

2 April, 11:00am

Digital humanities

Dr Miguel Escobar Varela



The 17th Annual Library Leadership Institute
Library Leadership in the Asia Pacific Century
Macau, 31 March - 4 April 2019

Digital Humanities

Challenges for libraries

Miguel Escobar Varela, Ph.D.
Assistant Professor, National University of Singapore
Academic Advisor on Digital Scholarship, NUS Libraries
miguelescobar.com
[@miguelJogja](https://twitter.com/miguelJogja)

This talk

What are the digital humanities?

What this means for libraries: challenges and opportunities

Some ideas for how to tackle these challenges and capitalize on these opportunities

What are the digital humanities (DH)?

The usage of digital data and digital methods for carrying out research in the humanities

Areas in the DH

Databases

Data-driven analysis

Digital publishing

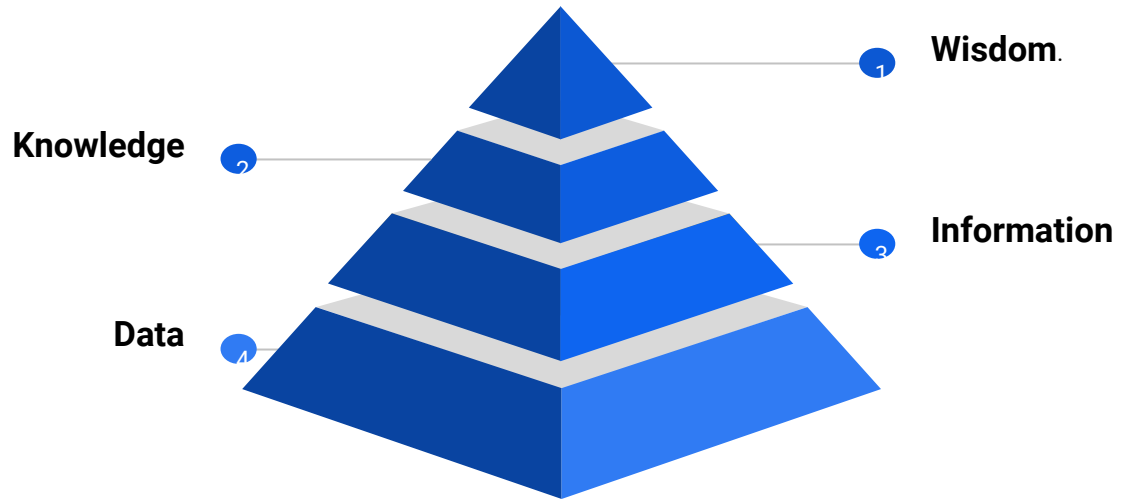
The DH are also a community, a space for meeting up and discussing how technology is changing the humanities

Databases

Databases

Data is 'potential information'. A process needs to be applied to data to become meaningful. This is a challenge to how people usually think about research in the humanities.

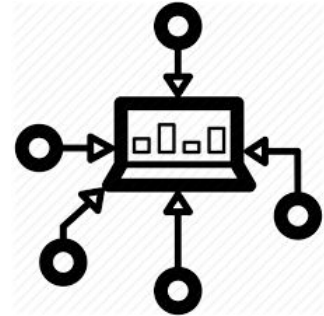
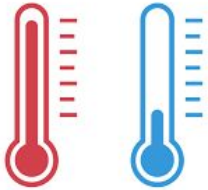
DIKW pyramid



Data is context dependant

- Data is context-dependant
- The question should not be what are data but "when are data" (Borgman, 2015)

That which is data in one context might be information or knowledge in another.



Sources of data in the humanities

Digitized books, films, artworks, video recordings

Motion capture

3d models

Surveys

Production details (film, theatre)

Historical records

Problems of DH data

Standardization

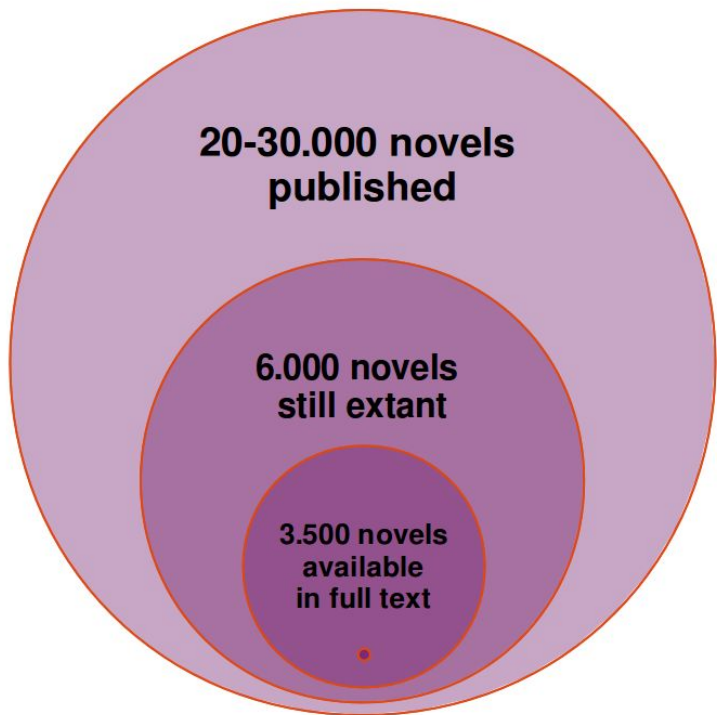
Incompleteness

Inaccuracy

These problems are important for historical data.

