

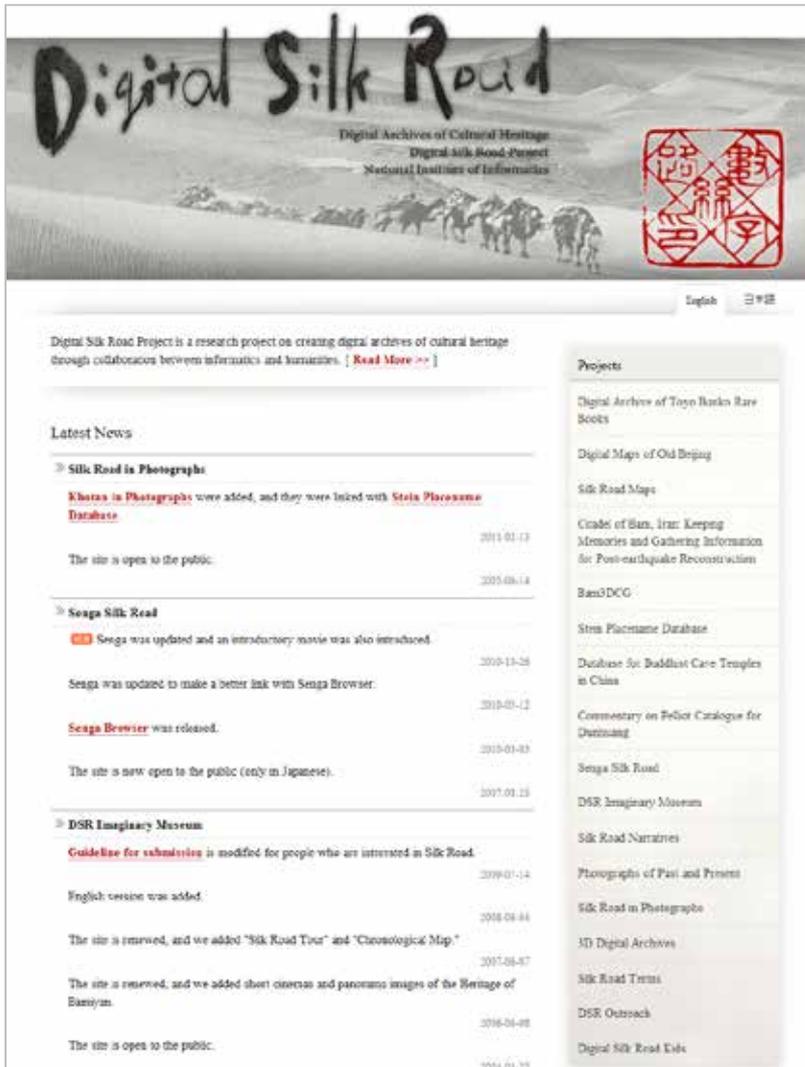
# Geohumanities and the Digital Silk Road

Asanobu KITAMOTO, National Institute of Informatics

Collaborator: Yoko NISHIMURA, Toyo Bunko

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

# Digital Silk Road Project



- Started in 2001.
- **Digital Humanities:** Collaborative work = informatics + humanities scholars.
- Databases and digital resources are **publicly accessible on the Web.**

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

# Toyo Bunko Rare Books

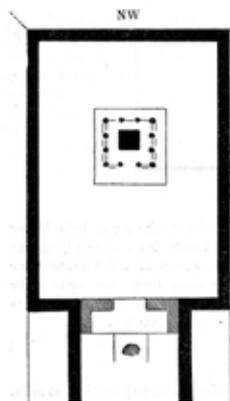


- Digitization of 203 books, 59358 pages.
- Collection of relevant academic references.
- Manual input of captions and TOCs.
- OCR for full-text search (with errors).

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

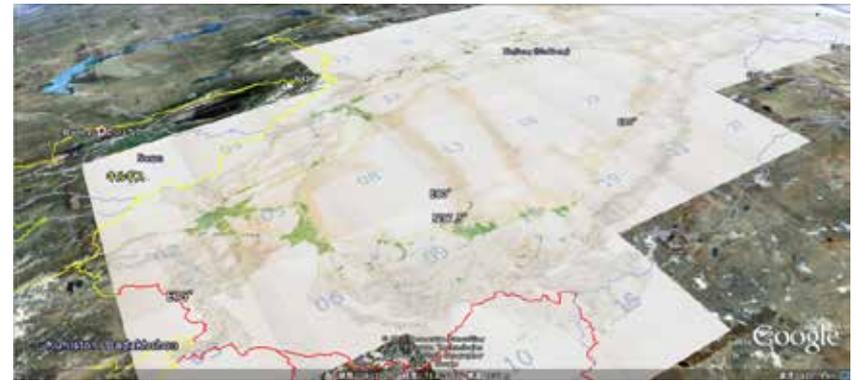
# Dealing with Heterogeneous Data

## 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 Stäben 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

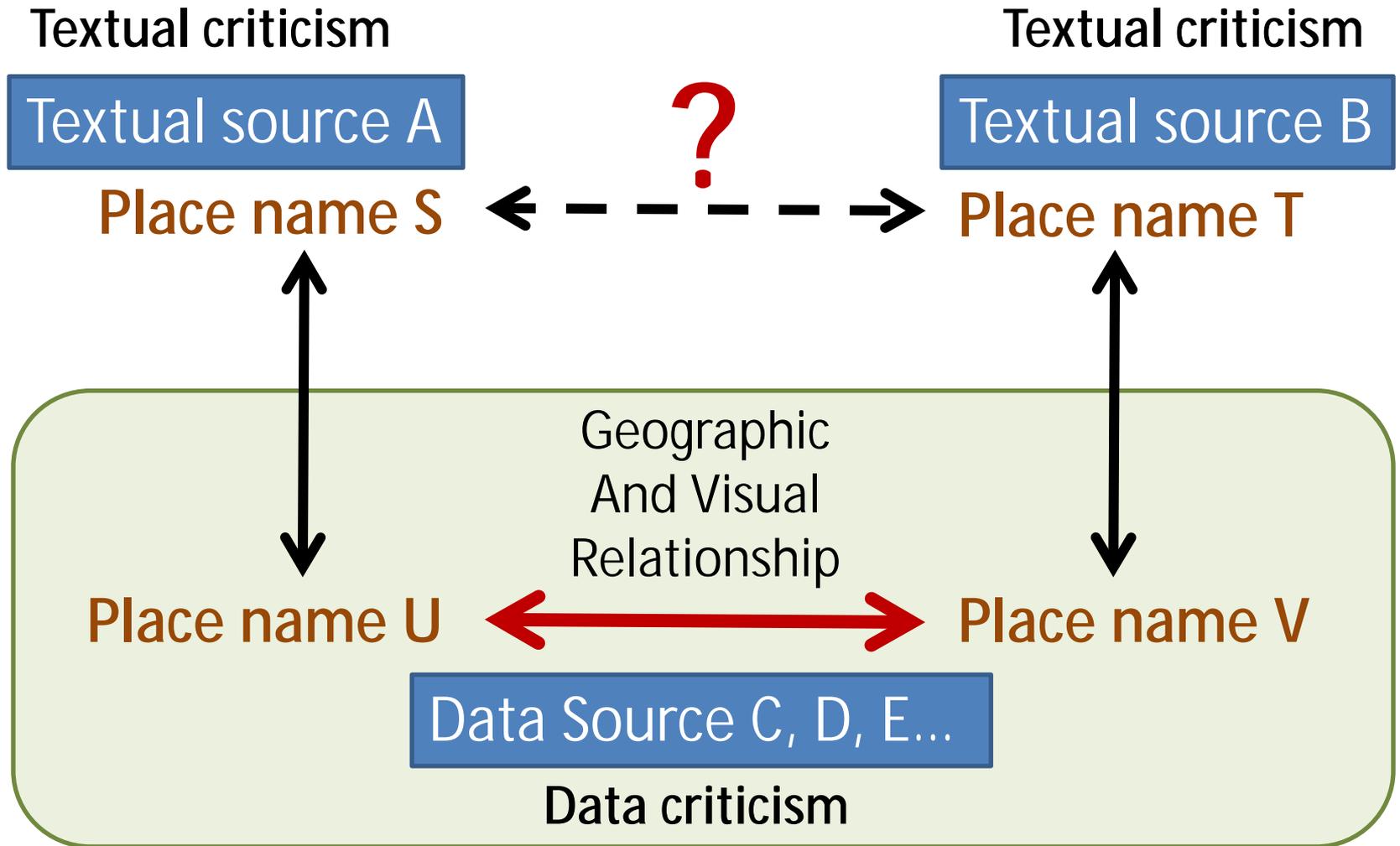


Y  
1904(Le Coq, 1913, Tafel. 70, I)

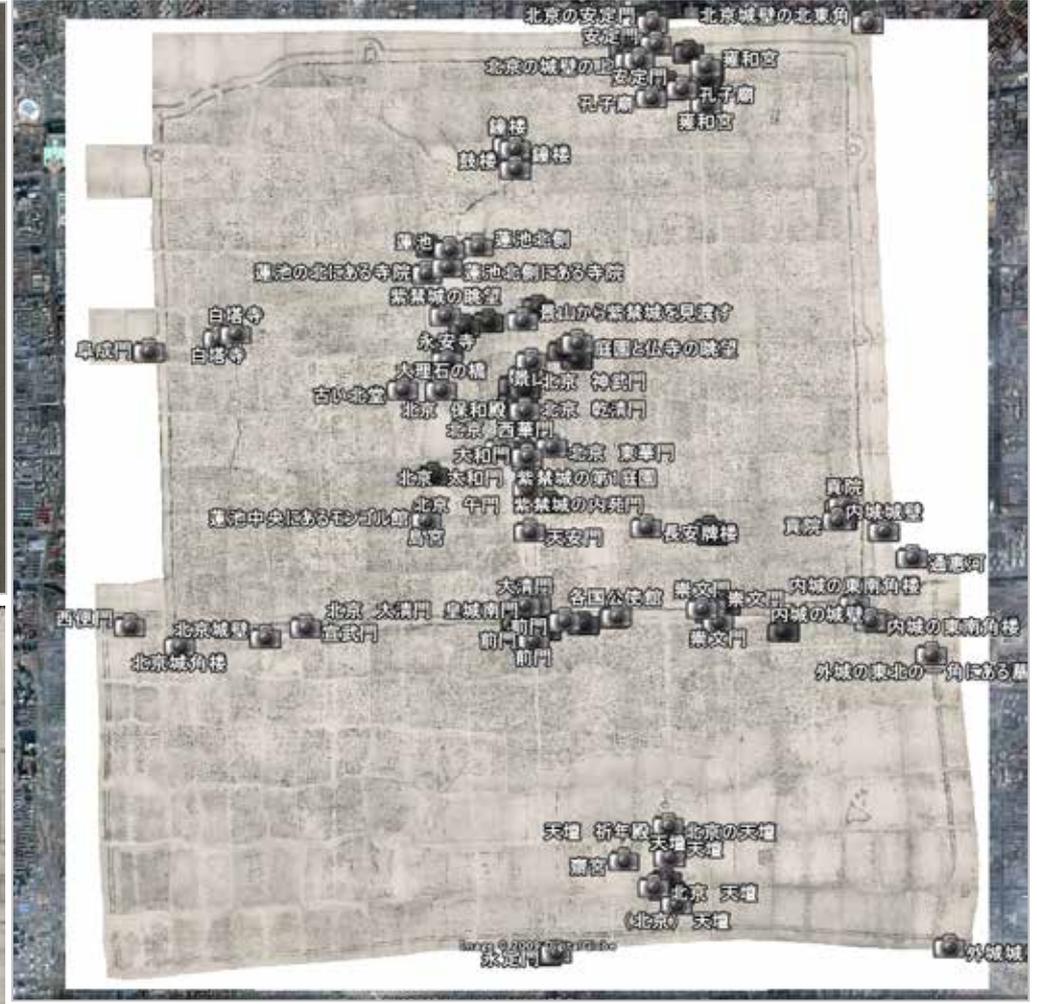
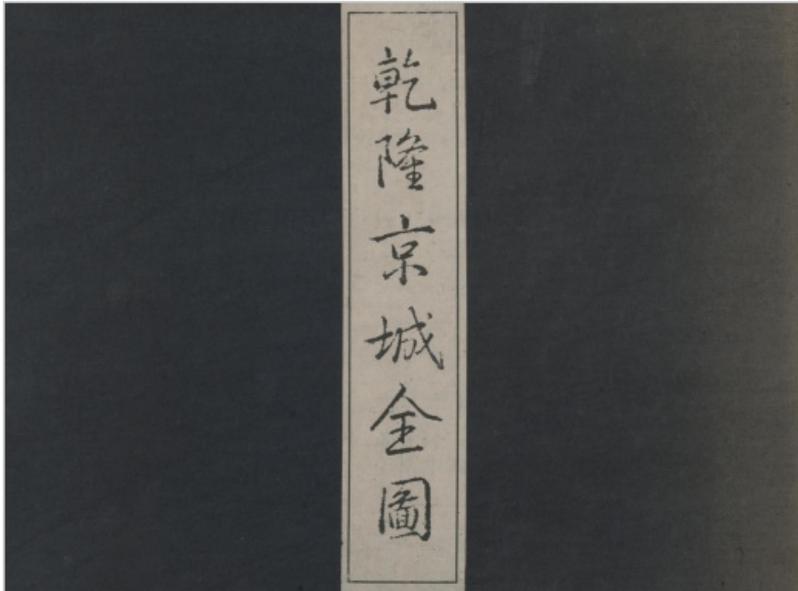
## Gazetteer

