International Joint Research and Training Program for Development and Application of Digital Humanities

「數位人文的發展與應用」國際培訓研習會

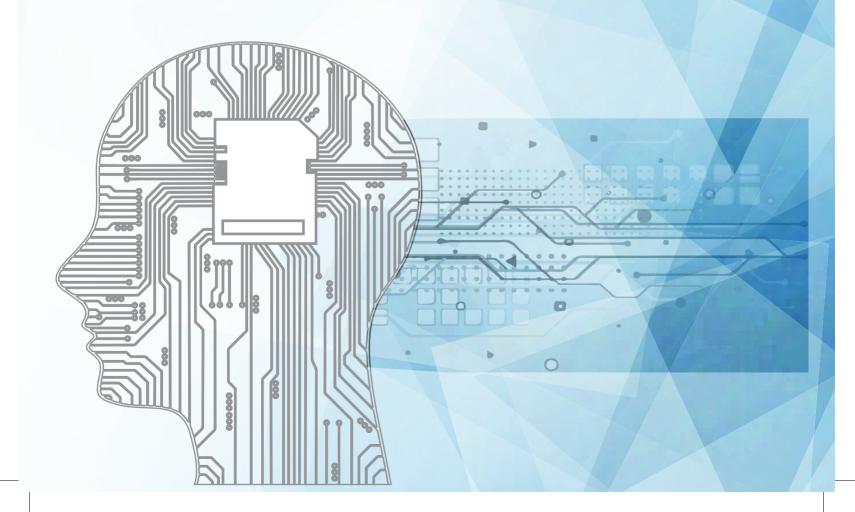


Table of Contents 目次

Introduction 會議簡介	3
Agenda 議程	5
Speaker Profile 講者簡歷	10
Lectures 課程內容	
♦A Journey from Digital Archives to Digital Humanities to	
Personal PlatformsJieh Hsiang	23
◆The View from Seattle: The Digital Humanities at the	
University of Washington ······Brian Reed	49
◆DocuSky: Your Powerful Personal	
Digital Humanities Platform ······Chi-jui Hu	81
• The Development of Information Platform for Humanities Computing:	
the Case of Center for Southeast Asian Studies (CSEAS)	
Kyoto University ······Shoichiro Hara	99
◆An Introduction of Social Analytical	
Approach to Digital Humanities	133
◆Digital Infrastructure, Bibliographic Data,	
and Chinese Print History Paul Vierthaler	153
♦Playing with CORPRO: a Chinese Text-mining Tool for	
Humanities Scholars·····Ho-chia Chueh	171
◆Taiwan Biographical Ontology (TBIO) and the	
Study of Taiwanese Elites Táňa	187
Dluhošová	
◆Development of Digital Humanities in the Field of	205
Buddhist Studies in Taiwan·····Jen-jou Hung	
◆Experiences and Perspectives – The Tools of Digital Humanities in	233
Taiwan and Their Future Che-chia Chang	
How to Develop a Digital Humanities Research Platform in an	050
Academic Library for Supporting Humanists' Research: An Experience	259
from National Chengchi University Libraries ······Chih-ming Chen	
1	

•Development and Trends of Digital Scholarship	
Services in Academic Libraries······Hao-ren Ke	287
Further Information 注意事項	321
List of International Participants 國際學員名單	323

International Joint Research and Training Program for Development and Application of Digital Humanities

Date: 28 October~1 November, 2019

Venue: International Conference Hall, National Central Library

No.20, Zhongshan South Road, Zhongzheng District, Taipei 10001, Taiwan (R.O.C.) Supervisor: Ministry of Education, Ministry of Science and Technology Organizer: National Central Library Conference Language: English Conference Website: http://workshop2019.ncl.edu.tw/

Introduction

National Central Library(NCL) has launched "International Workshop for Professional Librarians" once every two years for enhancing academic exchange among East Asian Libraries and Chinese Studies Institutes since 2009. Topics were "Sinological Resources Applications and Services" (2009), "Sinological Resources and Services" (2011), "Preservation and Innovation of Special Collections and Archives" (2013), and "From Ancient Times to the Present: Preservation of Ancient Chinese Books and Development of Contemporary Chinese Digital Resources" (2015), "Nation's Memories and Archives Management" (2017), respectively.

In 2019, the topics of Research and Training Program for Development and Application of Digital Humanities are organized according to the topics as follows:

- The introduction of digital humanities and the current developments of different countries
- ◆ The interdisciplinary research in the digital humanities
- The innovative services of digital humanities and libraries
- The application of digital technologies to the disciplines of humanities
- ◆ The introduction of digital humanities research and platforms.

We've invited 12 scholars and experts, at local and abroad, to share their experiences, research, and technology related to digital humanities, and look forward to creating new milestones in application of digital humanities through dialog and exchange.

The workshop receive over 50 applied or invited registrations of international participants from 11 countries. Most of them are executives of libraries or practitioners from various countries. It also opens to professional librarians and participants from libray and information science field in Taiwan.

After the lecture sessions, we arrange some cultural visits for international participants to National Taiwan Library on October 31st, Juming Museum and Tamsui Historical Museum on November 1st.

「數位人文的發展與應用」國際培訓研習會

會議時間: 2019年10月28日至11月1日

- 會議地點:國家圖書館國際會議廳
 - 臺北市中山南路 20 號
- 指導單位:教育部、科技部
- 主辦單位:國家圖書館
- 會議語言:英文
- 會議網頁:http://workshop2019.ncl.edu.tw/

會議簡介

國家圖書館為增進與全球東亞圖書館及漢學研究機構之交流,自 2009 年開始,每兩 年舉辦「國際圖書館專業館員研習班」,先後以「漢學研究文獻運用及服務」(2009 年)、 「漢學研究資源及服務」(2011 年)、「特藏文獻及檔案維護與創新」(2013 年)、「古與 今~古籍維護與當代華文電子資源創新發展」(2015 年)以及「國家記憶與檔案管理」(2017) 等為主題。

今(2019)年第六屆研習班,訂名為「數位人文的發展與應用」,以下列主題規劃 國際研習課程:

- ◆數位人文理論及各國發展現況
- ◆跨領域的數位人文研究
- ◆數位人文與圖書館創新服務
- ◆人文領域研究之相關數位工具應用
- ◆數位人文研究計畫及系統平台介紹

