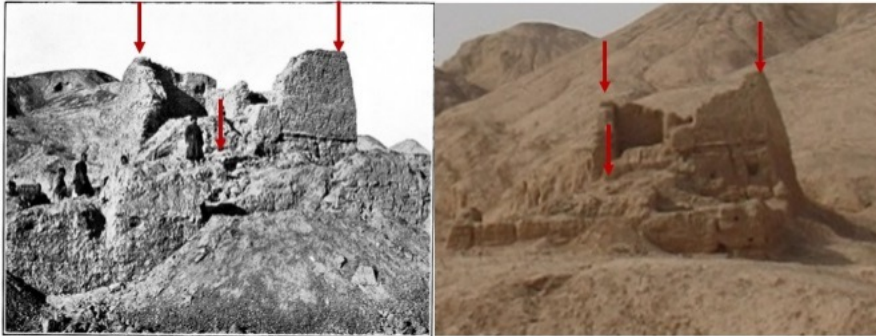
Evidence	Relationship between sources .
Hypothesis	Relationship between concepts .
Fact	Relationship between evidences and hypotheses that claims reusable knowledge .
Reliability	An attribute for an evidence and a hypothesis to represent the degree of reliability defined by a scholar.

Case Studies

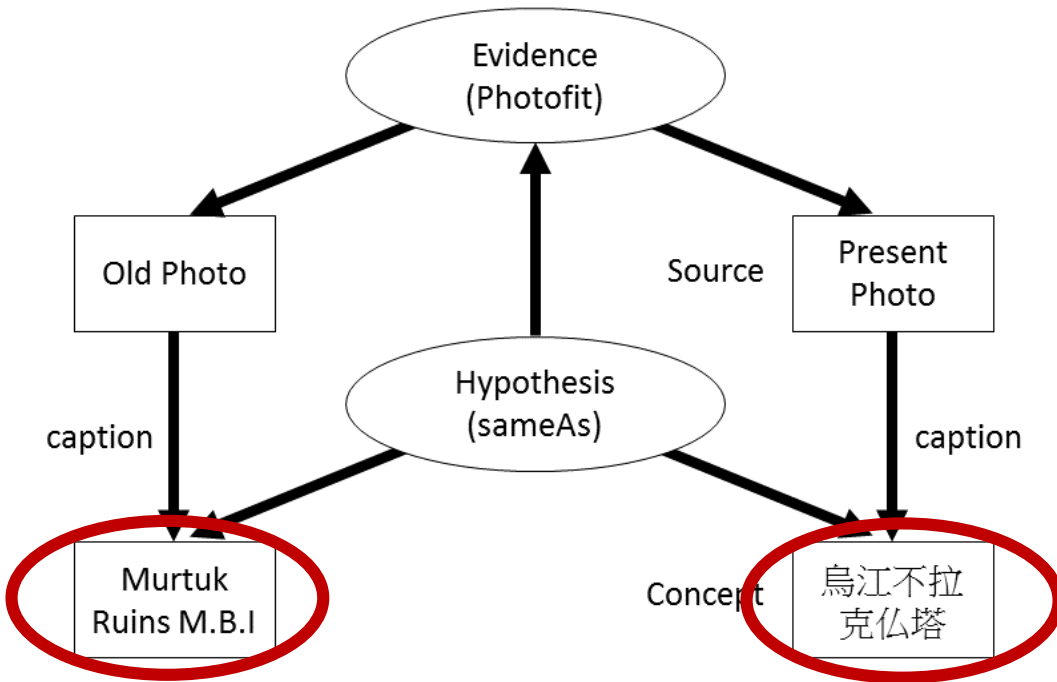
- Our past work [1] was re-examined to check how the reasoning process could be if we had the evidence network at the time of research.
- We verified the benefit of clarifying the bottom-up reasoning process of research using the evidence network.