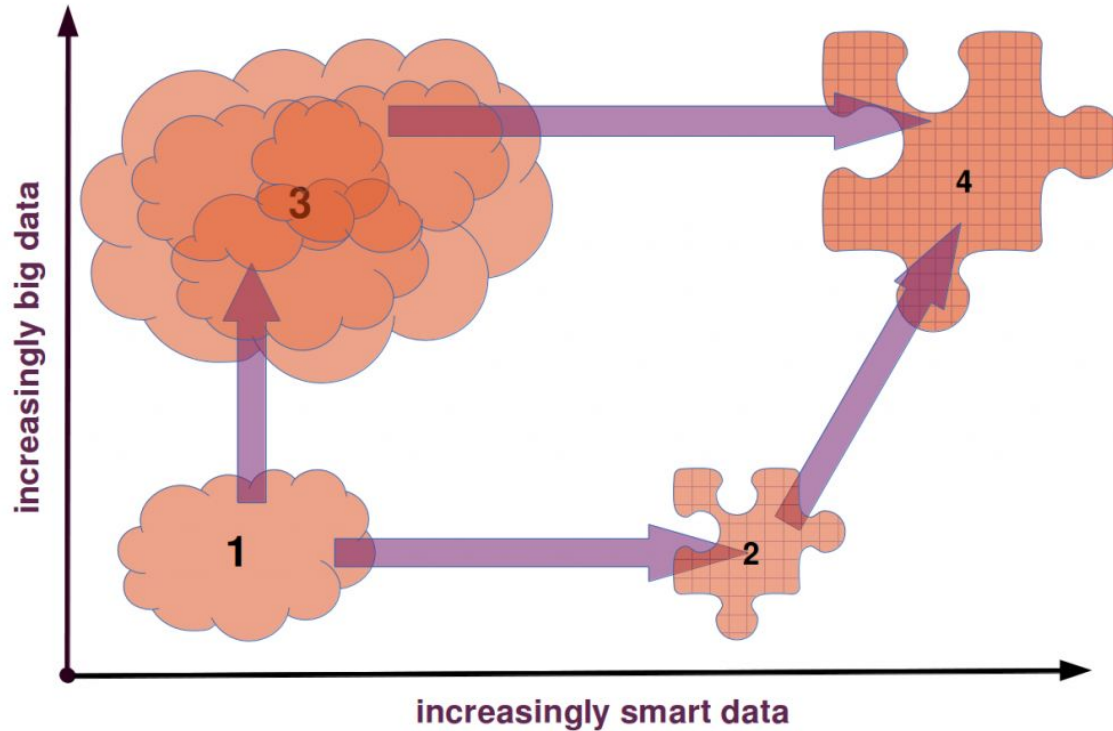
Annotation and discussion are the best way to deal with this problems.

Limited data in the humanities



Digitization and availability of British Nineteenth-Century Novels. Image published under a CC-BY license (Schöch).

The future of data in DH



From
Schöch
(2013)

Smart data

Markup, annotations and metadata

Clear data models

Clear relations to external entities

“Smart data to be semi-structured or structured, clean and explicit, as well as relatively small in volume and of limited heterogeneity” Schöch

Chinese Text Project

中文版

简体

About the site

Introduction

Font test page

Help us

Frequently Asked Questions

Instructions

Tools

System Statistics

Digital humanities

Pre-Qin and Han

Confucianism

Mohism

Daoism

Legalism

School of Names

School of the Military

Mathematics

Miscellaneous Schools

Histories

Ancient Classics

Etymology

Chinese Medicine

Excavated texts

Post-Han

Wei, Jin, and North-South

Sui-Tang

Song-Ming

Qing

Republican era

Notes

Resources

Dictionary

Discussion

Library

Wiki

Search

Pre-Qin and Han

for:

Search

Advanced

Title search:

Search

Log in

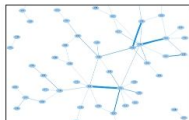
百諸家子 Chinese Text Project

Welcome

Welcome to the Chinese Text Project homepage. The Chinese Text Project is an online open-access digital library that makes pre-modern Chinese texts available to readers and researchers all around the world. The site attempts to make use of the digital medium to explore new ways of interacting with these texts that are not possible in print. With over thirty thousand titles and more than five billion characters, the Chinese Text Project is also the largest database of pre-modern Chinese texts in existence.

You may wish to [read more about the project](#), view the [pre-Qin and Han](#), [post-Han](#) or [Wiki](#) tables of contents, or [consult the instructions](#), [FAQ](#), or list of [tools](#). If you're looking for a particular Chinese text, you can [search for texts by title](#) across the main textual sections of the site.

網站有中、英文版本，也有繁、簡體版，可通過每頁左上角的連結隨時調整。



This site is provided as an open-access resource. If you use this site in your teaching or research, please ask your university library to [subscribe](#). Donations are also greatly appreciated.

Latest additions

Date	Content
2016-10-10	Harvard Yenching Library Chinese materials added Thanks to the support of Harvard Yenching Library, over 5 million pages of scanned materials from the Yenching Library collection have been added to the Library section of the site, including high quality images from the Chinese Rare Books Collection. Approximate transcriptions created using the ctext.org OCR procedure have also been added to the Wiki , making these materials full-text searchable. In future we hope to collaborate with other libraries to include materials from their Chinese language collections.
2015-07-02	Support for Unicode 8.0 adds new characters A new version of the Unicode standard has been released, defining thousands of additional rarely used and variant Chinese characters. Support for these has been added to the dictionary section of the site; to view these characters, please install the latest version of the Hanazono font . Many new characters belong to "CJK Extension E" - you can confirm system support for these from the Font Test Page .