本次五天的研習會,邀請到12位國內外數位人文相關領域之學者專家,講述有關數 位人文之相關經驗、研究以及科技或工具,並期待透過彼此之對話與交流,再創數位人 文應用的新里程碑。

此次研習會有來自 11 個國家超過 50 位海外學員報名或受邀參加,大多為各國之圖 書館行政主管暨實務工作者;同時開放臺灣圖資領域界有興趣人士報名參加。

研習會後特別安排海外學員於 10 月 31 日參訪國立臺灣圖書館以及 11 月 1 日參訪朱 銘美術館及淡水古蹟博物館。

2019 International Workshop for Professional Librarians 2019 國際圖書館專業館員研習班

Agenda 議程

Day 1, 28 October 2019 (Monday) 2019 年 10 月 28 日 (星期一)

Time	Activity		
0900 - 0930	Registration / 報到		
0930 – 1000	Opening ceremony / 開幕典禮 Hosted by Director- General Shu-hsien Tseng, National Central Library 國家圖書館曾淑賢館長主持 Addresses by Honorable Guests / 貴賓致詞		
Time	Topic Speaker		
1000 – 1130	A Journey from Digital Archives to Digital Humanities to Personal Platforms 從數位典藏、數位人文到個人化平台 的歷程	Jieh Hsiang 項潔 Distinguished Professor in Computer Science & Director of Research Center for Digital Humanities, National Taiwan University 國立臺灣大學資訊工程學系暨研究所特聘教授暨 數位人文研究中心主任	
1130 - 1140	Break / 休息		
1140 - 1220	Visit National Central Library / 參訪國家圖書館		
1210 - 1340	Lunch / 午餐		
1400 – 1530	The View from Seattle: The Digital Humanities at the University of Washington 西雅圖觀點:華盛頓大學的數位人文	Brian Reed Divisional Dean of the Humanities & Professor of English and Comparative Literature, University of Washington, USA 美國華盛頓大學人文學院院長暨英文及比較文學 系教授	
1530 - 1550	Tea Break / 茶敘		
1550 – 1720	DocuSky: Your Powerful Personal Digital Humanities Platform DocuSky: 專為您打造的數位人文研究 平台	Chi-jui Hu 胡其瑞 Postdoctoral Researcher, Research Center for Digital Humanities, National Taiwan University 臺灣大學數位人文研究中心 博士後研究人員	
1800 - 2000	Welcome Reception / 歡迎晚宴	Special Program For Foreign Participants	

Day 2, 29 October 2019 (Tuesday) 2019 年 10 月 29 日 (星期二)

Time	Торіс	Speaker
0830 - 0900	Registration / 報到	
0900 – 1030	The Development of Information Platform for Humanities Computing: the Case of Center for Southeast Asian Studies (CSEAS) Kyoto University 數位人文資訊平台之建置:以日本 京都大學東亞地域研究所為例	HARA Shoichiro 原正一郎 Professor, Center for Southeast Asian Studies (CSEAS), Kyoto University, Japan 日本京都大學東南亞地域中心研究所教授
1030 - 1050	Break / 休息	
1050 – 1220	An Introduction of Social Analytical Approach to Digital Humanities 社會網路分析應用於數位人文	Muh-chyun Tang 唐牧群 Associate University Librarian & Professor, Library & Information Science, National Taiwan University 國立臺灣大學圖書資訊學系教授兼圖書館副館長
1220 - 1340	Lunch/午餐	
1340 - 1510	Digital Infrastructure, Bibliographic Data, and Chinese Print History 數位基礎建設、書目資料以及中國 印刷史	Paul Vierthaler Assistant Professor, Chinese Studies at the College of William & Mary, USA 美國威廉與瑪麗學院漢學研究助理教授
1510 - 1530	Tea Break / 茶敘	
1530 – 1700	Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars 玩庫博:人文社會學者的中文文本 探勘工具	Ho-chia Chueh 闕河嘉 Associate Professor, Department of Bio-Industry Communication and Development, National Taiwan University 國立臺灣大學生物產業傳播暨發展學系副教授

Day 3, 30 October 2019 (Wednesday) 2019 年 10 月 30 日 (星期三)

Time	Topic	Gy) 2017 年 10 万 30 日 (生約二) Speaker
0830 - 0900	Registration / 報到	
0900 - 1030	Taiwan Biographical Ontology (TBIO) and the Study of Taiwanese Elites 臺灣人物誌本體資料庫以及臺灣菁 英份子研究	Táňa Dluhošová 路丹妮 Research Fellow, Oriental Institute of the Czech Academy of Sciences, Czech 捷克科學院亞非研究所研究員
1030 - 1050	Break / 休息	
1050 – 1220	Development of Digital Humanities in the Field of Buddhist Studies in Taiwan 臺灣佛學研究領域的數位人文發展	Jen-jou Hung 洪振洲 Associate Professor, the Department of Buddhist Studies and Director of the Library and Information Center at Dharma Drum Institute of Liberal Arts (DILA) 法鼓文理學院佛教學系副教授暨圖書資訊館館長
1220 - 1320	Lunch / 午餐	
1320 - 1400	National Central Library → Academia Si 國家圖書館至中央研究院	nica
1400 - 1515	Visit Museum of Institute of History and 參訪中研院歷史文物陳列館	Philosopy
1515 - 1530	Break / 休息	
1530 – 1700	Experiences and Perspectives - The Tools of Digital Humanities in Taiwan and Their Future 經驗與展望一臺灣的數位人文建設 與發展方向 Visit Academia Sinica Center for Digital Cultures (ASCDC) 介紹中央研究院數位研究中心	Che-chia Chang 張哲嘉 Associate research fellow of Institute of Modern History & Deputy Director of Center for Digital Cultures, Academia Sinica 中央研究院近代史研究所副研究員暨數位文化中 心副召集人

Day 4, 31 October 2019 (Thursday) 2019 年 10 月 31 日 (星期四)