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.	

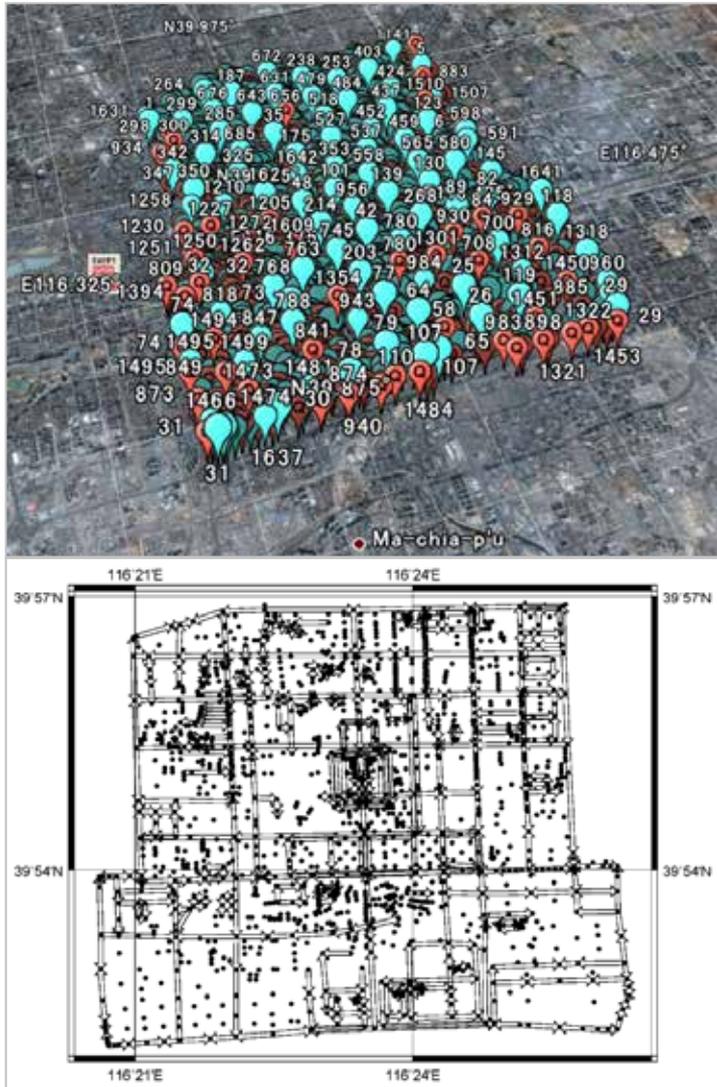
# Linking Entities Across Sources



# Qianlong Map (Beijing)



# Massive Geometric Correction



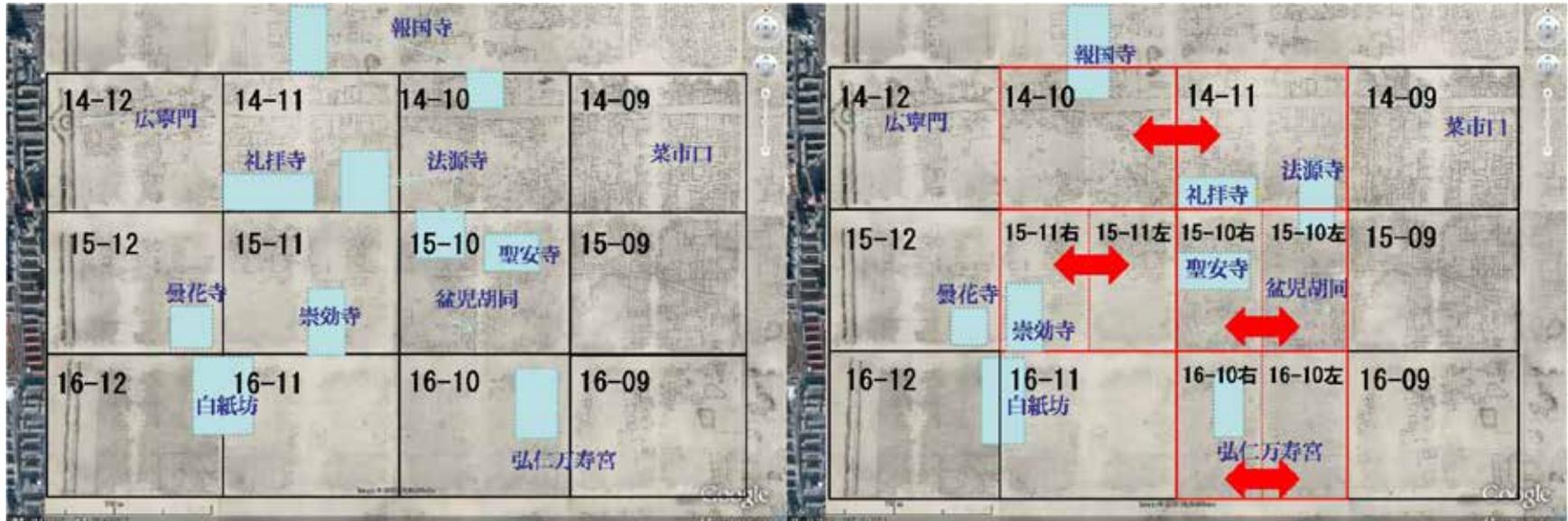
**Huge size** = W 13 m x H 14 m

**Many sheets** = 203 sheets in total

**Massive pixels** = 29 billion pixels

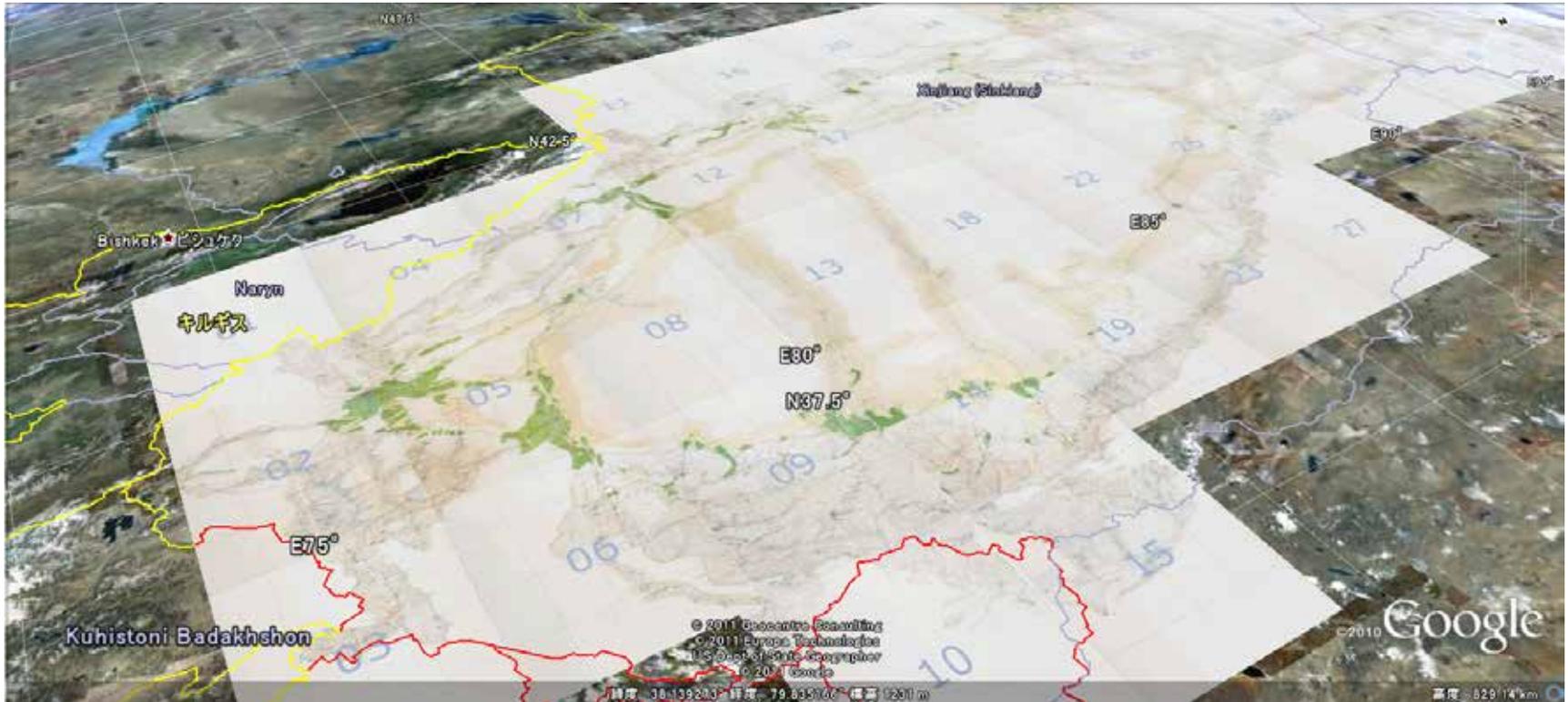
- **Control points + lines:**  
We proposed a new geometric correction.