Recent technical updates are listed in the ["Latest Updates"](#) forum. Please note that only major technical updates are listed here; for content updates and recently added texts (which occur daily and hourly), please [log in](#) and refer to the [Wiki Recently Updated](#) sections.

Recent discussion topics

Date	Content
2019-03-20 by oao1006	麻怎么读呢？ 谁知道麻怎么读的呀？最好有出处明示。[More...] [1 reply]
2019-03-17 by reimu106	注释混在原文里了 「陽虎，陽貨，魯季氏家臣也。天理人欲，不容並立。虎之言此，恐為仁之害於富也；孟子引之，恐為富之害於仁。」[More...]
2019-03-16 by good0812	《四庫全書》景上海涵芬樓藏宋刊本《資治通鑑》第20冊電子版有缺 《四庫全書》景上海涵芬樓藏宋刊本《資治通鑑》第20冊電子版有缺，電子版第10頁第11行「王莽將三」[More...]

Fonts and Compatibility

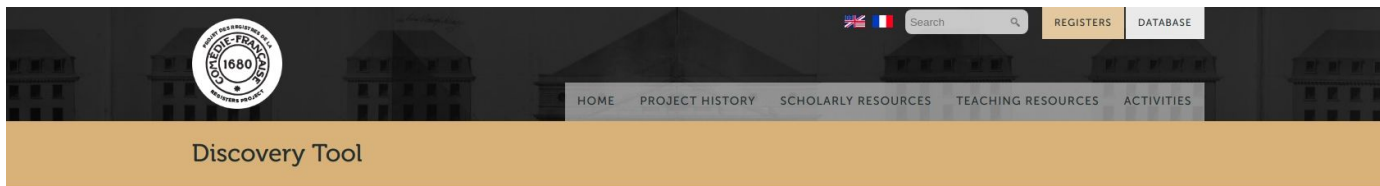
In order to view the Chinese text on this site, you will need to have a Chinese font and compatible web browser installed. In addition, because the ancient texts on this site incorporate many now obsolete and rarely-used characters, you will need to install a Unicode font which provides support for such characters (often referred to as CJK Extensions A through D). A suitable font is the freely downloadable ["Hanazono" font](#); ["Han Nom"](#) also includes many (though not all) of the required characters. To confirm whether your system properly supports these characters, please see the [Font test page](#). Proper support for Unicode (including so-called "surrogates") is also needed, which may cause additional problems for users with older operating systems.

The site is designed to comply to relevant W3C standards, and so should be usable in any standards-compliant web browser. However, as the emphasis on this site is on usability and not compatibility with legacy browsers, it is recommended that you upgrade to a browser with good CSS support such as Mozilla Firefox to get the most out of the site.



<https://ctext.org/>

The Comédie-Française Registers Project



Titre ^v	Auteur ^v	Genre ^v	Nombre de représentations ^v	Nombre de reprises ^v	date de la première
<input type="text" value="Rechercher un titre"/>	<input type="text" value="Rechercher un auteur"/>	<input type="text" value="Rechercher un genre"/>			
Tartuffe ou l'Imposteur	Jean-Baptiste Poquelin dit Molière	comédie	956	0	1680-06-03
Médecin malgré lui (Le) / Le Médecin forcé (titre alt.)	Jean-Baptiste Poquelin dit Molière	comédie	892	0	1680-10-31
École des femmes (L')	Jean-Baptiste Poquelin dit Molière	comédie	744	0	1680-06-25
Crispin médecin	Hauteroche (Noël Lebreton, sieur de)	comédie	735	0	1680-12-12
Esprit de contradiction (L')	Du Fresny (Charles)	comédie	690	0	1700-08-27
Avocat Patelin (L')	Brueys (David-Augustin de)	comédie	685	0	1706-06-04
Avare (L')	Jean-Baptiste Poquelin dit Molière	comédie	674	0	1680-05-14
George Dandin ou le Mari confondu	Jean-Baptiste Poquelin dit Molière	comédie	652	0	1680-05-16
Plaideurs (Les)	Jean Racine	comédie	645	0	1680-09-04
École des maris (L')	Jean-Baptiste Poquelin dit Molière	comédie	636	0	1680-05-16
Joueur (Le)	Regnard (Jean-François)	comédie	633	0	1696-12-19
Amphitryon	Jean-Baptiste Poquelin dit Molière	comédie	616	0	1680-06-06
Misanthrope (Le)	Jean-Baptiste Poquelin dit Molière	comédie	591	1	1680-05-27
Florentin (Le)	La Fontaine (Jean de)	comédie	564	0	1685-07-23
Comtesse d'Escarbagnas (La)	Jean-Baptiste Poquelin dit Molière	comédie	548	0	1680-05-25

<https://www.cfregisters.org/en/the-data/basic-tool>

Data-driven research

Network analysis, geospatial visualizations, time series analysis and textual analysis

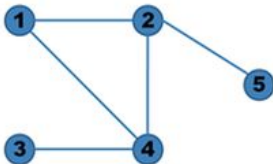
Networks

A network consists of:

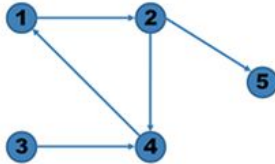
Nodes (things that are connected).