Time	Торіс	Speaker	
0830 - 0900	Registration / 報到		
0900 – 1030	How to Develop a Digital Humanities Research Platform in an Academic Library for Supporting Humanists' Research: An Experience from National Chengchi University Libraries 如何在學術圖書館發展數位人文研 究平台支援數位人文研究——國立 政治大學圖書館的經驗	Chih-ming Chen 陳志銘 University Librarian & Professor, Graduate Institute of Library, Information and Archival Studies, National Chengchi University 國立政治大學圖書資訊與檔案學研究所教授暨圖 書館館長	
1030 - 1050	Break / 休息		
1050 – 1220	Development and Trends of Digital Scholarship Services in Academic Libraries 學術圖書館數位化學術研究服務發 展與趨勢	Hao-ren Ke 柯皓仁 University Librarian & Professor, Graduate Institute Library & Information Studies, National Taiwan Normal University 國立臺灣師範大學圖書資訊學研究所教授暨圖書 館館長	
1220 - 1330	Lunch / 午餐		
1330 - 1440	Closing Ceremony & Awarding of Certi 閉幕典禮與頒發證書	ificate	
1440 - 1520	National Central Library → National Ta 國家圖書館至國立臺灣圖書館	uwan Library	
1520 - 1700	Visit National Taiwan Library 參訪國立臺灣圖書館		
1800 - 2000	Farewell Banquet 歡送晚宴	Special Program For Foreign Participants	

Day 5, 1 November 2019 (Friday) 2019 年 11 月 1 日 (星期五)

Time	Activity	
0830 - 0930	National Central Library → Juming Museum 國家圖書館至朱銘美術館	
0930 - 1140	Visit Juming Museum 參訪朱銘美術館	
1200 - 1330	Lunch / 午餐	Special Program For
1330 - 1430	Juming Museum → Tamsui Historical Museum 朱銘美術館至淡水古蹟博物館	Foreign Participants
1430 - 1830	Visit Tamsui Historical Museum 參訪淡水古蹟博物館	

Speaker Profile 講者簡歷



Jieh Hsiang 項潔 Distinguished Professor in Computer Science Director, Research Center for Digital Humanities, National Taiwan University 國立臺灣大學資訊工程學系 暨研究所特聘教授暨數位人文研究中心主任

Jieh Hsiang is a Distinguished Professor in Computer Science of the National Taiwan University and the founding director of its Research Center for Digital Humanities. Before returning to Taiwan in 1993, he was a full professor at Stony Brook University. He had also been a visiting professor at the University of Paris, University of Nancy and University of Milan. He was the founding dean of the College of Science and Technology of the National Chi-Nan University from 1998 to 2001. He also served as the University Librarian of NTU from 2002 to 2008, and the director of University Press of NTU from 2008 to 2016. He has been working on Digital Archives and Digital Humanities and served as its first president. Professor Hsiang has written over 200 scholarly papers and chapters, as well as 6 books on Digital Humanities.

項潔教授現任國立臺灣大學資訊工程學系暨研究所特聘教授 及國立臺灣大學數位人文研究中心主任。項教授於 1993 年返台 前,為紐約州立大學(石溪分校)計算機科學系專職教授,亦曾 擔任法國巴黎大學、法國南錫第一大學、義大利米蘭大學的客座 教授。1998 年至 2001 年,項教授擔任國立暨南國際大學科技學 院的創始院長,並曾於 2002 年至 2008 年擔任國立臺灣大學圖書 館館長,於 2008 年至 2016 年擔任國立臺灣大學出版中心主任。 項教授自 1996 年以來即持續投入於數位典藏及數位人文的相關 研究工作,並成立了臺灣數位人文學會,並擔任第一任理事長。 項教授已撰寫發表了 200 多篇學術論文,以及 6 本關於數位人文 領域的專書著作。



Brian Reed

Divisional Dean of the Humanities & Professor of English and Comparative Literature, University of Washington, USA 美國華盛頓大學人文學院院長暨英文及比較文學系教授

Brian Reed is the Dean of Humanities and the Milliman Endowed Chair in the Humanities at the University of Washington, Seattle. He is also a Professor of English and Comparative Literature.

He graduated summa cum laude from Harvard University in 1992, and he received a second BA, with first-class honors, from Oxford University in 1994. In 2000 he completed his Ph.D. in English and American Literature from Stanford University. He has been a Rhodes and a Fulbright Scholar, as well as a Visiting Scholar at Australian National University.

He is the author of three books—most recently *Nobody's Business: Twenty-First Century Avant-Garde Poetics* (2013)—and more than forty essays and articles on modern and contemporary poetry. A new book—*Out of the Pen: Essays on Poetry at Its Limits*—is forthcoming from the University of Alabama Press. He is currently researching inter- and multimedia poetries, especially those written in and for digital environments.

Brian Reed 教授現任美國華盛頓大學人文學院院長、 Milliman 特聘人文教授,以及比較文學系教授。Brian Reed 教 授於 1992 年以優異成績畢業於哈佛大學,並於 1994 年獲得牛 津大學第二等榮譽學士學位。2000 年獲得史丹佛大學英美文學 博士學位。他曾為羅德及傅爾布萊特獎助學者,以及澳洲國立 大學的訪問學者。Brian Reed 教授曾撰寫三本專書,最近的一 本著作為 «Nobody's Business: Twenty-First Century Avant-Garde Poetics»,並曾發表過四十多篇關於現代和當代詩歌的學術論文 及專文。最新撰寫的著作為 «Out of the Pen: Essays on Poetry at Its Limits»。Brian Reed 教授目前正在專注於研究多媒體和多媒 體詩篇,尤其是那些以數位形式呈現或於數位環境中創造出的 詩篇文本資料。



Chi-jui Hu 胡其瑞 postdoctoral researcher, the Research Center for Digital Humanities of National Taiwan University 國立臺灣大學數位人文研究中心博士後研究員

Dr. Chijui Hu is a postdoctoral researcher at the Research Center for Digital Humanities of Taiwan University. Dr. Hu was graduated from the Graduate Institute of Religious Studies from Chengchi University, Taiwan. His research was focused on the cultural and history of the minority groups in Southwest China, and now, he wants to put these research issue into the digital humanities tools to see what may happen.