# Discovery and Digital Survey



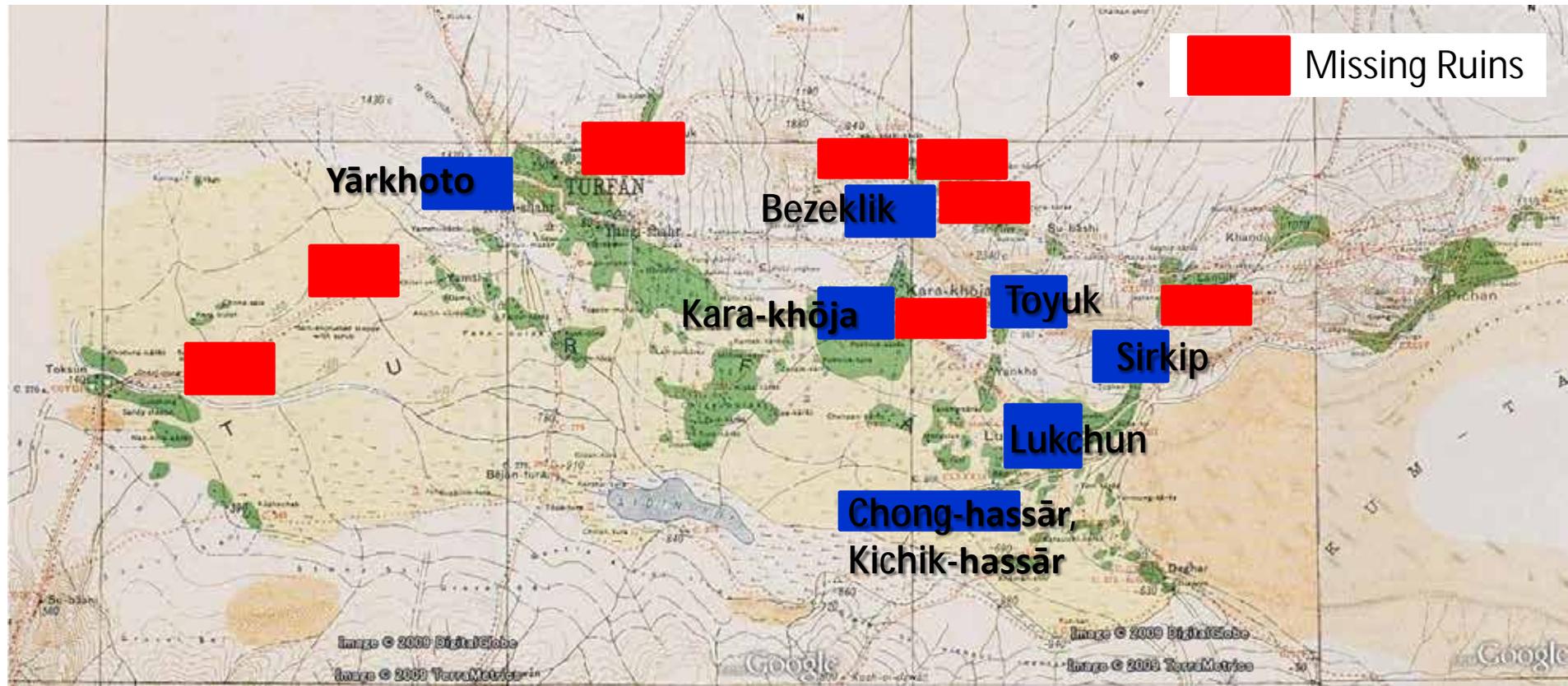
- **Discovery:** 5 sheets have misarrangements due to improper reconstruction in the past.
- **Digital Survey:** Place-names were checked to make a comprehensive gazetteer.

# Stein Map (Silk Road)



- Stein's map "Innermost Asia" was registered and displayed on Google Earth satellite images.

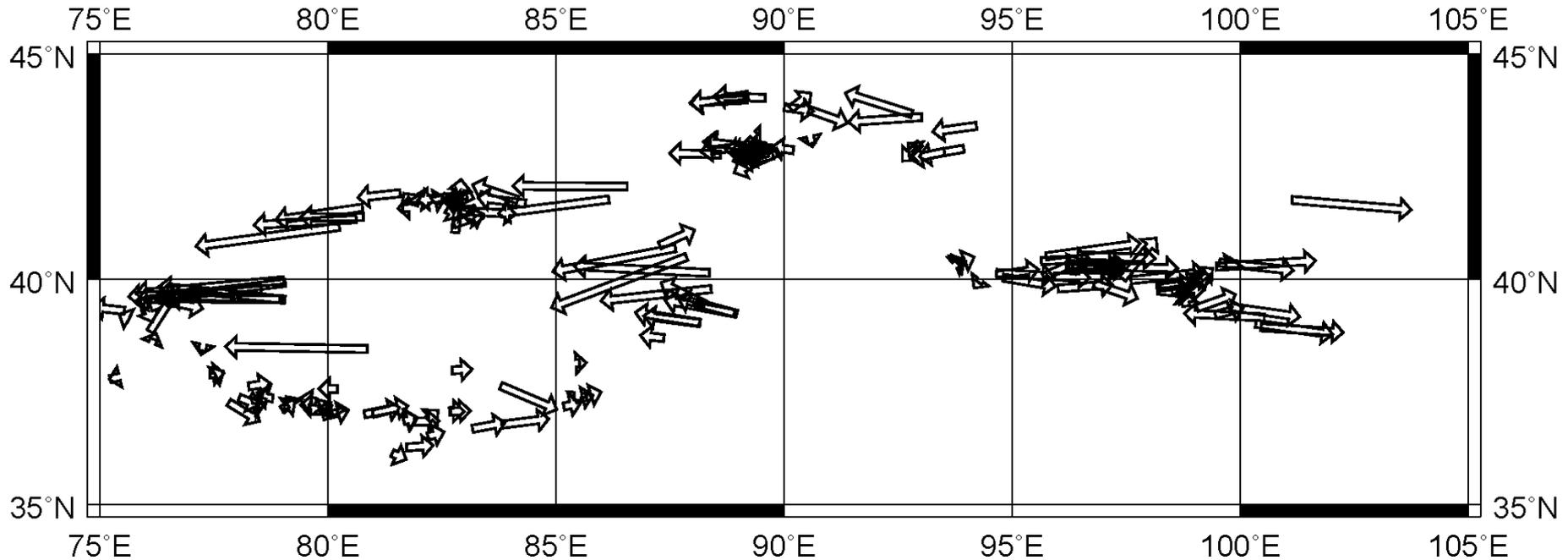
# Problem of "Missing" Ruins



Oi-tam, ruined fort  
Bögan-tura  
Buluyuk ( Shipang, Sassik-bulak, Kazma )  
Murtuk-ruins

Yoghan-tura  
Chikkan-köl  
Bedaulat's town, Bēsh-kāwuk, Kosh-gumbaz  
Yutōgh

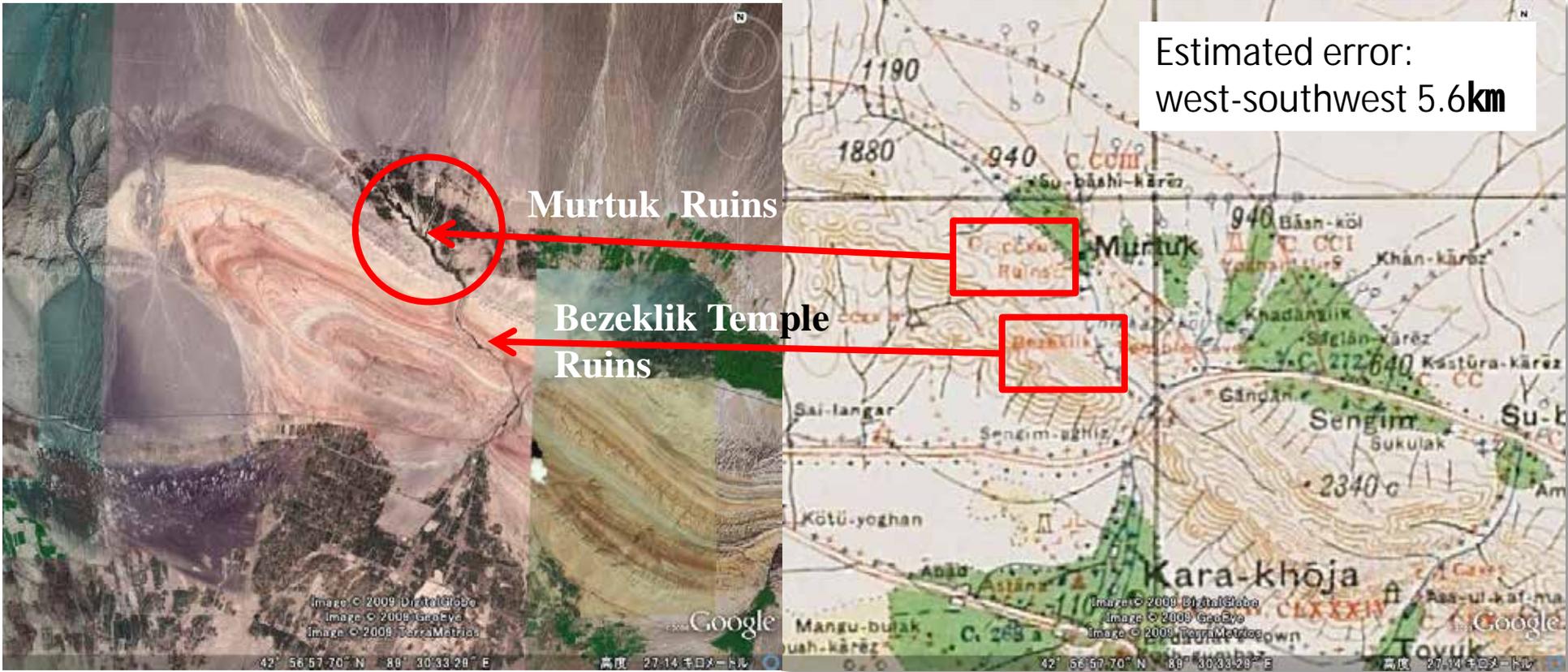
# Error Distribution of Maps



Error Distribution of Innermost Asia, Aurel Stein, 1928.