Edges (connections between those things). Specific, explicit connections between things. They can be directed or undirected.

Undirected Graph



Directed Graph



Examples: social networks, communication networks, citation networks, collaboration networks.

Network measurements

Degree. Total number of edges it has to other nodes.

Density. The portion of the potential connections in a network that are actual connections.

Average path length. The average steps (edges) to go from one node to another node.

Network analysis of wayang kulit characters

Digital Wayang Encyclopedia

About

Network analysis

Characters

Lakons (stories)

Search

all pages ▾

Werkudara

[\[full description\]](#)

Terms of address: Raden

Type: Human

Origin: India

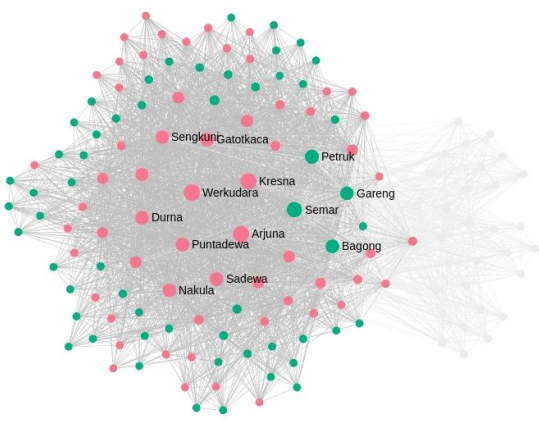
Alternative names: Gandawatrutmaja, Panduputra, Bayuputra, Bayusiwi, Bayusuta, Bayutanaya, Pawanasuta, Bimasena, Arya Sena, Sena, Bratasena, Ballawa, Bhima, Birawa, Dandun, Nagata, Kusumayuda, Kowara, Kusumadilaga, Wijasena

Mother: Kuni


Father: Bayu (Pandu)

Degree: 116
Weighted degree: 734
Eccentricity: 2
Closeness centrality: 0.83
Betweenness centrality: 321.31
Clustering: 0.35

The non-directed network visualization represents co-occurrences of characters in the scene-by-scene plot descriptions. The nodes (circles) indicate the different characters. Green nodes are Javanese in origin and pink ones are also found in the Indian versions of the stories. The size of the nodes corresponds to their degree (larger nodes have a higher number of connections). An edge (line) between two characters means they are present in the same scene. The thickness of the edges represents the number of times two characters appear together in the same scene. Clicking on a character's node reveals more information



Programming and information by [Miguel Escobar Varela](#). Research assistants: Hariyanto, Losheini Ravindran, F.X. Satriyo Nugroho, Yosephin Novi Marginingrum and Andi Wicaksono. Scientific advisor: Andrew Schauf.

 This work is licensed under a [Creative Commons Attribution 4.0 International License](#).

<https://villaorlado.github.io/wayangnetworks/html/canonical.html>



Network analysis of Javanese wayang kulit characters

	Percentage	Average Weighted Degree
Indian Characters	53%	153.47
Javanese Characters	47%	73.36

Textual analytics

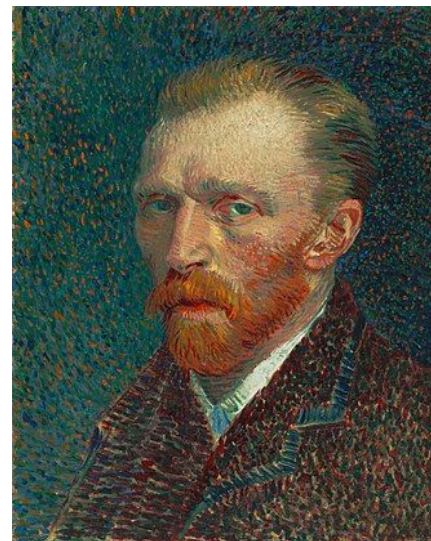
Many different approaches (most developed area of DH).

- Corpus linguistics
- Topic modelling
- Sentiment analysis

The screenshot displays the Corpus Workbench interface with three main panels:

- Left Panel (Word Cloud):** A word cloud visualizing the frequency of words in the corpus. Prominent words include "good", "write", "beautiful", "know", "hope", "evening", "hague", "dear", "things", "time", "paintings", "just", "wish", "day", "regards", "home", "painters", "glad", "little", "send", "don't", "letter", "london", "think", "soon", "i've", "come", "theo", "m", "regard", "soon", "i've", "come", "theo", "m", "regard", "soon", "i've", "come", "theo", "m".
- Center Panel (Document Text):** Displays the text of document `let001.xml`. The text is a letter from Theo to Vincent, dated 29 September 1872. It discusses the weather, the auction, and the family. The text is:

let001.xml
 The Hague, 29 September 1872. My dear Theo, Thanks for your letter, I was glad to hear that you got back safely. I missed you the first few days, and it was strange for me not to find you when I came home in the afternoon. We spent some pleasant days together, and actually did go for some walks and see a thing or two whenever we had the chance. What terrible weather, you must feel on your walks to. Yesterday there were postponed races on the occasion of the exhibition, but the illumination and fireworks were postponed because of the bad weather, so it's just as well you didn't stay to see them. Regards from the and the. I ever Your loving Vincent.
- Right Panel (Frequency Chart):** A line chart showing the relative frequency of words across documents. The x-axis is labeled "Corpus (Documents)" and the y-axis is "Relative Frequencies". The chart shows multiple lines for different words, with a legend indicating colors for "I'm", "good", "write", "it's", and "like". A specific data point for "write" is highlighted with a value of 0.0064835.



