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

Photograph

Map

Gazetteer
Linking Entities Across Sources

Textual criticism

Textual source A

Place name S

Textual criticism

Textual source B

Place name T

Geographic
And Visual
Relationship

Place name U
Non-textual Source C, D, E...

Place name V

Digital criticism
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

<table>
<thead>
<tr>
<th><strong>Evidence</strong></th>
<th>Relationship between <em>sources</em>.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis</strong></td>
<td>Relationship between <em>concepts</em>.</td>
</tr>
<tr>
<td><strong>Fact</strong></td>
<td>Relationship between evidences and hypotheses that claims <em>reusable knowledge</em>.</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td>An <em>attribute</em> for an evidence and a hypothesis to represent the degree of reliability defined by a scholar.</td>
</tr>
</tbody>
</table>
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.

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?
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.


3. MemoryHunt: Spatial and visual matching of photographs and the real world (Mobile app).
Mappinning
http://dsr.nii.ac.jp/digital-maps/mappinning/

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

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

```sparql
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 ;
FILTER (?CF > 80)
}
```

<table>
<thead>
<tr>
<th>instance_a</th>
<th>Relation</th>
<th>instance_b</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>萩津湖遺跡 (Chotetsu)</td>
<td><a href="http://schema.org/sameAs">http://schema.org/sameAs</a></td>
<td>萩津湖遺跡 (Chotetsu)</td>
<td>85</td>
</tr>
<tr>
<td>萩津湖遺跡 (Chotetsu)</td>
<td><a href="http://schema.org/sameAs">http://schema.org/sameAs</a></td>
<td>Chikkan-kou</td>
<td>80</td>
</tr>
<tr>
<td>萩津湖遺跡 (Chotetsu)</td>
<td><a href="http://schema.org/sameAs">http://schema.org/sameAs</a></td>
<td>七澤湖遺跡 (Shirakawa)</td>
<td>80</td>
</tr>
</tbody>
</table>

2016/07/15 Digital Humanities 2016
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/

• **Mappinning**
  • http://dsr.nii.ac.jp/digital-maps/mappinning/

• **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]