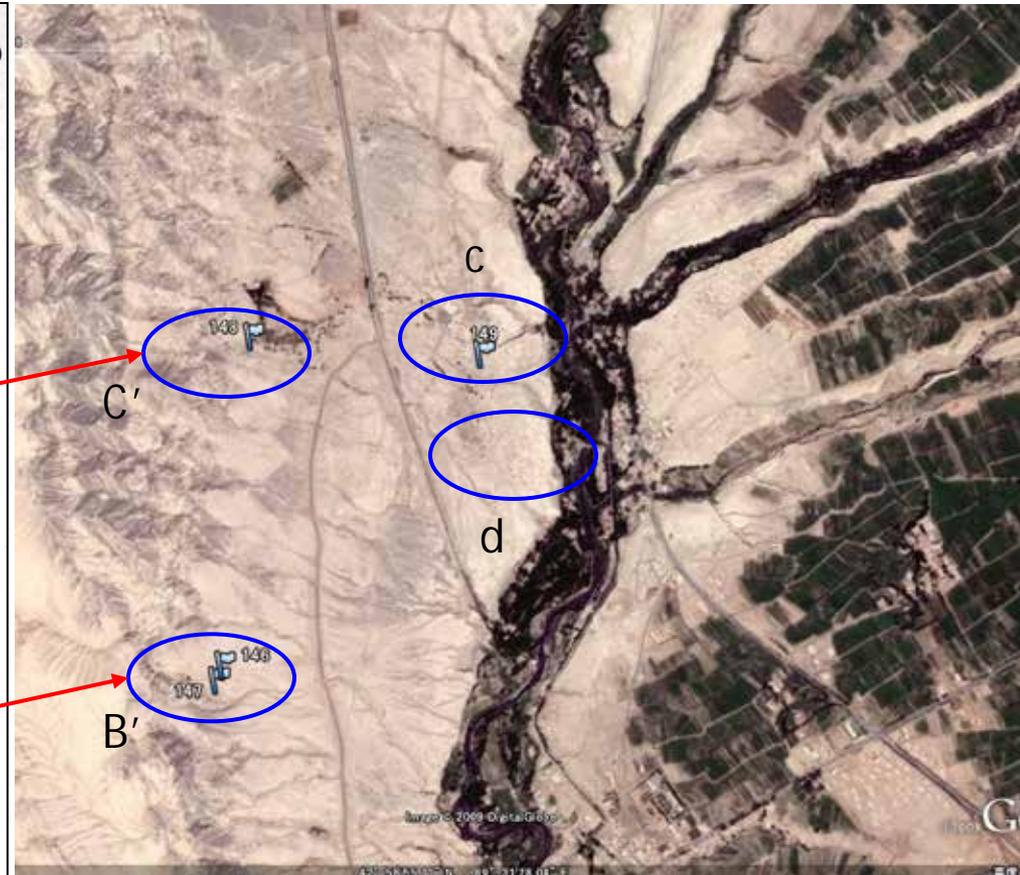
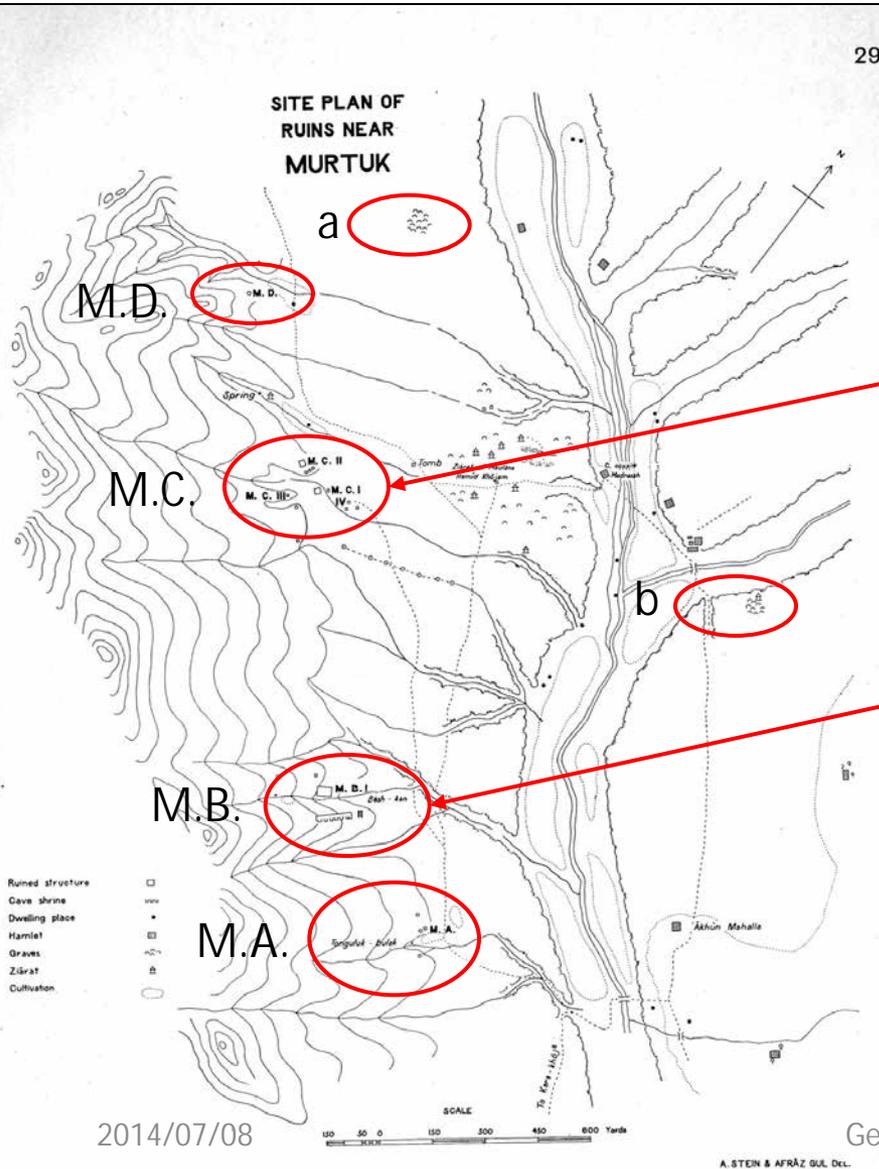
- Errors are bigger for longitude due to the limitation of survey technology around 1900.

# Estimating the Current Location



Based on the error information of maps, we estimate the location of Murtuk Ruins at .

# Matching Entities



Stein's map and satellite images for the same area. Each source reports different ruins due to different conceptualization.

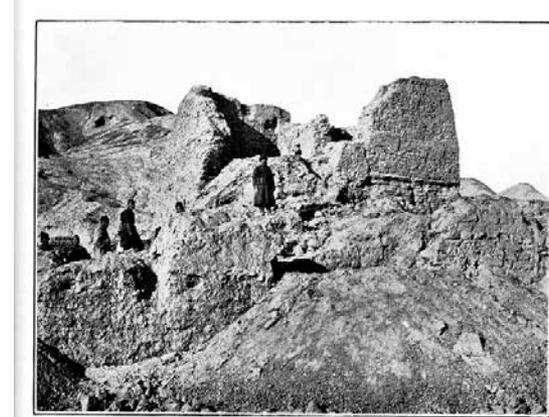
# Photographs as Evidence



伯西哈石窟(烏江不拉克仏塔)



烏江不拉克烽火台



305. RUIN, M.B. I, AT MURTUK IN COURSE OF CLEARING.

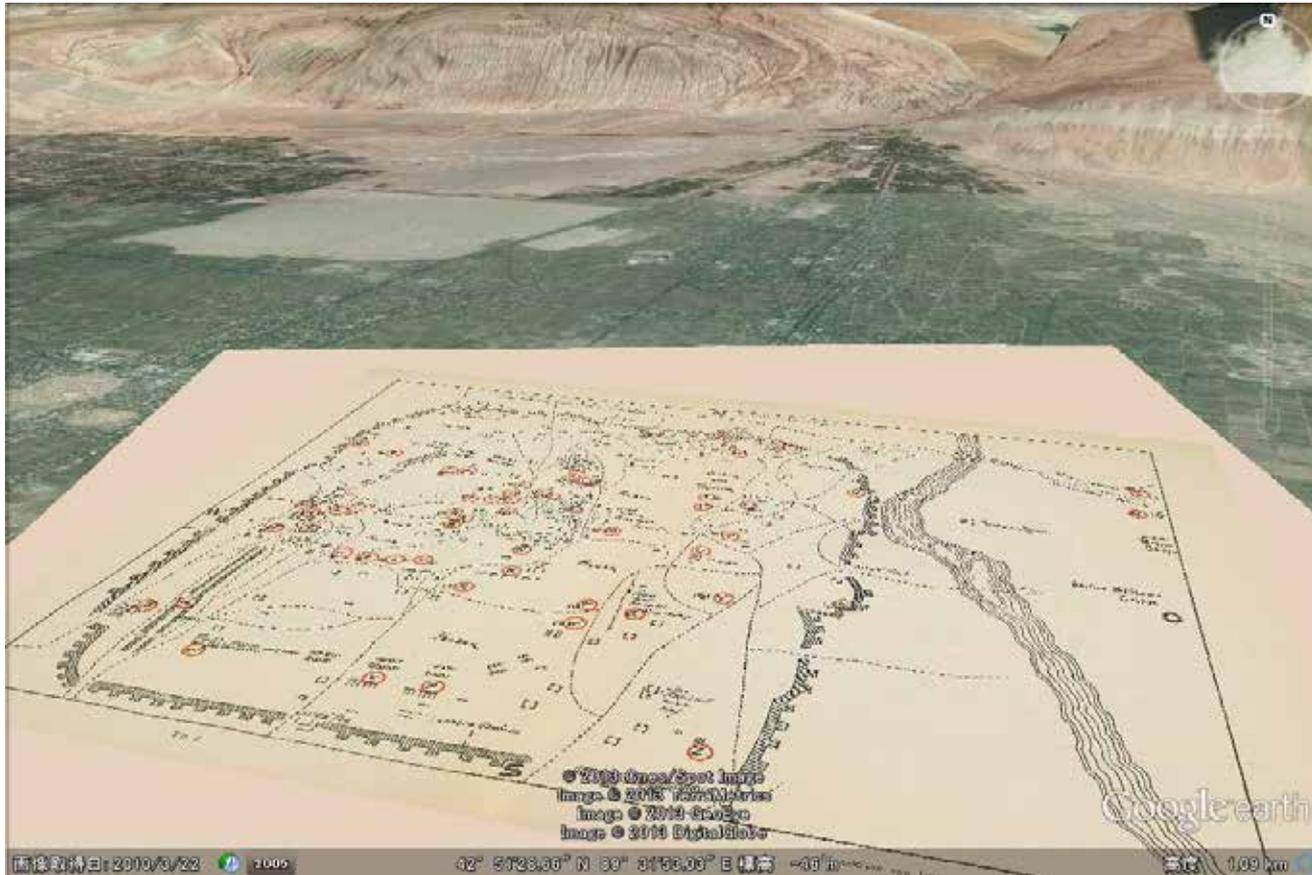
Murtuk Ruins (M. B. I)



303. RUINED SHRINES, M.C. I, AT MURTUK, TURFAN.

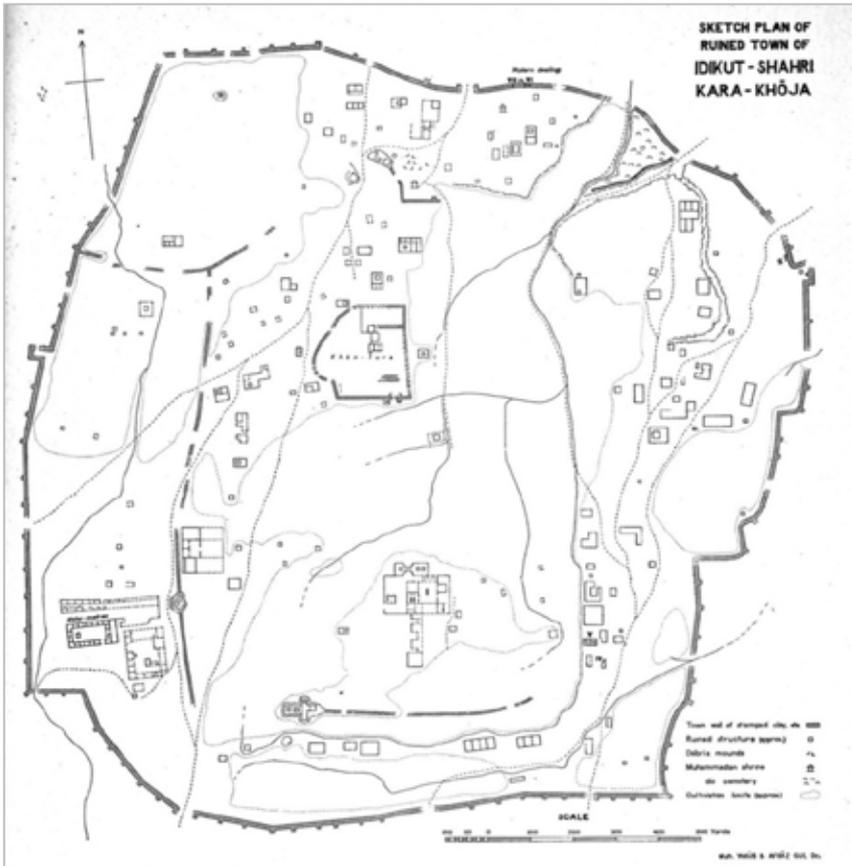
Murtuk Ruins (Ruined Shrine M. C. I)