Analysis of letters by Vincent Van Gogh (1853-1890)

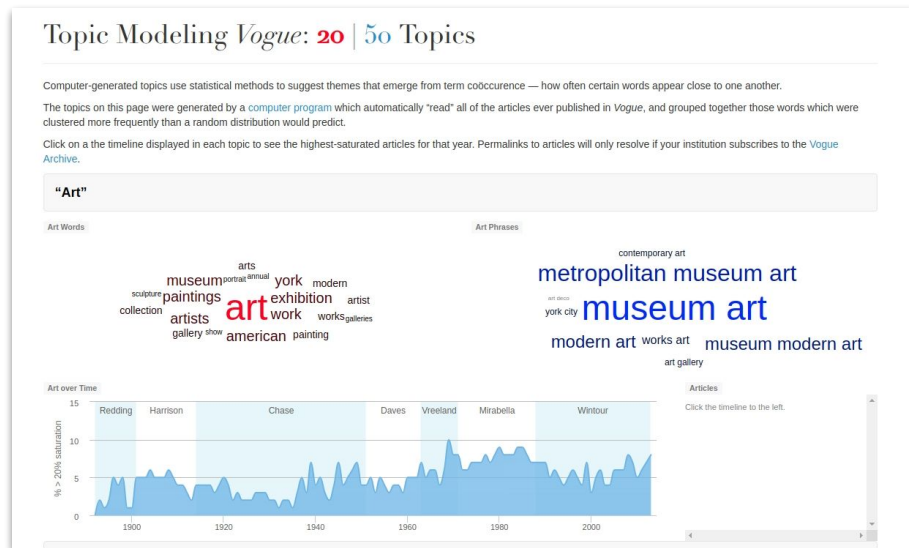
Robots reading vogue

Bookworm



<http://bookworm.library.yale.edu/>

Topic modelling



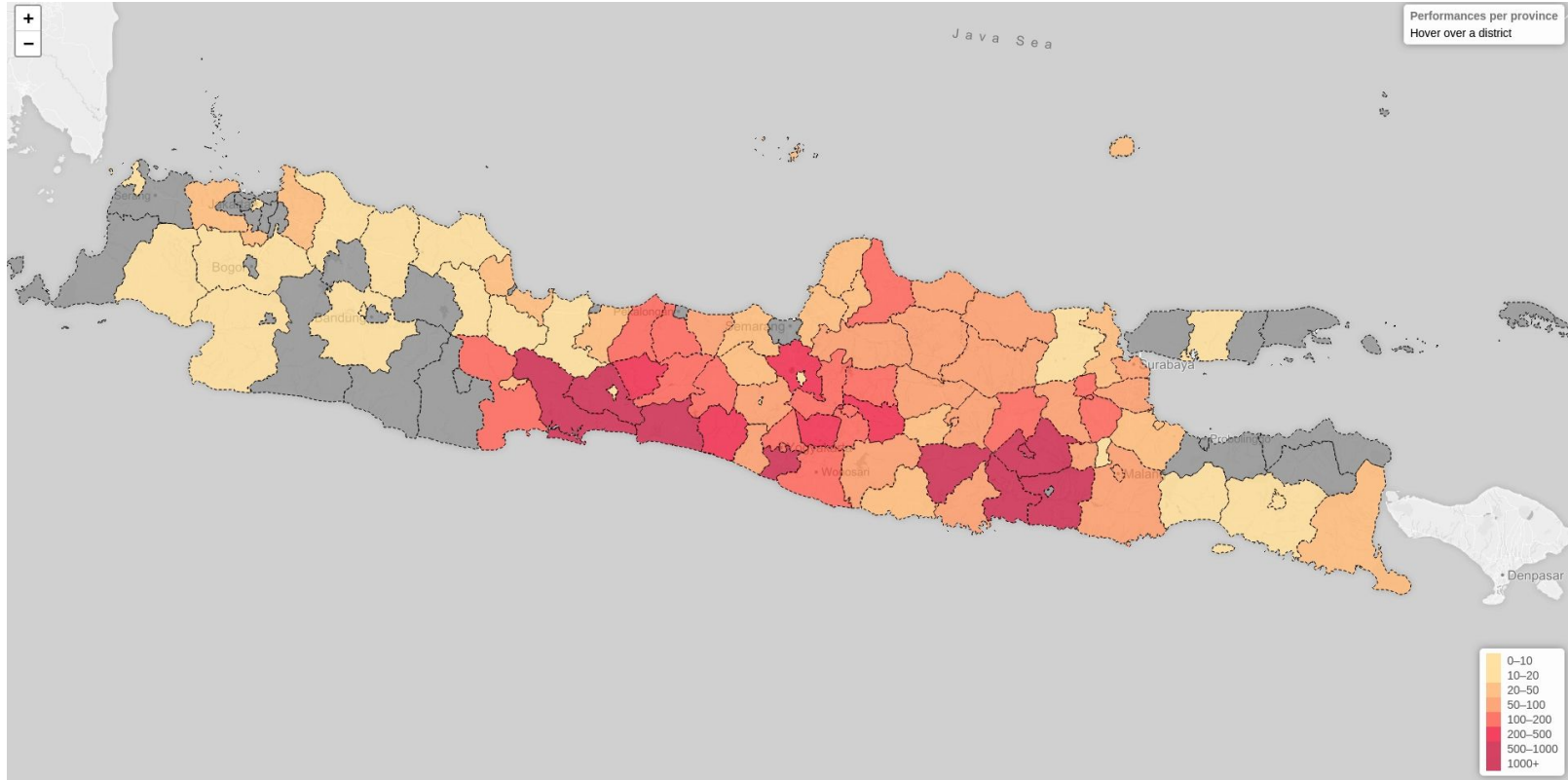
<http://dh.library.yale.edu/projects/vogue/topics/>