胡其瑞博士,臺灣大學數位人文研究中心博士後研究人員。 政治大學宗教研究所畢業。胡博士的研究重心在中國西南少數民 族的宗教、歷史與文化,近年來則嘗試以數字人文的研究方法, 用以審視以往所進行的相關研究,以期能在原有的研究基礎上, 有所突破。



Hara Shoichiro 原正一郎 Professor in Center for Southeast Asian Studies (CSEAS), Kyoto University 日本京都大學東南亞地域中心研究所教授

Prof. Hara Shoichiro was graduated from the Department of Medical Science (Doctor Course), University of Tokyo on March 30, 1987, receiving his Doctor's Degree in Medical Science (Ph.D.). He joined at the National Institute of Japanese Literature(NIJL) since September 1, 1991, and was assigned to the Research Information Department as an Associate Professor, then moved to the Center for Integrated Area Studies, Kyoto University from August 1, 2006, and was assigned to the Division of Area Informatics as a Professor, serving in that position up to date. The former director of CIAS and the former deputy director of CSEAS, Kyoto University. Prof. Hara Shoichiro's research field includes humanities computing, geographic information systems, machine learning, bigdata, database, health informatics, etc.

原正一郎教授於 1987 年取得東京大學醫學博士學位。現任京 都大學東南亞地域中心教授。其曾任國立日本文學研究所(NIJL) 研究資訊所副教授、京都大學綜合區域研究中心主任及京都大學 東南亞地域研究中心副主任。原正一郎教授其研究領域包括:數 位人文、地理資訊系統、機器學習、大數據、健康資訊學 ... 等。



Muh-chyun Tang 唐牧群Associate University Librarian & Professor, Library &Information Science, National Taiwan University國立臺灣大學圖書資訊學系教授兼圖書館副館長

Prof. Muh-Chyun Tang graduated from Department of Advertising at National Chengchi University, and after that, he had worked at Radio Taiwan International as editor for one year. He received the M.A. degree from the University of North Carolina at Chapel Hill and the Ph.D from the Department o Library and Information Science at Rutgers University. Prof. Muh-Chyun Tang is currently is the Associate University Librarian & Professor, Library & Information Science, National Taiwan University, and he ever served as the executive editor of the Library Association of the Republic of China (Taiwan) and the Secretary-general of Chinese Association of Library & Information science education. His research areas include social network, scholarly communication, reading study, performance evaluation, and digital humanities.

唐牧群教授畢業於國立政治大學廣告系,曾於中央廣播電台 擔任新聞編譯約一年的時間,隨後即赴美深造,先後取得美國北 卡羅萊納大學教堂山分校新聞與大眾傳播碩士,以及美國羅格斯 大學傳播資訊暨圖書館學研究院博士學位。其現任國立臺灣大學 圖書館副館長暨臺灣大學圖書資訊學系教授,並曾任中華民國圖 書館學會執行編輯及中華資訊圖書資訊教育學會秘書長。研究專 長領域為社會網絡分析、學術傳播、閱讀研究、績效評估及數位 人文。



Paul Vierthaler Assistant Professor, Chinese Studies at the College of William & Mary, USA 美國威廉瑪麗學院漢學研究助理教授

Paul Vierthaler is currently an assistant professor of Chinese Studies at the College of William & Mary in the United States. From 2016 until August 2019, he was an assistant professor of the Digital Humanities at Leiden University in the Netherlands. While there, he helped found the Leiden University Centre for Digital Humanities. He has previously held a postdoctoral fellowship in digital humanities at Boston College and an An Wang postdoctoral fellowship at Harvard University. Paul also holds an MA in Asian Studies from Yale, and a BA in Chinese Language from the University of Kansas. In 2014 he

was awarded a PhD in East Asian Languages and Literatures from Yale University. During the 2013-2014 academic year he held a Fulbright Fellowship, which he held at the Academia Sinica Institute for Chinese Literature and Philosophy. Paul has published a number of articles that feature digital analysis of Chinese literary history. He has published a methods paper that outlines how to detect intertextuality in large collections of Chinese documents and an article on the stylistic relationships among historical and fictional texts written in late imperial China in the Journal of Cultural Analytics. He has also published an article on using large bibliometric datasets to study Chinese printing in the Harvard Journal of Asiatic Studies, the first digital humanities article published in that journal.

Paul Vierthaler 教授目前為美國威廉瑪麗學院漢學研究的助 理教授。自 2016 年到 2019 年 8 月,曾任荷蘭萊頓大學數位人文 助理教授。於萊頓大學任職期間,Vierthaler 教授協助建立了萊 頓大學數位人文中心。在此之前,他曾獲美國波士頓學院數位人 文博士後獎學金,並曾在哈佛大學擔任博士後研究員。Vierthaler 教授擁有耶魯大學亞洲研究碩士學位和堪薩斯大學漢語學士學, 2014 年,獲得耶魯大學授予東亞語言和文學博士學位。在 2013-2014 學年期間,亦獲得了中央研究院中國文學與哲學研究所的 傅爾布萊特獎學金。Vierthaler 教授發表了許多以中國文學史其 數位資料分析為主題的文章。他曾發表過一篇概述如何檢測大 量中文文獻其文本脈絡的關聯性的學術論文,以及曾於 Harvard Journal of Asiatic Studies 期刊,發表關於中國晚期歷史和虛構文 本間其文體關係的文章,亦曾於哈佛大學亞洲研究期刊上發表了 一篇關於使用大量的書目計量數據資料來研究中國印刷的文章, 也是該期刊發表的第一篇有關數位人文文章。



Ho-chia Chueh 闕河嘉 Associate Professor, Department of Bio-Industry Communication and Development National Taiwan University

Ho-chia Chueh is currently Associate Professor at Department of Bio-Industry Communication and Development, National Taiwan University. Her research interests fall under four main areas: Rural Studies, Digital Humanities, Ethical issues of Agriculture, Philosophy of Education. Trained as an educational philosopher, Ho-chia's early work was focused on critical deconstruction of theories of identity through post-strucutralist thinking. She published a book in this field, *Anxious Identity* from the Praeger Publishers. Later, she applied poststructuralist analysis to the field of rural studies, in particular examining how new agricultural concepts are represented in social texts. She is now focused on humanities approach to agricultural practices and rural activities. Ho-chia is also interested in digital humanities. She has developed a corpus analysis software CORPRO that enables humanities/social scientists to 'independently' conduct text-mining analysis of a large amount of texts. CORPRO is designed in the consideration of special contexts of Chinese language use.