# Grünwedel Map (Gaochang)

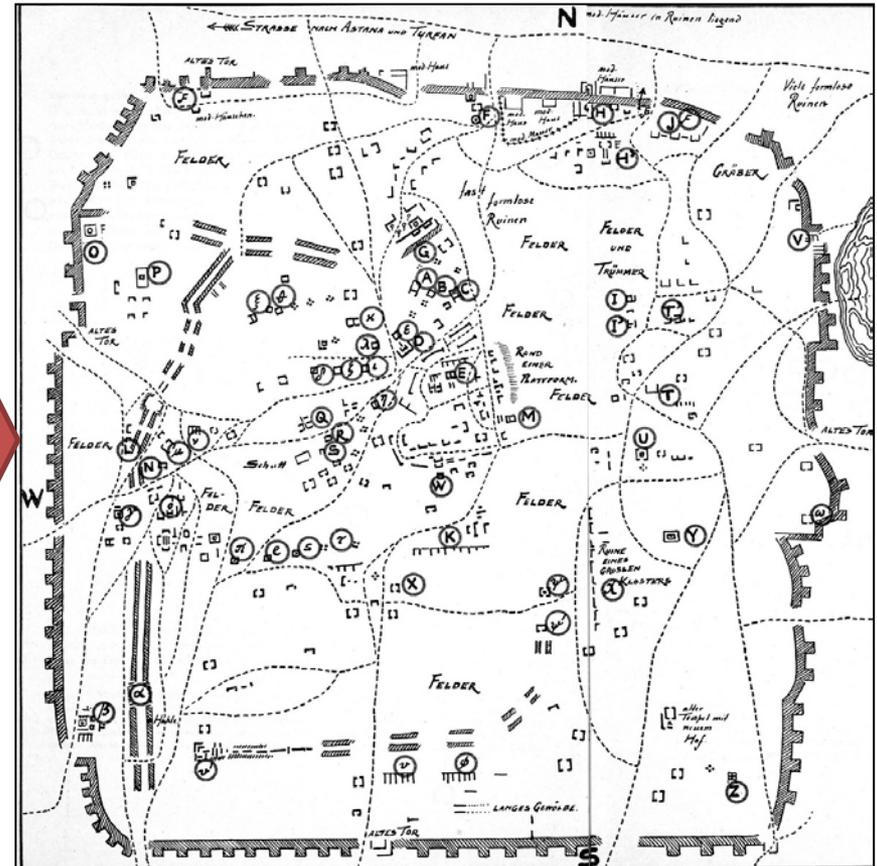


How to use the previously “untrusted” map?

# Inconsistencies in Gaochang Maps



Aurel Stein

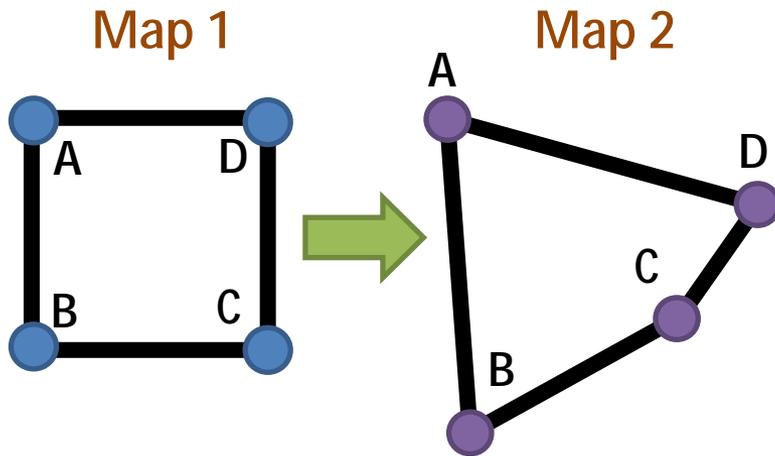


Albert Grünwedel

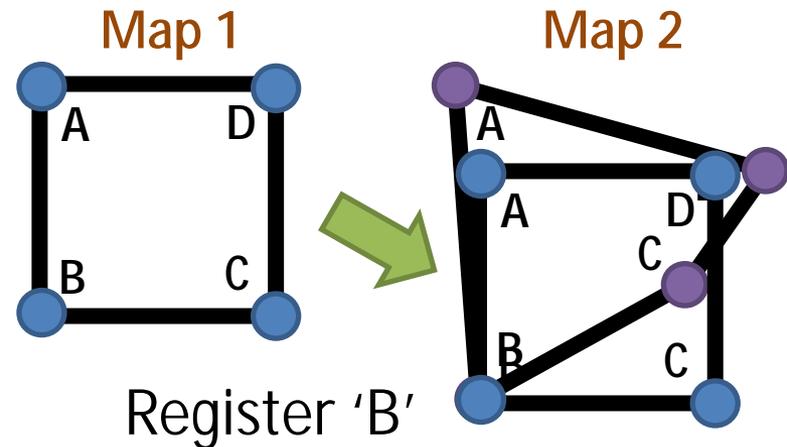


# Registration of Maps

## Geometric Correction



## Interactive Georeferencing



- All points are registered.
- Shapes are distorted.

- Single point is registered (but no other points).
- Shapes are not distorted.

# Mapping (Map+Pinning)

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

**MAPPING**  
Interactive georeferencing by pinning old maps

Digital Silk Road > Database of Ruins in Silk Road > Mapping

日本語 | English | 简体中文

Map Satellite

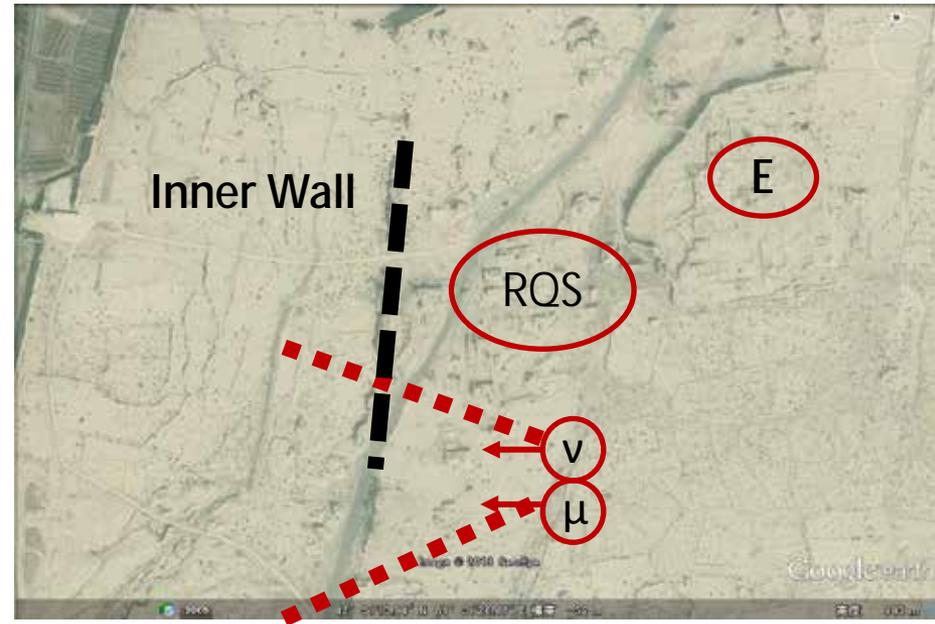
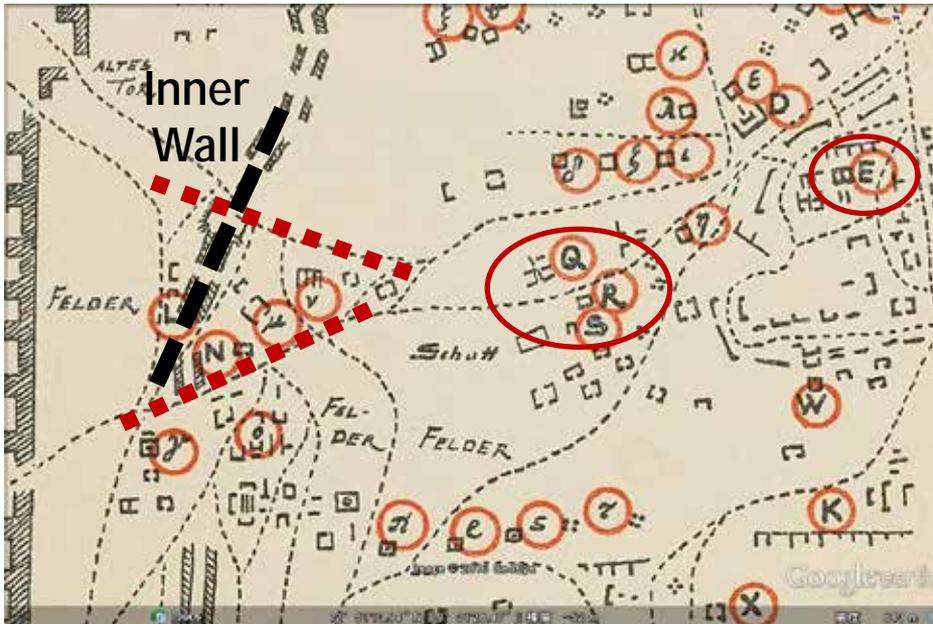
Inmst Asia

**kara-khoja**  
Registrant:yoko  
Location:42.831,89.464  
Error:-0.02,-0.064  
5.69km  
高昌故城

TURFAN  
Yangi-shahr  
Kara-khoja  
Tuyuk  
Yankhe

Map data ©2013 Old Map: National Institute of Informatics "Digital Silk Road"  
Map data ©2013 AutoNavi, Google Imagery ©2013 TerraMetrics - Terms of Use

# Topological Interpretation

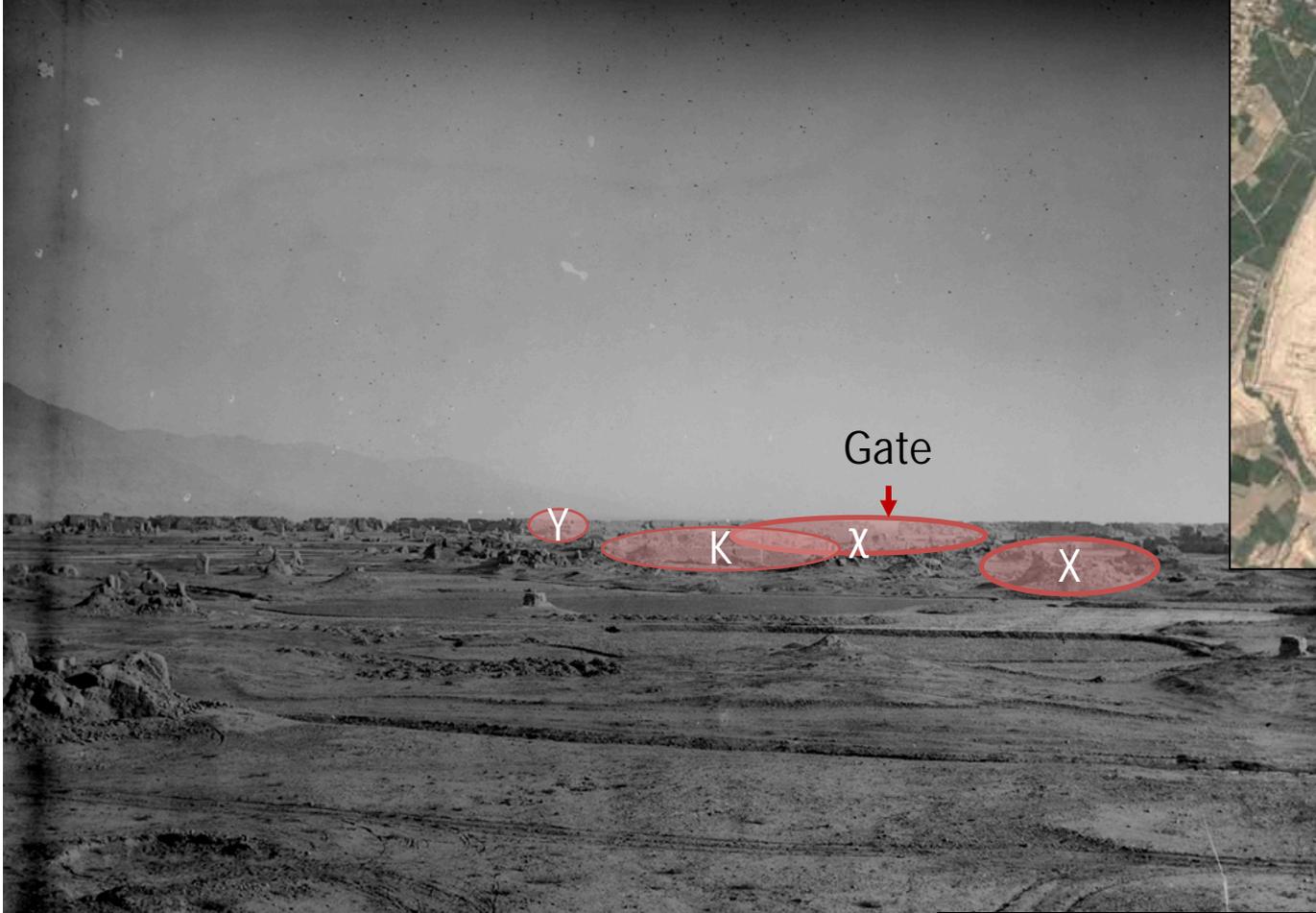


Where is  $v$  and  $\mu$ ?  
We have multiple candidates...