Geospatial analysis

Increasingly common approach

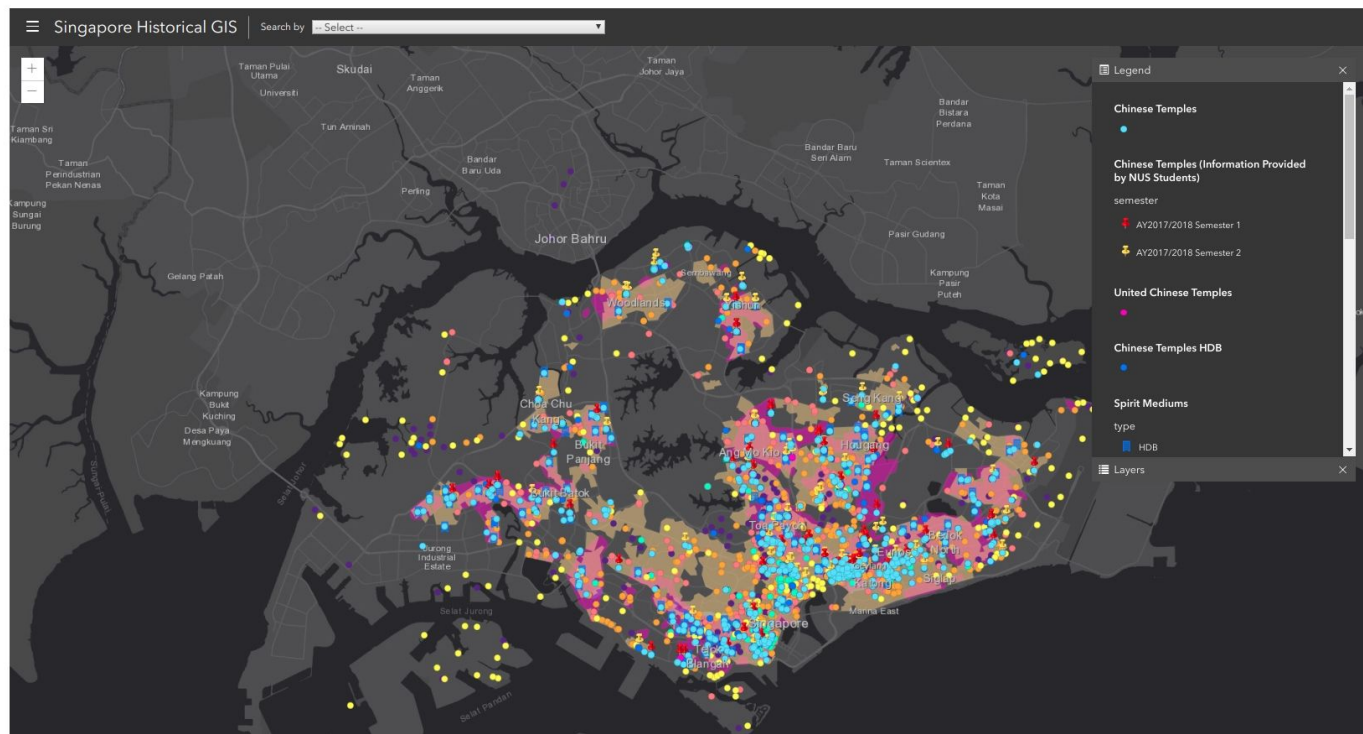
- Geospatial visualizations
- Geostatistics

Visualization of wayang kulit performances (Java, Indonesia)



Spatial distribution of Chinese Culture in Singapore

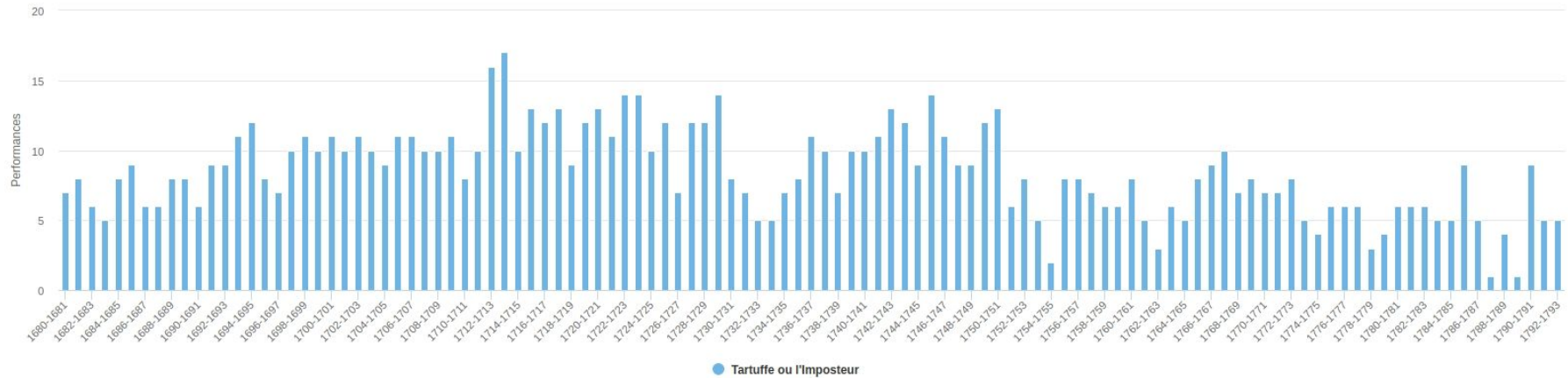
Spatial Distribution of Singapore Chinese Culture/新加坡华人文化空间分布