闕河嘉教授為人文社會學家,現任國立台灣大學生物產業傳 播暨發展學系副教授。她的主要研究領域有鄉村研究、數位人 文、農業倫理、教育哲學。闕河嘉教授以符號學分析媒體再現的 鄉村意象,因而踏進數位人文領域。

她開發了「庫博中文語料庫分析工具」(CORPRO)。「庫 博」是人文學者的玩具,不需要會寫程式的人也可以獨立對大量 文本進行文本探勘分析,藉而一窺語料庫展現出模式。當代圖書 館數位化許多文學典籍,包括小說、傳記、書信,也訂購許多 數位形式的資料,包括新聞資料庫、電子雜誌等。她認為人文研 究學者可以根據研究興趣去自建語料庫,藉由玩「庫博」進行初 步文本分析而獲取研究靈感,或檢驗先前的論點和假說。她結合 數位人文工具和鄉村研究,譬如瞭解當代新農業概念的形成與發 展。



Táňa Dluhošová 路丹妮 Research Fellow, Oriental Institute of the Czech academy of Sciences 捷克科學院亞非研究所研究員

Prof. Táňa Dluhošová is Research Fellow, Oriental Institute of the Czech academy of Sciences, Editor in Chief, Archiv orientální (AHCI, Scopus). Her previous employment incudes Assistant Professor of Center for Chinese Studies, Department for the Study of Religions, Faculty of Arts, Masaryk University, Project manager of Center for the Study of China and Japan Religions, Department for the Study of Religions, Faculty of Arts, Masaryk University. And she have received the Academic Awards includes 2nd prize in the Young Scholar Award granted by the European Association of Taiwan Studies for the article: Taiwan funü in the Early Post-War Period (1945–1949) Literary Field, and Annual Prize of French-Taiwanese Cultural Foundation of the French Académie des sciences morales et politiques as a recognition of efforts for a better understanding of Taiwan in Central and Eastern Europe.

The area of her interest includes Digital humanities, History and Culture of Taiwan, Chinese literature, Sociology of Literature, Early Postwar Taiwanese Literature (1945-49), Ideology and language.

路丹妮 (TáňaDluhošová) 教授是捷克科學院東方研究所的研 究員,也是 Archiv orientální (AHCI, Scopus)的主編。其曾擔 任馬薩里克大學藝術學院宗教研究系中國研究中心助理教授以及 中國和日本宗教研究中心專案經理。路丹妮教授曾榮獲多項學術 研究獎,其中包括歐洲臺灣研究協會頒發的青年學者獎,以及法 國年度獎-由法蘭西學院人文政治科學院的臺灣文化基金會所頒 發。路丹妮教授主要專長研究領域包括數位人文,臺灣歷史與文 化,中國文學,文學社會學,戰後早期臺灣文學(1945-49), 意識形態與語言…等。



Jen-jou Hung 洪振洲

Associate Professor in the Department of Buddhist Studies and Director of the Library and Information Center at Dharma Drum Institute of Liberal Arts (DILA), Taiwan 法鼓文理學院佛教學系副教授暨圖書資訊館館長

Jen Jou (aka "Joey") HUNG 洪振洲 was born in Taiwan in 1976. He received his Ph.D. in 2006 from the National Taiwan University of Science and Technology, Taipei. Now, he is Associate Professor in the Department of Buddhist Studies and Director of the Library and Information Center at Dharma Drum Institute of Liberal Arts (DILA), Taiwan. He is currently engaged in a variety of digital archives projects at DILA. His research interests include authorship attribution for ancient Buddhist literature, the construction of digital archives, and digital text processing.

洪振洲教授於 2006 年於臺灣科技大學資訊管理系取得博士 學位,目前任教於法鼓文理學院佛教學系,為專任副教授,並兼 任圖書資訊館館長一職。洪教授目前參與許多由法鼓法鼓文理學 院執行的數位典藏計畫。他的研究興趣包含漢譯佛典作譯者分 析、數位典藏專案建構、數位人文研究資源開發與數位文字處理 議題。



Che-chia Chang 張哲嘉

Associate research fellow of Institute of Modern History & Deputy Director of Center for Digital Cultures, Academia Sinica

中央研究院近代史研究所副研究員暨數位文化中心副召集人

Prof. Che-chia Chang is Associate Research Fellow of Institute of Modern History, Academia Sinica. He received his PhD degree from University of Pennsylvania with a dissertation on the Imperial physician-patient relationship in Empress Dowager Cixi's era from 1874 to 1908, under the supervision of Professor Nathan Sivin. Since then he has worked on a variety of topics on history of science in China and Japan, including materia medica, popular medical education, astrological history, history of cartography, drug commercials, forensic medicine, bone-setting, acupuncture, and the translations of anatomy. For conducting those researches he had been Visiting Fellows of International Research Center for Japanese Studies, Wellcome Trust Center for History of Medicine, Harvard-Yenching Institute, and IKGF (Internationales Kolleg für Gesiteswissenschaftliche Forschung) of the Universität Erlangen-Nürnberg. In 2008 he won the "Zhu Kezhen Award" from International Society for History of East Asian Science Technology and Medicine with the paper "The Myth of Rhubarb: The Strategic Rationale and Cultural Implications of China's Prohibitions on the Export of Rhubarb to Britain and Russia in the Qing Period". His current major research project is about the Manchu texts Ge ti ciowan lu bithe with which the Jesuit Dominique Parrenin (1663-1741) tutored the Kangxi Emperor (reign 1661-1722) to introduce the Western anatomy and pathology. When he studied in University of Pennsylvania, he learned computing research methods from Professor Robert Hartwell and first touched Hartwell's biographical database of the Chinese Medieval history (later developed into today' s CBDB). Since 2011 he was assigned the duty of supervising the Database projects of the Institute of Modern History, Academia Sinica, and then reformed the scattered databases into an integrated system MHDB (Modern History DataBases) and pushed its best openness to the users all over the world. Other than that, he drives the constructions of new databases such as the early modern English-Chinese Dictionaries. In 2015 he was appointed Director of the Archives, Institute of Modern History, Academia Sinica to upgrade the Archives' digitalization. In the next year he was appointed Deputy Director of Digital Cultural Center to promote digital humanities in Academia Sinica.