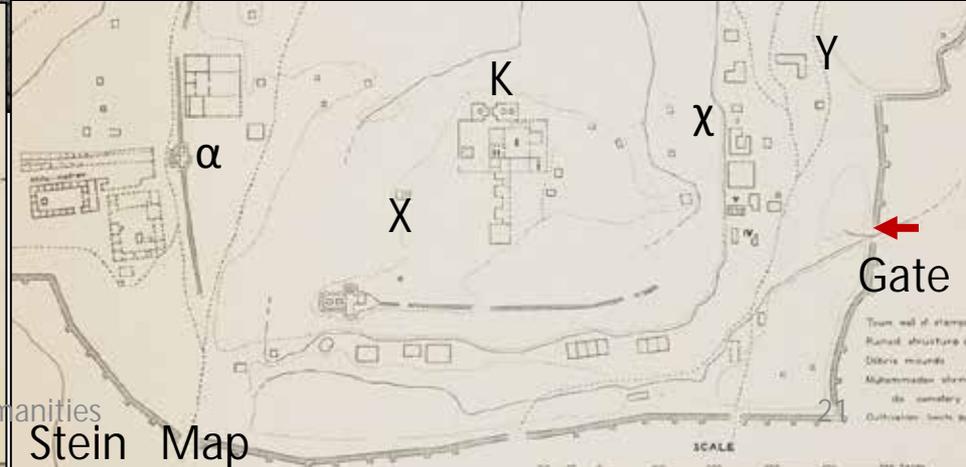
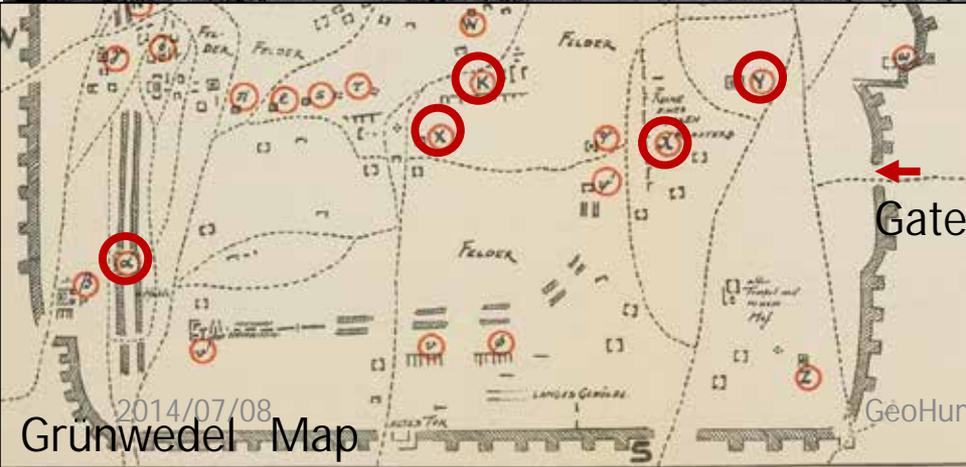
Search for a road in north of  $\gamma$  and  $o$ , and a road between the inner wall to the wall gate.

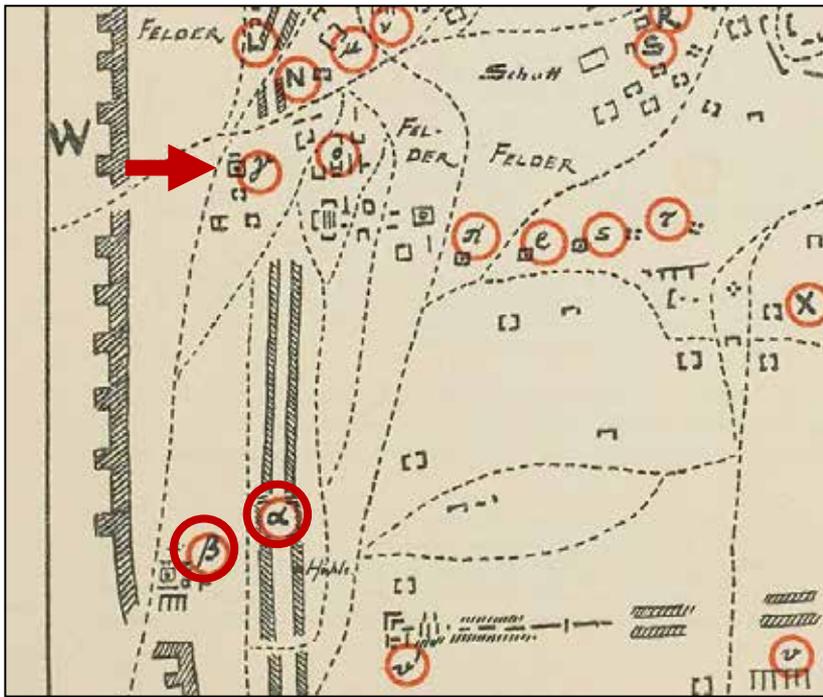
Hypothesize the location of  $v$  and  $\mu$  (to be verified later).

Maps designed for navigation purposes should preserve the topology, not to get lost when visiting there again.

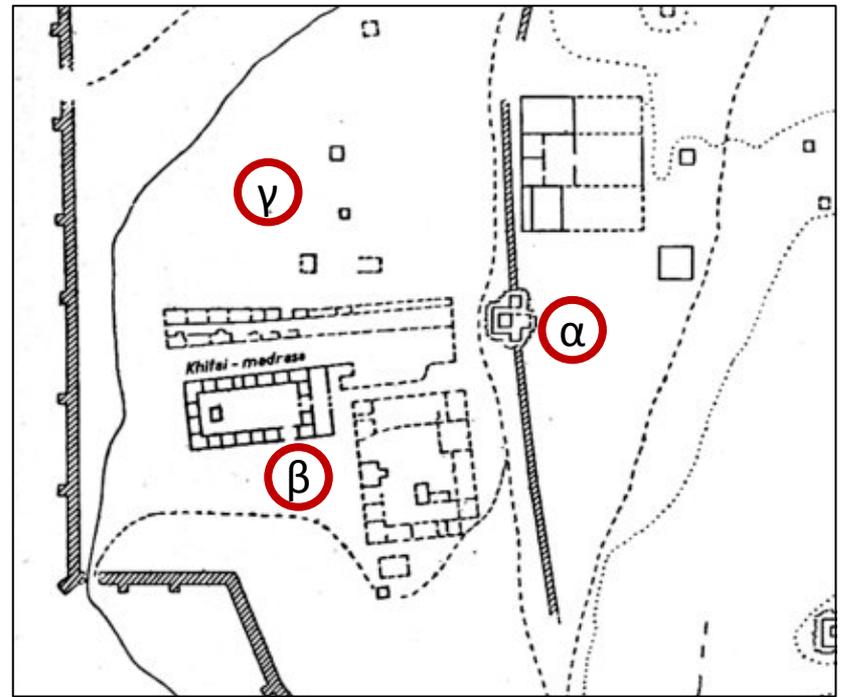


↑ Viewpoint from  $\alpha$  to east  
 ← MIK B 1278, IDP Database

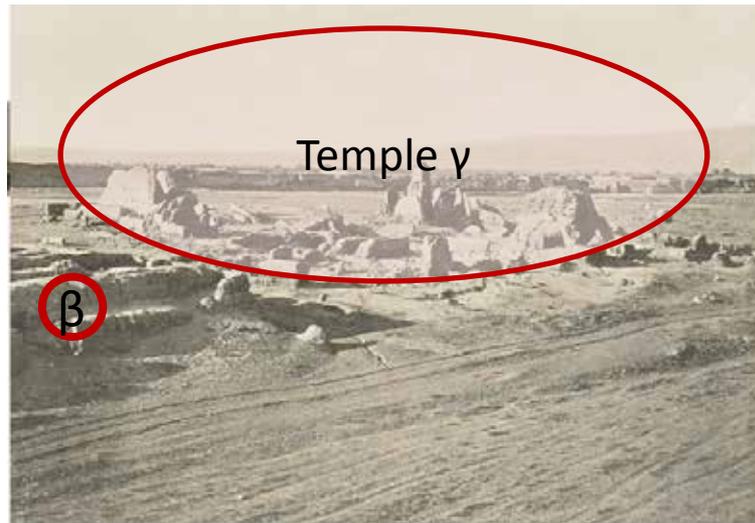




Grünwedel Map

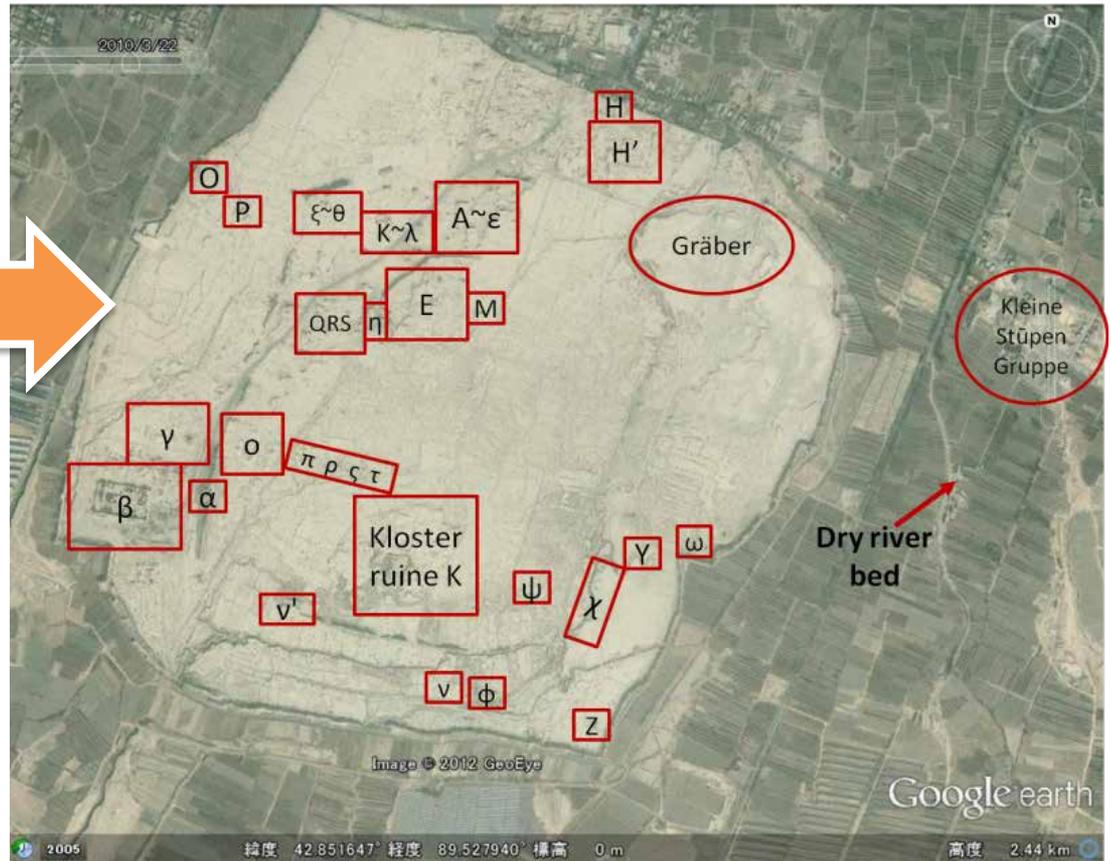
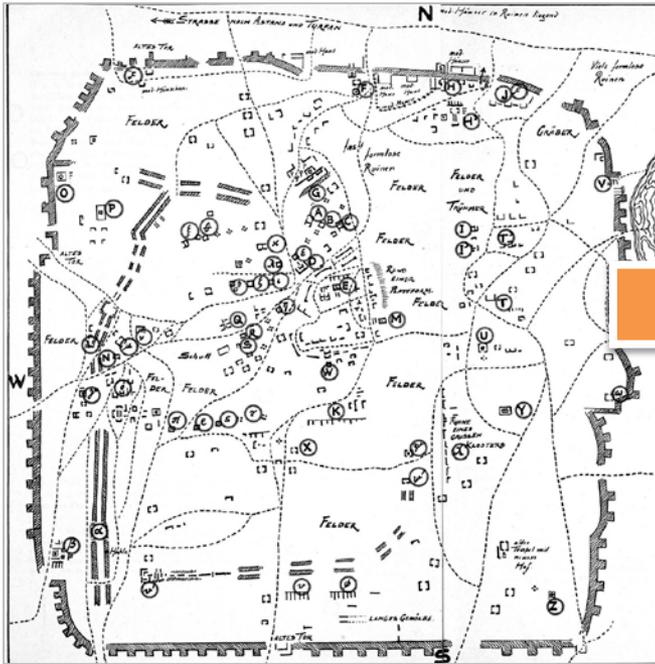