<http://shgis.nus.edu.sg/>

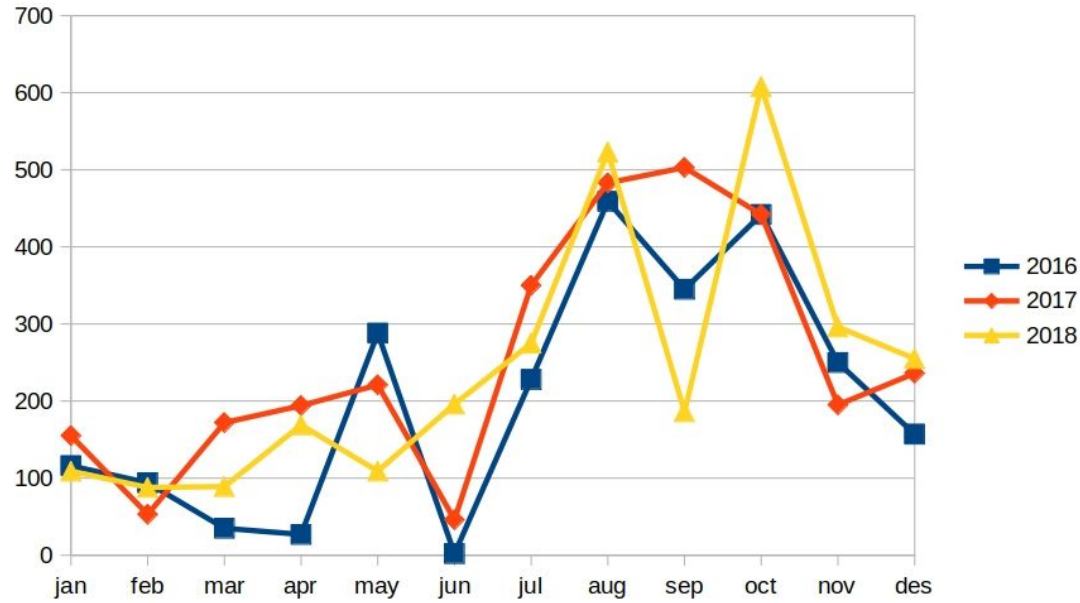
Time-series analysis

PERFORMANCES (956)



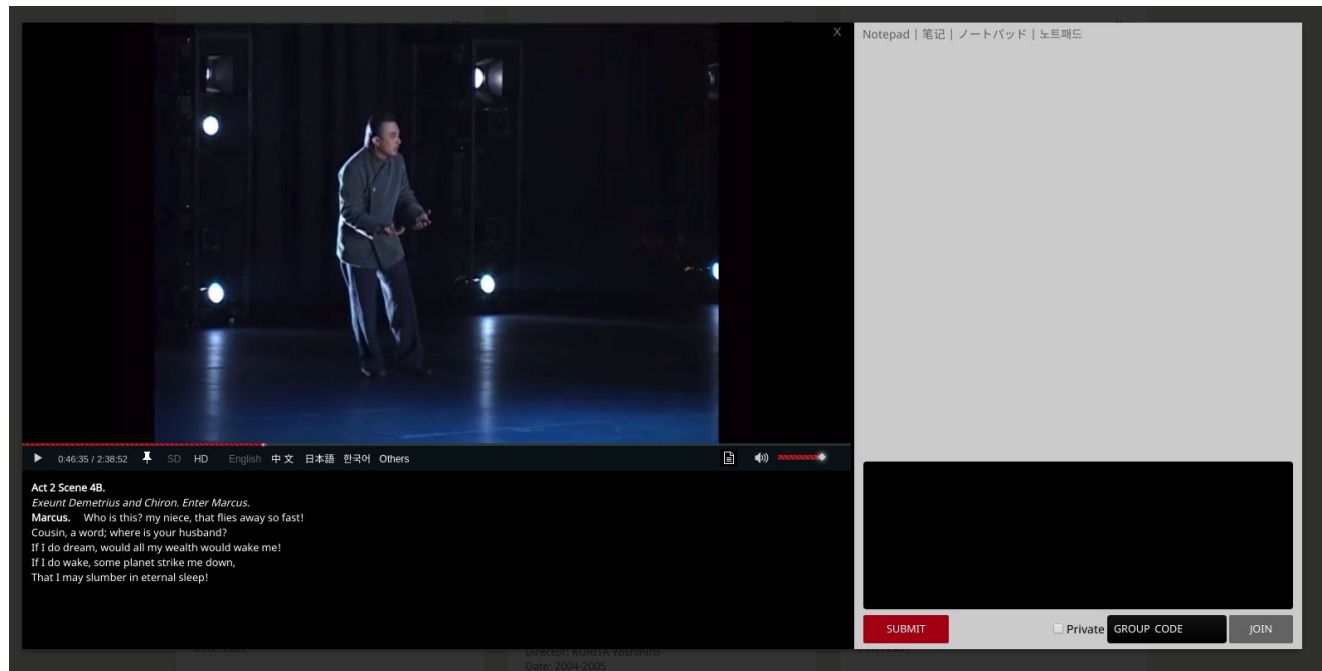
<https://www.cfregisters.org/en/the-data/basic-tool>