[1] Yoko NISHIMURA, Asanobu KITAMOTO, "Identification of Ruins Excavated by Silk Road Expeditions through Matching Names and Locations by Stein Maps and Google Earth", IPSJ SIG Computers and the Humanities Symposium 2010, pp. 255-262, 2010-12 (in Japanese)

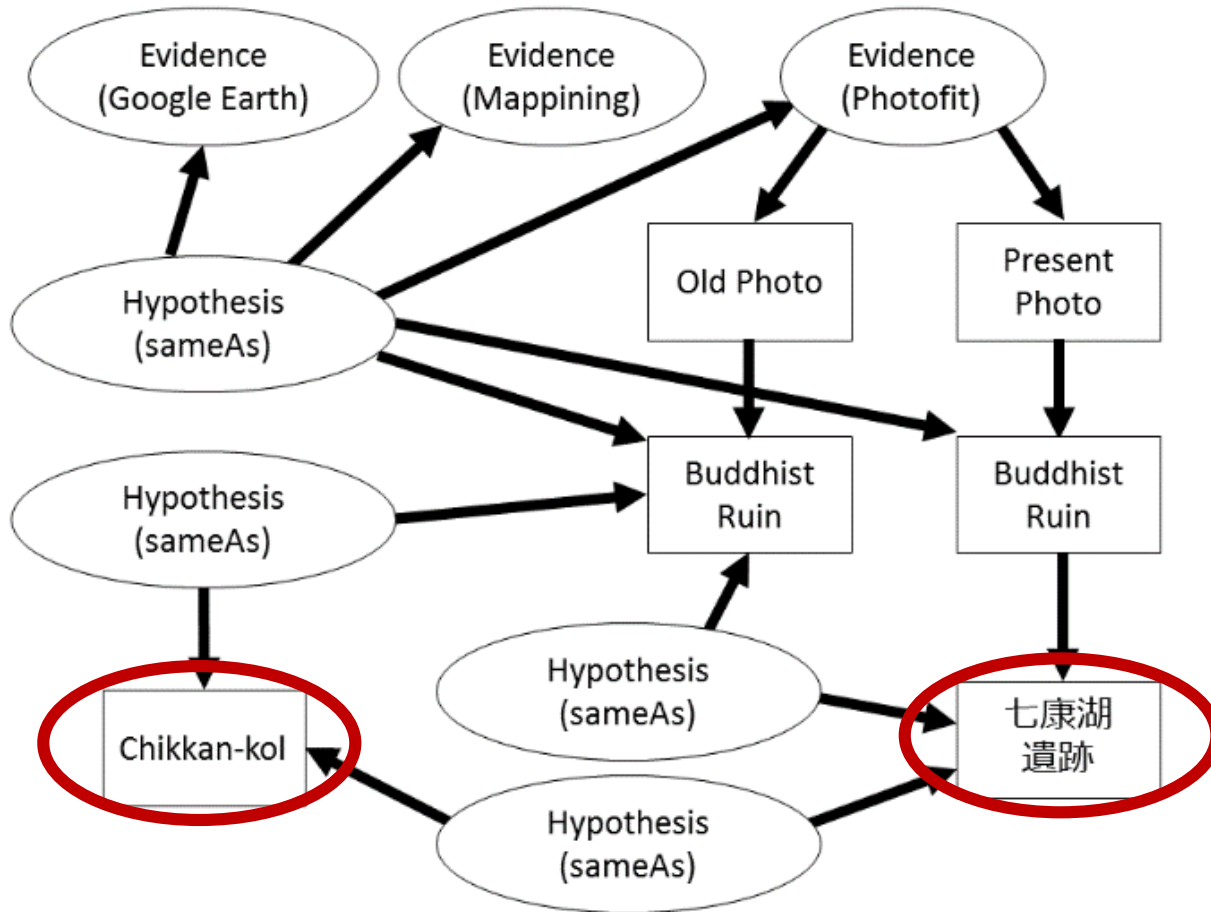
Murtuk Ruins and 烏江不拉克仏塔



Question: A ruin described as “Murtuk Ruins” in an old book, and a ruin described as 「烏江不拉克仏塔」 in a modern book is the same or not?