Stein Map



Tempel  $\gamma$ , seen from Tempel  $\alpha$   
 Le Coq, *Chotscho*, Tafel 70, I

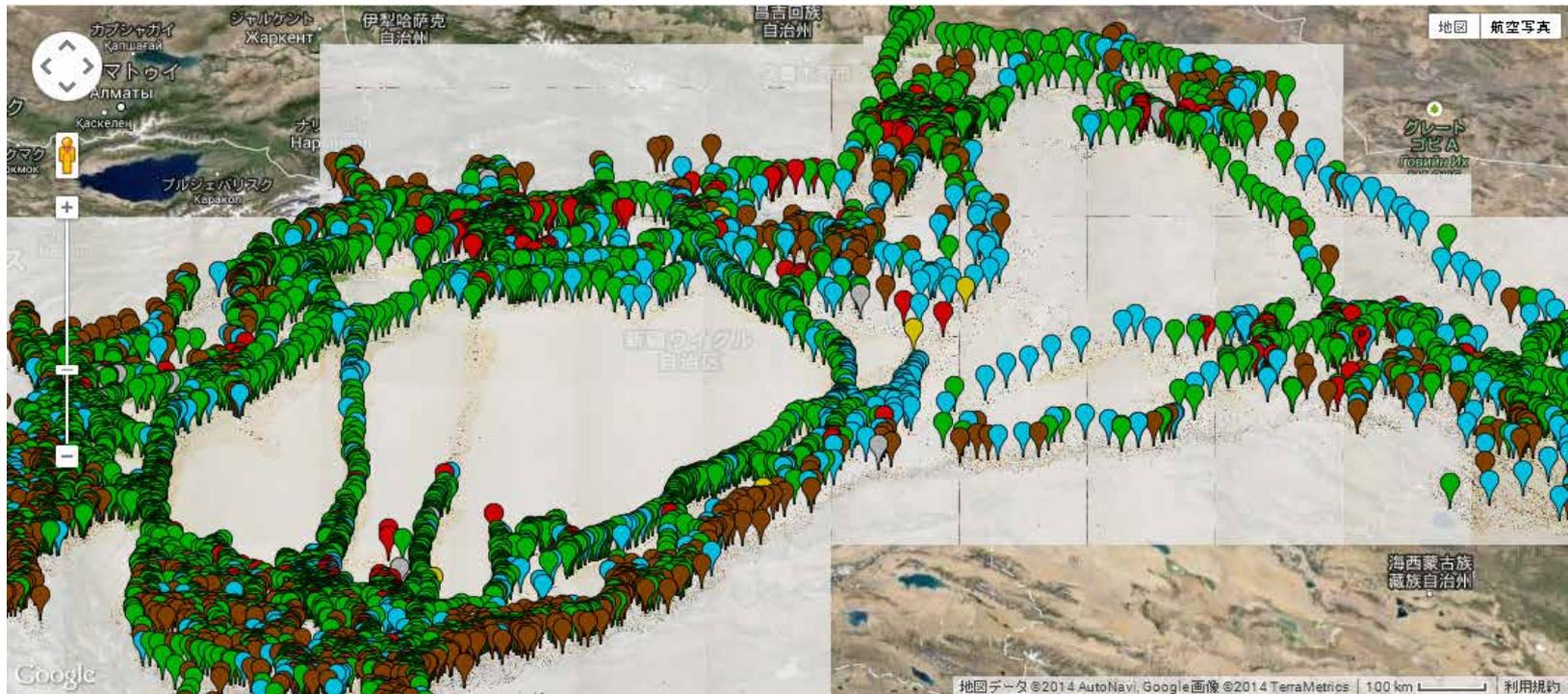
# Most Ruins are Identified



Linking expedition results to recent surveys may lead to new interpretations.

# Stein Placename Database

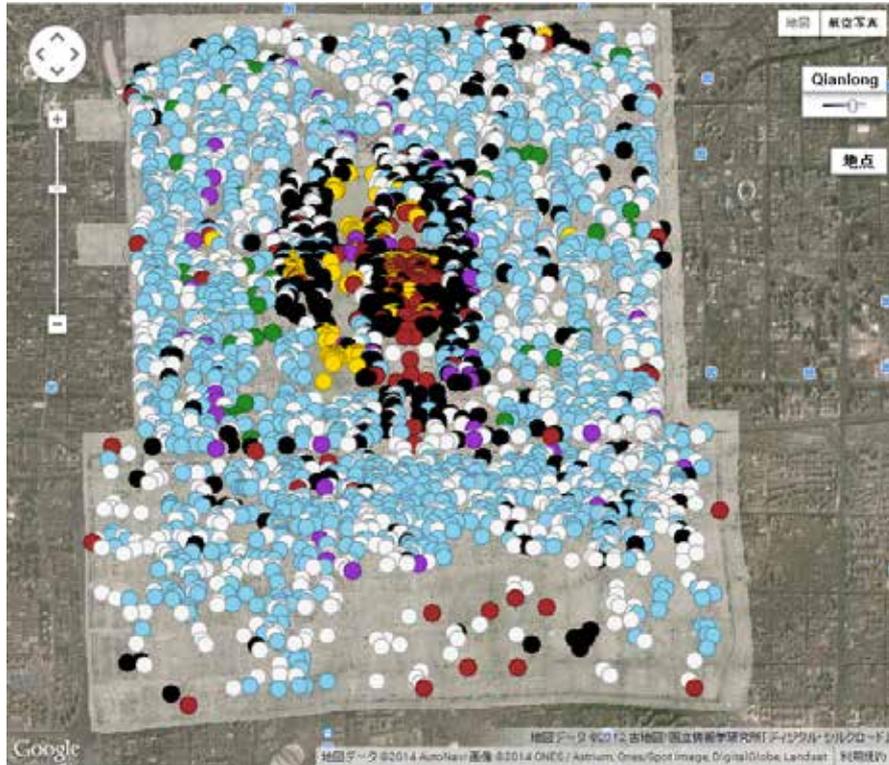
<http://dsr.nii.ac.jp/digital-maps/stein/place-names/map/>



- 5528 place names in “Innermost Asia” maps (still adding new names not in the Index).

# Beijing Placename Database

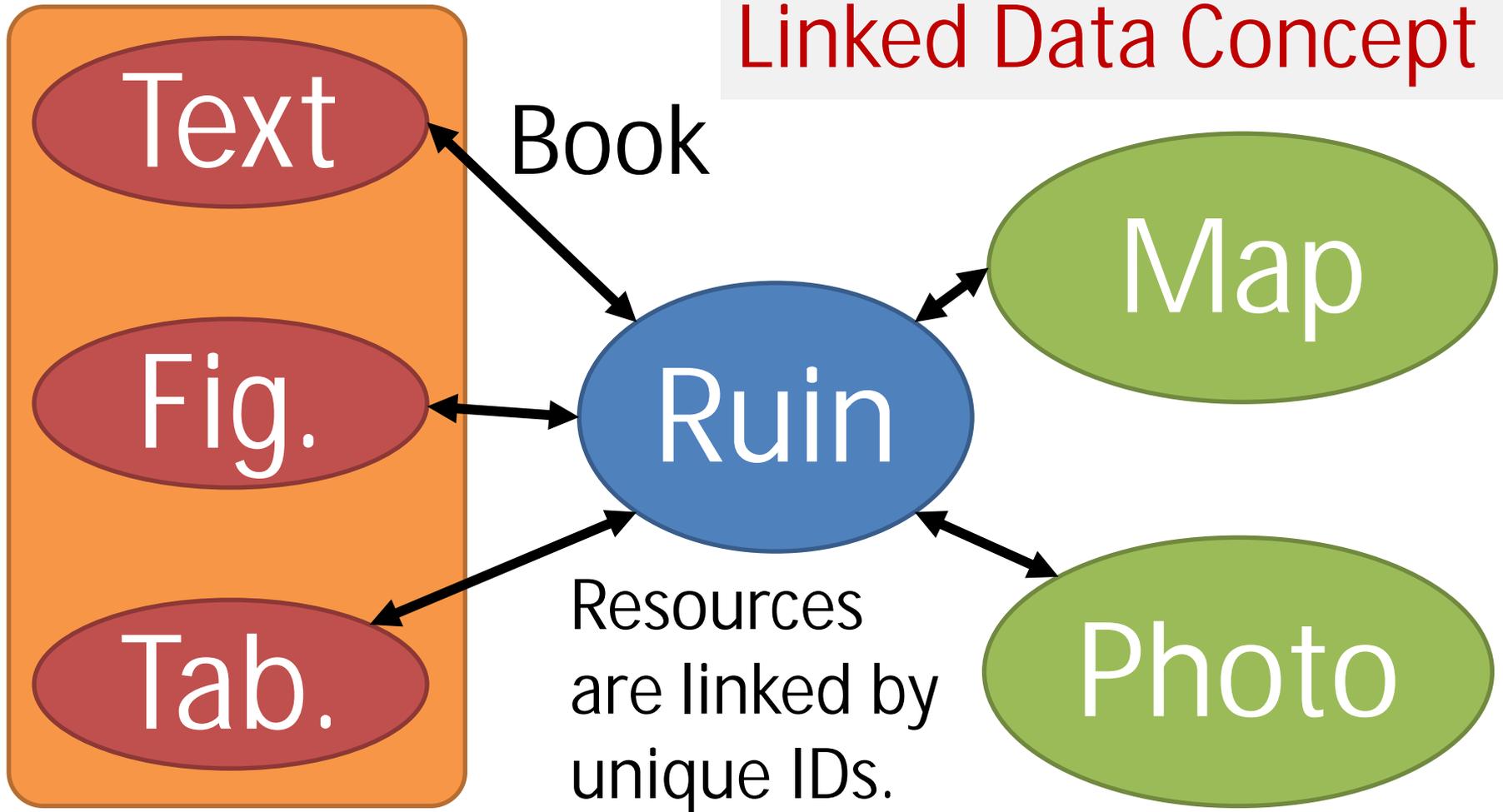
<http://dsr.nii.ac.jp/beijing-maps/search/>



- 3609 place names in the Qianlong Map, manually identified and classified.