Javanese Wayang Kulit Performances Per Year



Digital publishing

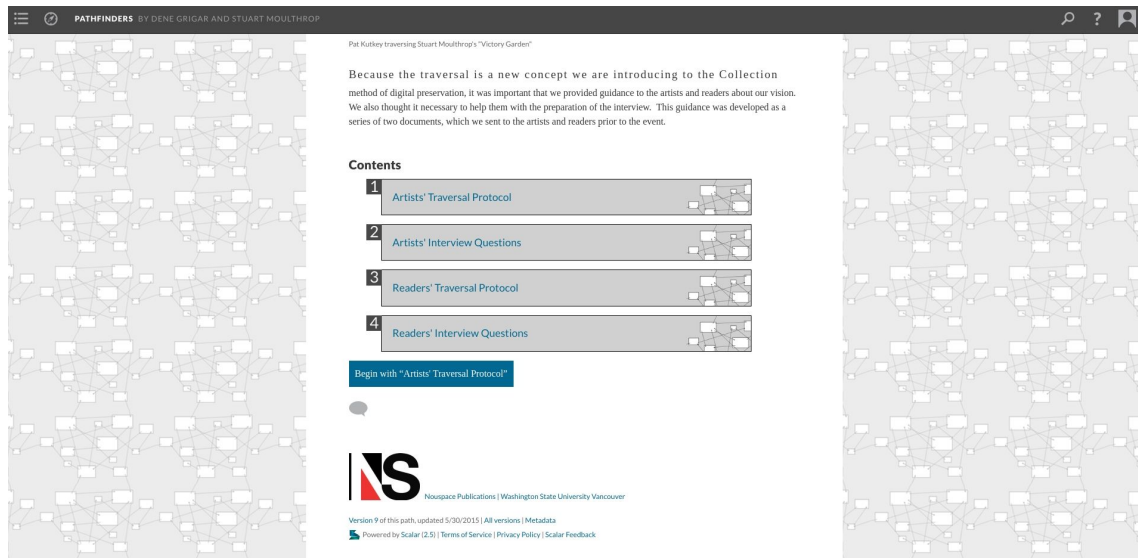
A|S|I|A



<http://a-s-i-a-web.org/en/productions.php>

Titus Andronicus, Hong Kong Arts Festival and No Man's Land
Director: TANG Shu-wing, Date: 2008

Pathfinders (project built with Scalar)



PATHFINDERS BY DENE GRIGAR AND STUART MOULTROP

Pat Kutley traversing Stuart Moulthrop's "Victory Garden"

Because the traversal is a new concept we are introducing to the Collection method of digital preservation, it was important that we provided guidance to the artists and readers about our vision. We also thought it necessary to help them with the preparation of the interview. This guidance was developed as a series of two documents, which we sent to the artists and readers prior to the event.

Contents

- 1 Artists' Traversal Protocol
- 2 Artists' Interview Questions
- 3 Readers' Traversal Protocol
- 4 Readers' Interview Questions

Begin with "Artists' Traversal Protocol"

NS
Nonpace Publications | Washington State University Vancouver

Version 9 of this path, updated 5/30/2015 | All versions | Metadata
Powered by Scalar (2.5) | Terms of Service | Privacy Policy | Scalar Feedback

<http://scalar.usc.edu/works/pathfinders/traversals-and-interviews-documentation>

The Chinese Deathscape

STANFORD UNIVERSITY PRESS

Documentation Web Archive Stanford Digital Repository

Edited by Thomas S. Mullaney

The Chinese Deathscape

Grave Reform in Modern China

ENTER >

In the past decade alone, more than ten million corpses have been exhumed and reburied across the Chinese landscape. The campaign has transformed China's graveyards into sites of acute personal, social, political, and economic contestation.

In this digital volume, three historians of China, Jeffrey Snyder-Reinke, Christian Henriot, and Thomas S. Mullaney, chart out the history of China's rapidly shifting deathscape. Each essay grapples with a different dimension of grave relocation and burial reform in China over the past three centuries: from the phenomenon of "baby towers" in the Lower Yangzi region of late imperial China, to the histories of death in the city of Shanghai, and finally to the history of grave relocation during the contemporary period, examined by Mullaney, when both its scale and

Edited by Thomas S. Mullaney, Stanford University Press

<http://chinesedeathscape.org/>

Challenges and suggestions

Challenges for libraries

Changing roles for people and spaces

Sustainability

Development support

Two ideas for discussion

- The library should be the home of DH in the university
- A simple distinction between researcher and librarian doesn't make sense in the 21st century

Some suggestions for libraries

Rethink the distinction between librarians
and academics

Rethink the library as a space for dialogue
and exchange

Thank you

Comments and questions

m.escobar@nus.edu.sg

miguelescobar.com

[@miguelJogja](#)