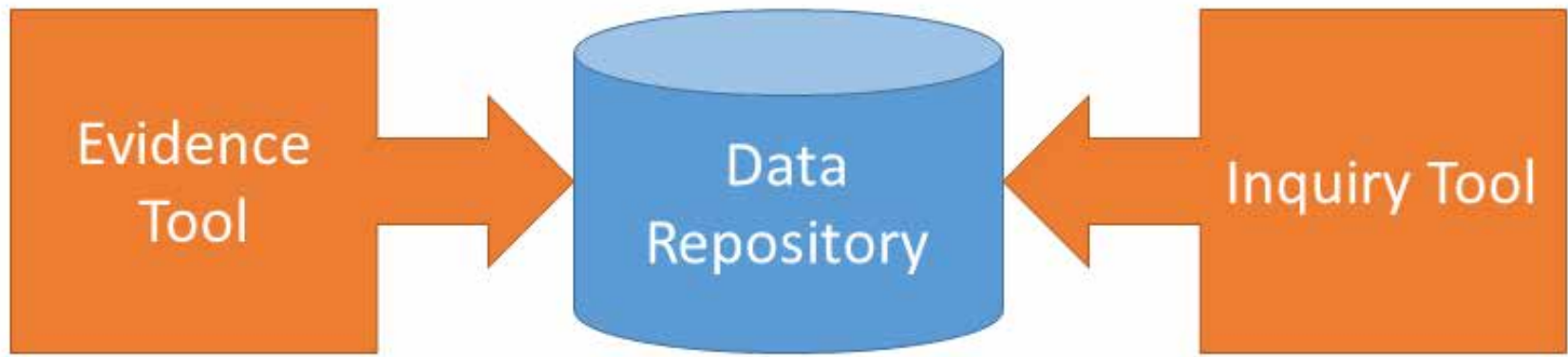


Chikkan-köl and 七康湖遺跡



Question: A Buddhist ruin in one source, and another Buddhist ruin in another source is the same?

Three Components of DCP



- **Data repository** has a role of storing and managing evidence network.
- **Evidence tool** and **inquiry tool** has a role of growing and refining evidence network.

Data Repository

1. **Sources and concepts** : assign an identifier and community-shared metadata.
2. **Evidences and hypotheses** : assign an identifier and per-user metadata.
3. **APIs** : talk with evidence and inquiry tools.

DSpace was selected for open source project that is surviving, and the availability of a SPARQL endpoint within linked data support.

Evidence Tools

Grow the evidence network on semantic relationships.

1. **Mapping** : Spatial matching of maps (Web-based tool).
2. **Photofit** : Visual matching of photographs (Web-based tool).
3. **MemoryHunt** : Spatial and visual matching of photographs and the real world (Mobile app).

Mapping

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



Web-based tool for matching maps based on **interactive georeferencing** in contrast to geometric correction.

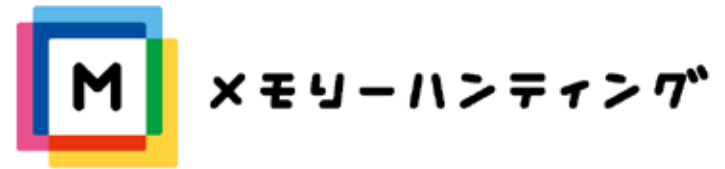
Photofit



Web-based tool for matching photographs. It allows **planar shift and zooming transformation** to find the best match.

MemoryHunt

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



Mobile app for matching a photograph with the real world. **Active viewfinder with controllable transparency** to find the best match.

Inquiry Tools

Refine the evidence network by the inquiry of scholars.

- Review the evidence network to add, modify or delete evidences and hypotheses.
- **RDF (Resource Description Framework) and SPARQL** : semantic search (such as extracting a subgraph of evidences with high reliability) to help refine the network.

SPARQL Endpoint

```
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)
}
```

- SPARQL query is used to filter evidences having higher reliability.
- A subgraph is more convenient for studying a new hypothesis.

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

The Goal of DCP

Using digital technology as a leverage for upgrading the methodology and expand the possibility of historical studies.

- Digital technology is not for automation, but for enhancement, of research workflow.
- The workflow should be similar to the current workflow so that scholars can adapt to it without friction.

More Information

- **Digital Silk Road**
 - <http://dsr.nii.ac.jp/>
- **Toyo Bunko Archive of Rare Books**
 - <http://dsr.nii.ac.jp/toyobunko/>
- **Mapping**
 - <http://dsr.nii.ac.jp/digital-maps/mapping/>
- **MemoryHunt**
 - <http://dsr.nii.ac.jp/memory-hunting/>

Funding support : Digital Criticism: Link-Oriented Research Infrastructure for Evidence-Based Digital Humanities, Grant-in-Aid for Scientific Research (B), Japan Society for the Promotion of Science [No. 16H02920]