張哲嘉教授出生於臺北市,1988年於美國賓夕法尼亞大學

亞洲與中東研究系取得博士學位,師從席文教授,學位論文主題 為慈禧太后與宮廷御醫之間的醫病關係。畢業後即進入中央研究 院近代史研究所工作,主要研究領域為中國與日本之醫學、科學 史,如本草史、通俗醫學、占星學史、地圖學史、醫療廣告、針 灸史、骨科史、法醫史以及西方解剖學的翻譯等。他曾經在愛知 大學、國際日本文化研究中心、英國維康醫學史研究中心、哈 佛燕京學社以及埃爾朗根 - 紐倫堡大學的國際科學學術院訪問研 究。於 2008 年,他獲得國際東亞科學技術與醫學史學會頒贈之 「竺可楨學術論文獎」,論文題目為〈大黃迷思:清代制裁西洋 禁運大黃的策略思維與文化意涵〉。目前進行之主要課題為耶穌 會傳教士為康熙皇帝講授之西洋解剖學、病理學滿文教本《格體 全錄》。於美國求學期間,因郝若貝教授學習到以數位科技從事 歷史研究的方法,並接觸到中古中國傳記資料庫(今 CBDB 之 前身)。2011年受命執掌近史所之資料庫業務,將原本各自為 政的各庫整合成近代史數位資料庫(MHDB),並主導建置如《英 華字典》等新型資料庫。2015 掌近史所檔案館,從事檔案數位 化之升級。次年被任命為中央研究院數位文化中心副召集人,推 動數位人文研究。



Chih-ming Chen 陳志銘

Professor, Graduate Institute of Library, Information and Archival Studies at National Chengchi University (NCCU), Taiwan.

University librarian of NCCU.

Director of Social Sciences and Information Center of NCCU. Editor-in-chief of journal of InfoLib & Archives 國立政治大學圖書資訊與檔案學研究所教授暨圖書館館長

Chih-Ming Chen is currently a professor in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University (NCCU), Taipei, Taiwan. Now he is also the university librarian of NCCU, the director of Social Sciences and Information Center of NCCU, and the editor-in-chief of journal of InfoLib & Archives. He ever served as the chair of the graduate institute of library, information and archival studies, associate university librarian, associate director of computer center at NCCU. He received B.Sc. and M.Sc. degree from the Department of Industrial Education at National Taiwan Normal University in 1992 and 1997, and received Ph.D. degree from the Department of Electronic Engineering at National Taiwan University of Science and Technology in 2002. His research interests include e-learning, digital humanities, digital reading, digital library, data mining, machine learning, and intelligent agents on the web. In academic honors, he won the Memorial Award of Mr. Wu Dayu of Ministry of Science and Technology in Taiwan, Distinguished Professor of NCCU, 2 Research Outstanding Awards of NCCU, 9 Research Excellence Awards of NCCU, 9 Ministry of Science and Technology Awards for Special Talents, etc. So far, he has published over 80 journal papers in local and international areas. Among the 80 journal papers, there are 40 journal papers with SSCI index and 20 journal papers with SCI index.

陳志銘教授現任國立政治大學圖書資訊與檔案學研究所教授、圖書館館長、社會科學資料中心主任、圖資與檔案學刊主編。 他曾擔任國立政治大學圖書資訊與檔案學研究所所長、圖書館副 館長、電算中心副主任。他分別於 1992 年和 1997 年獲得國立台 灣師範大學工業教育學系的學士和碩士學位,並於 2002 年獲得 國立台灣科技大學電子工程系博士學位。他的研究興趣包括數位 學習、數位人文、數位閱讀、數位圖書館、資料探勘、機器學習 和網際網路智慧型代理人。在學術榮譽上曾獲得科技部吳大猶先 生紀念獎、國立政治大學特聘教授、2 次國立政治大學研究特優 獎、9 次國立政治大學研究優良獎、9 次科技部獎勵特殊優秀人 才等殊榮。到目前為止,他已在當地和國際領域發表了 80 多篇 期刊論文。在 80 篇期刊論文中,有 40 篇為具有 SSCI 索引的國 際期刊論文,20 篇為 SCI 索引國際期刊論文。



Hao-ren Ke 柯皓仁

University Librarian & Director of Center for Education Research and Evaluation

Professor, Graduate Institute of Library & Information Studies, National Taiwan Normal University

President, Library Association of the Republic of China (Taiwan)

國立臺灣師範大學圖書資訊學研究所教授兼圖書館館長 教育研究與評鑑中心代理主任 中華民國圖書館學會理事長

Hao-Ren Ke was born on June 29, 1967 in Taipei, Taiwan, Republic of China. He received his B.S. degree in 1989 and Ph.D. degree in 1993, both in Computer and Information Science, from National Chiao Tung University, Taiwan, R.O.C. Now he is a professor of Graduate Institute of Library and Information Studies and the University Librarian at National Taiwan Normal University (NTNU). He is the President of the biggest library association in Taiwan – Library Association of the Republic of China (Taiwan) (LAROC), and the

former Presidents of Interlibrary Cooperation Association (ILCA) and Chinese Association of Library and Information Science Education (CALISE).

His main research interests include library performance measurement, digital library/archives, library and information system, information retrieval, Web mining, and digital humanities. By using his technical expertise, he devotes himself to helping academic and research libraries in Taiwan move toward digital libraries. For example, he built mirror sites for many important A&I databases including SCI/SSCI, Ei Compendex in late 1990s. He developed the Interlibrary Loan Networking System for Taiwanese Libraries, which significantly expedites the interlibrary loan process in Taiwan. He also plays an important technical role in the operation of the biggest Taiwan academic & research library consortium– CONCERT (CONsortium on Core Electronic Resources in Taiwan).

Hao-Ren Ke has been participating in TELDAP (and the forerunner NSC projects on digital archives) for nearly ten years. The digital museums Web sites created by his team focus on Taiwanese native artists, including YuYu Yang (sculptor), Hsin-Chin Liu (cartoonist), Tai-Hsiang Li (musician), and Cloud Gate (dancing theatre).