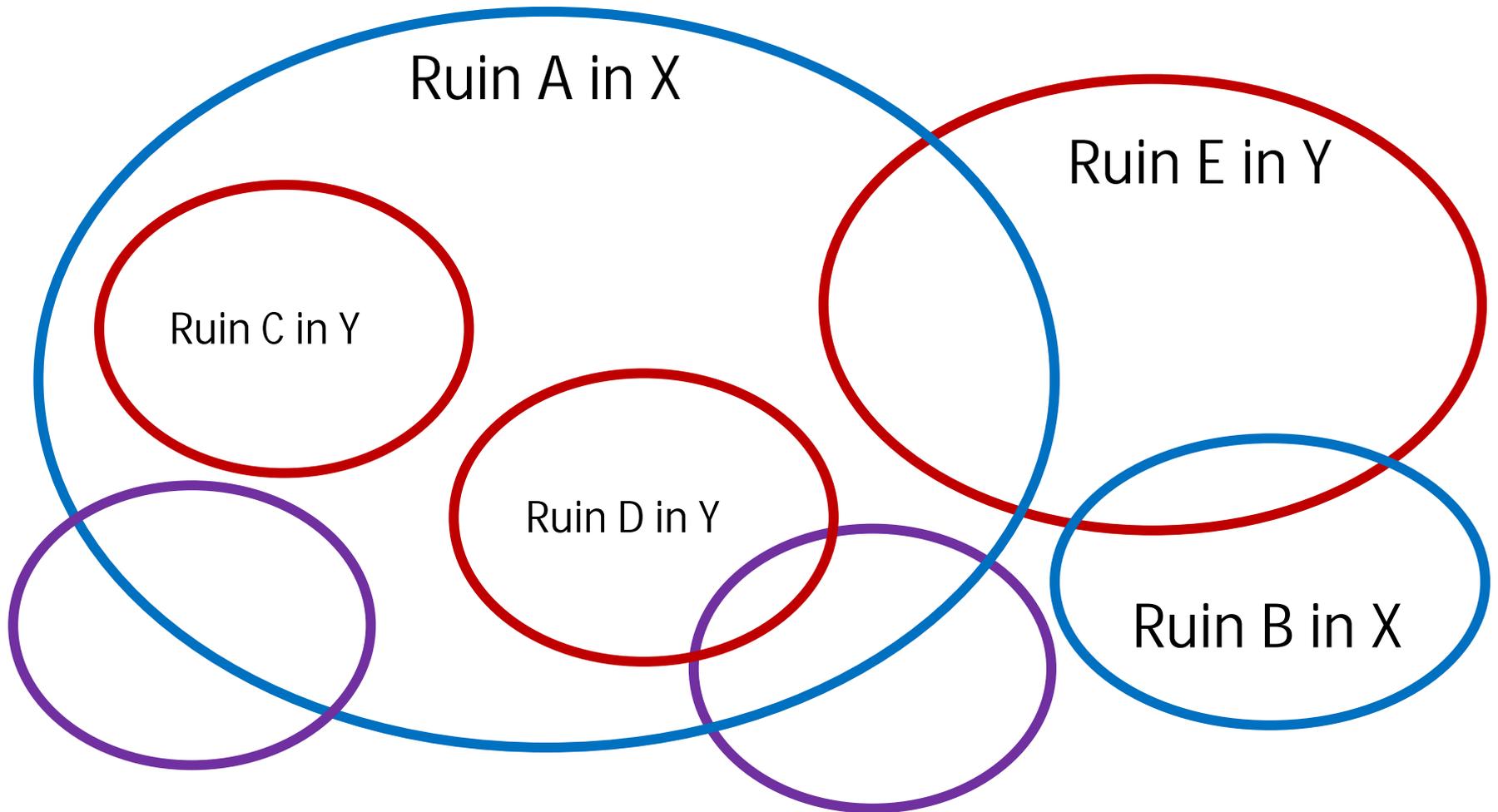
# Plan for the Ruin Database

Linked Data Concept



# Different Conceptualization

Source X and Y may have different concepts on ruins.



# GeoNLP (Geo+NLP)

<https://geonlp.ex.nii.ac.jp/>

Task: Named Entity Recognition

**Extract and disambiguate place names from natural language text.**

Do you like London? I like Bath.



Ambiguous = Multiple interpretations.

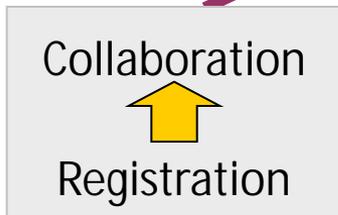
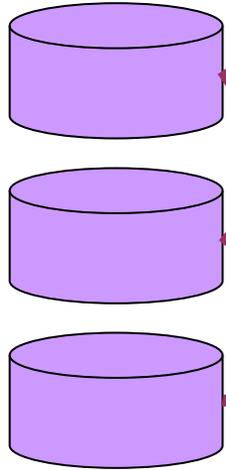
We have many places in the world named *London*.

*Bath* is a city or a hot spa?

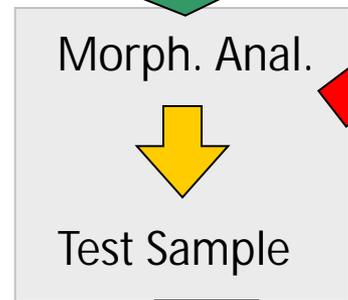
Geo/Non-Geo disambiguation and Geo/Geo disambiguation.

# GeoNLP Framework

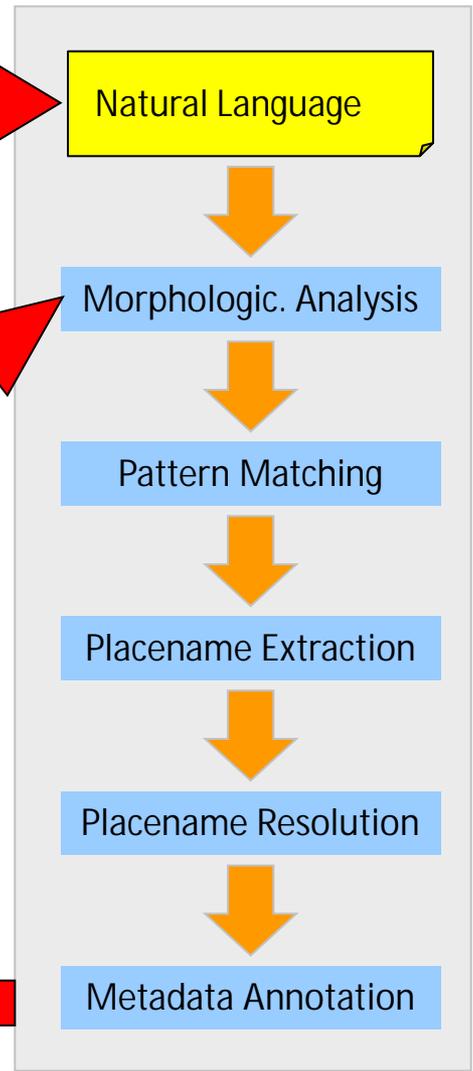
Public resources for placenames



Text or HTML for input into JSON-RPC API



GeoNLP Server



Response in JSON format or use it as CMS modules

Update costs

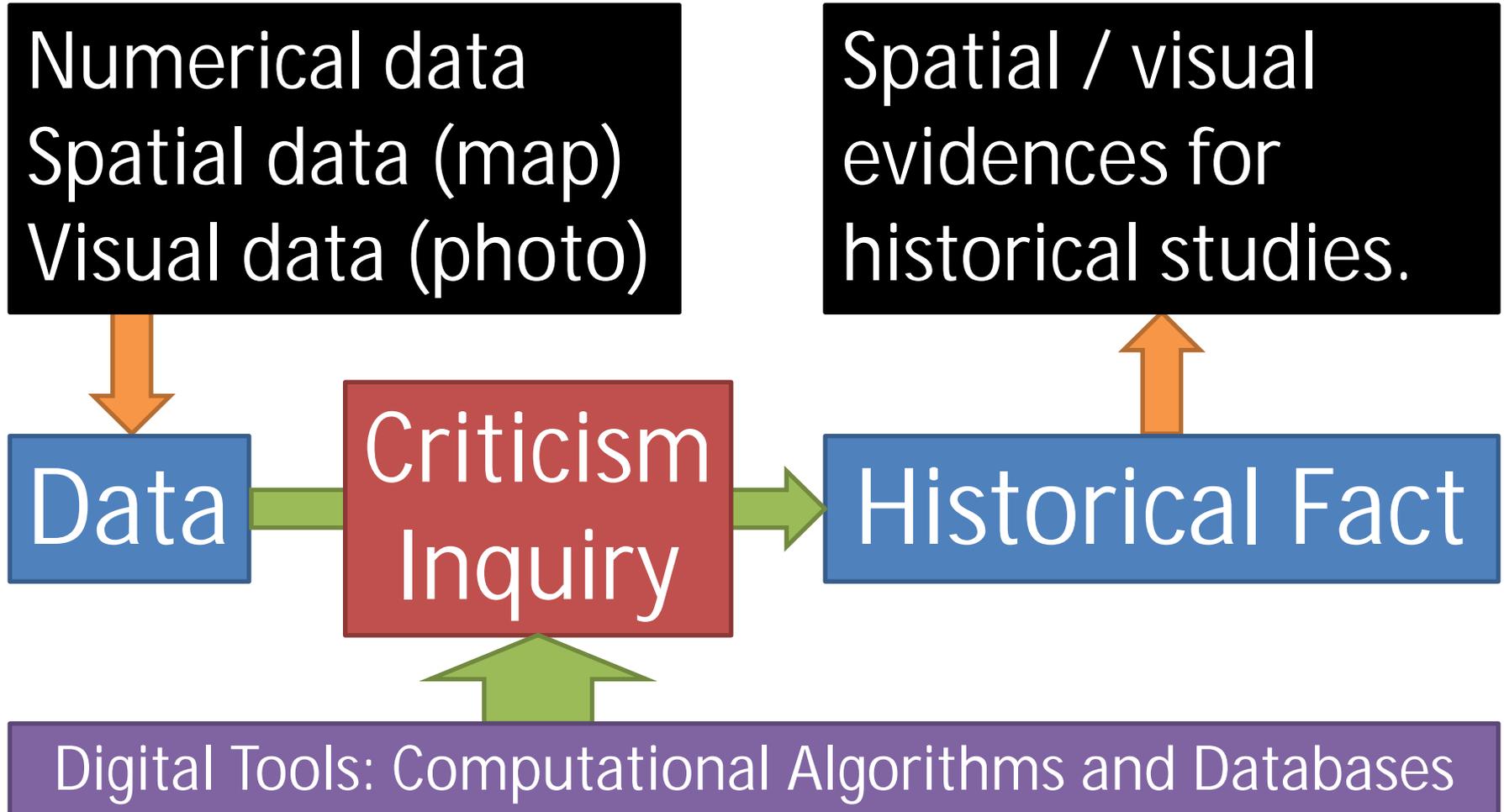
# Open Data and Software Cycle



<https://geonlp.ex.nii.ac.jp/>

- **Open data**: sharing place name gazetteers.
- **Open software**: sharing source codes.
- **Linked open data (GeoLOD)** using D2RQ.
- Goal: **ecosystem** for data infrastructure.

# Data Criticism Approach



# Digital Silk Road Project



數  
字  
絲  
綢  
之  
路

- Website

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

- Toyo Bunko Digital Archive

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

- Silk Road Maps

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

## Contact:

Asanobu KITAMOTO : [kitamoto@nii.ac.jp](mailto:kitamoto@nii.ac.jp)