柯皓仁教授,1967年次,1989年取得新竹交通大學資訊科 學系學士學位、1993年取得新竹交通大學資訊科學系博士學位。 目前為臺灣師範大學圖書館館長與圖書資訊學研究所教授,亦是 臺灣最大的圖書館協會-中華民國圖書館學會理事長,並曾任館 際合作協會與中華圖書資訊學教育學會理事長。柯皓仁教授的研 究興趣包含:圖書館評鑑、數位圖書館、數位典藏、圖書館資訊 系統、資訊檢索、資料探勘、鏈結資料等。藉由柯皓仁教授其專 業之資訊科技背景,多年來持續協助臺灣圖書館界邁向數字化圖 書館。在世紀交替前後,柯博士於臺灣建置了包含 SCI/SSCI, Ei Compendex, CSA 等索摘資料庫系統的鏡射站; 柯博士尚為臺灣 圖書館界設計與開發館際互借系統,對臺灣圖書館界的館際互 借業務具有重大貢獻;此外,柯博士在臺灣最大的圖書館聯盟 CONCERT (CONsortium on Core Electronic Resources in Taiwan) 亦 提供許多技術上的諮詢。柯皓仁博士過去十年在數位典藏方面著 力甚深,其所帶領的團隊主要著重於臺灣重要藝術家與團隊的典 藏數位化,包含楊英風(雕刻家)、劉興欽(漫畫家)、李泰祥(音 樂家)、朱銘(雕刻家)、雲門舞集(表演藝術)、優人神鼓(表 演藝術)等。



A Journey from Digital Archives to Digital Humanities to Personal Platforms

從數位典藏、數位人文到個人化平台的歷程

Jieh Hsiang 項潔

Distinguished Professor in Computer Science & Director of Research Center for Digital Humanities,

National Taiwan University

國立臺灣大學資訊工程學系暨研究所特聘教授暨數位人文研究中心主任

Abstract 摘要

The past twenty years witnessed a massive digitization effort of high quality primary research materials. Indeed, in Chinese language materials alone there are now several billion characters full-text of Chinese classics available for free over the Web in addition to digital libraries containing tens of millions of archival images and metadata. The availability of "big data" at the fingertip has changed the way humanities scholars conduct their research. It has also inspired IT researchers and library scientists to think deeper about how to provide the type of services that can better facilitate this massive quantity of data into new humanities research need.

In this talk we describe the evolution of relationship between digitization and humanities research in Taiwan during the past twenty years. We shall start from the National Digital Archives Program - a national project of digitization of cultural artifacts and its several precursors. Under NDAP, many high quality digital libraries of research significance are produced. In order to harness these cultural treasure troves for research, concepts of digital humanities, in particular a notion of textual context, have been developed and incorporated into the digital archives so that, in addition to providing information, digital libraries can also be used for *context discovery*, an important aspect of humanities research. As more tools and contents are developed, the need for better interaction between scholars, tools, and contents also emerged. This has lead to the so-called digital humanities research platforms. The most important

difference between a digital platform and a digital library is that the content and functions of a platform are not tightly coupled. In other words, one can select content as one wishes and choose applicable tools as one deems fit. A platform also allows the digitization process to move from the institutional level collections to personal materials. We shall illustrate our view by showing how different stages of development of DH affect the life cycle of humanities research.

過去二十年見證了針對高質量的研究資料進行大規模數位化工作。實際 上,在中文資料中,除了包含數千萬個檔案圖像和詮釋資料的數位圖書館外, 現在可以通過網路近用數十億字符的中文經典全文。唾手可及的"大數據"與其 可用性,改變了人文學者研究的方式。此外,此亦激發了資訊科技研究人員和 圖書館學家更加深入思考,如何提供可以更好且促進大規模數據融入於新型態 的人文科學研究需求及服務。

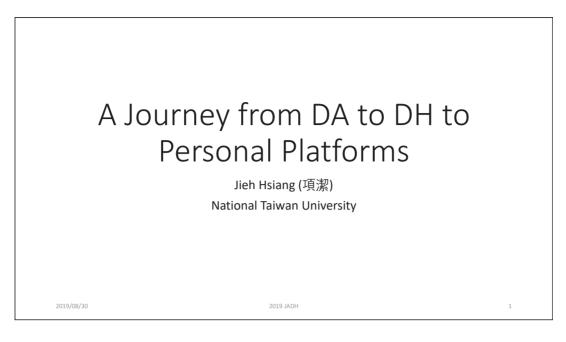
在本次演講中,將為大家說明過去二十年來,臺灣數位化與人文研究之間 其關係之演變。我們將從國家型數位典藏計劃開始談起-這是一個文化及藝術作 品之數字化及相關前身的的國家型計畫項目。在國家型數位典藏計畫下,建置 了許多具有重要研究意義的高質量數位圖書館。為了利用這些文化寶庫進行研 究,開發了數字人文的概念(特別是文本語意的概念)並融合於數位典藏資料 庫。

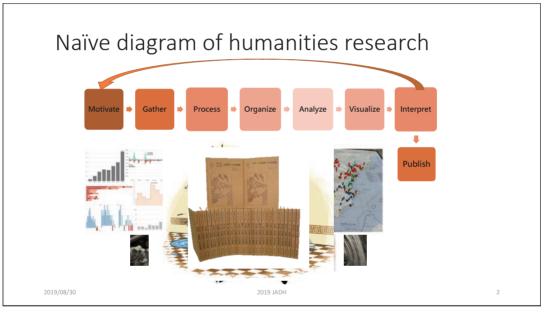
因此,除了提供資訊之外,數位圖書館也可應用於文本語意上下文脈絡的 探究,作為人文研究的一個重要面向。隨著更多工具和內容的開發,學者、工 具和內容之間擁有更良好互動的需求與期待也隨之出現。也因此,數位人文研 究平台便於此情境下因運而生。數位人文研究平台和數位圖書館之間最重要的 區別是,平台的內容和功能間的關係具個人化的彈性。換句話說,可以根據使 用者的需求選擇內容,並選擇適合的工具。此外,平台亦允許數位化過程由機 構層級的資料集轉移至個人的資料。我們將透過展示數位人文的不同發展階段 如何影響人文研究的生命週期來說明我們的觀點。

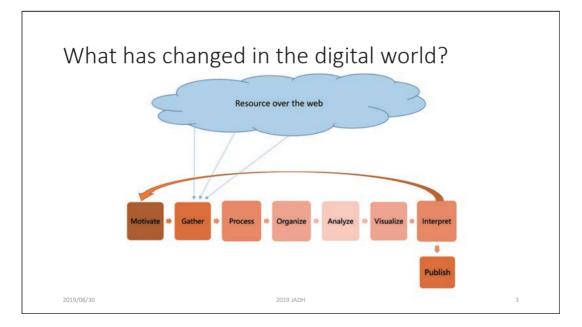
24

A Journey from Digital Archives to Digital Humanities to Personal Platforms

從數位典藏、數位人文到個人化平台的歷程



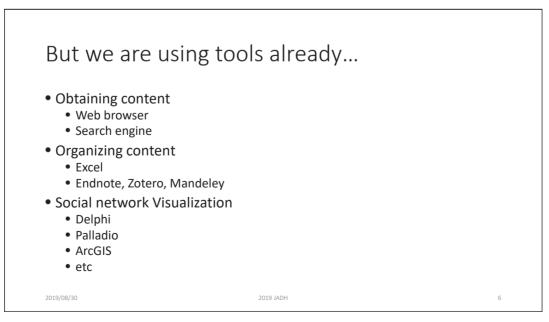




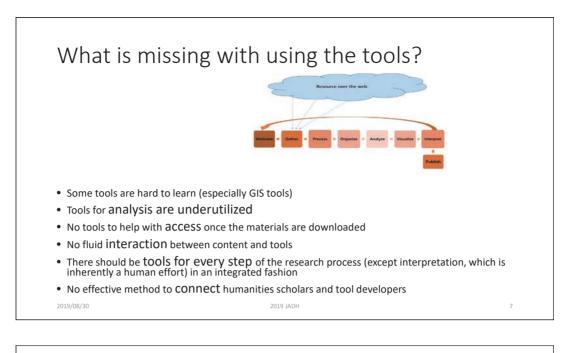
2019/08/30







從數位典藏、數位人文到個人化平台的歷程



Digital resources = digital contents + digital tools

Two main subjects in this talk

- Connecting digital contents and digital tools
- Connecting humanities scholars and IT researchers

Assumptions

- Contents are already in digital form
 Going from analog to digital is important, but that's a diffe
 - Going from analog to digital is important, but that's a different issue

2019 JADH

- Work mainly with Chinese texts, not images
 - The principle is the same, but tools are different

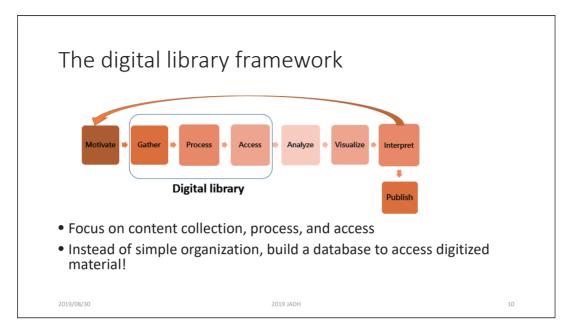
2019/08/30

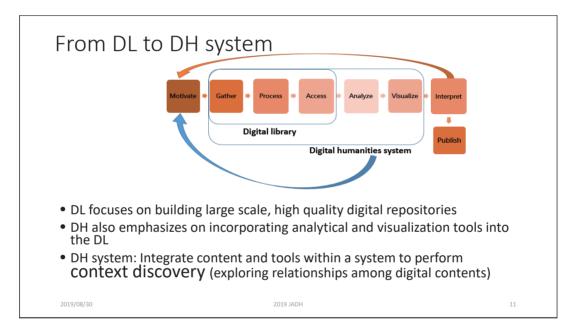
The digital library approach to develop scholarly content

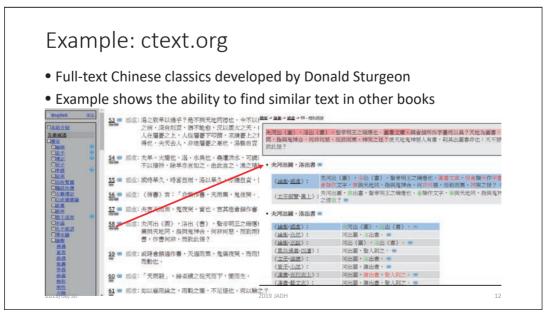
- Bypass individual research problem/motivation
- Data Collection
 - Comprehensively and systematically collect material of a specific domain (usually with institutional support)
- Data Processing
 - Organize and digitize material according to an SOP
- Data Access
 - Build a large scale digital library system
 - Access through browse and keyword search

2019/08/30

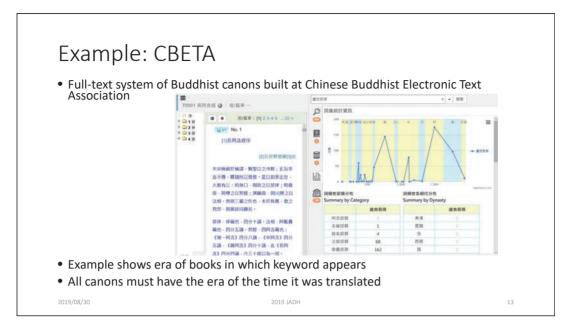
2019 JADH



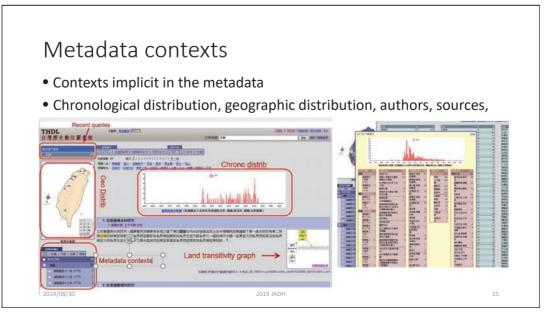


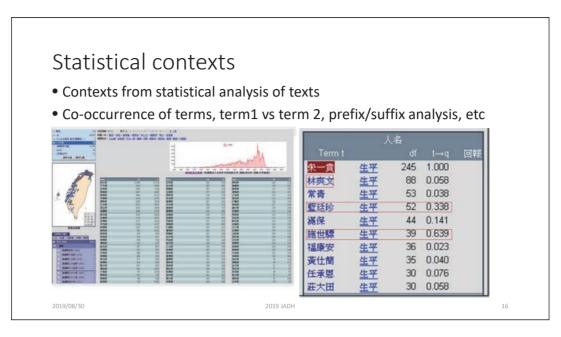


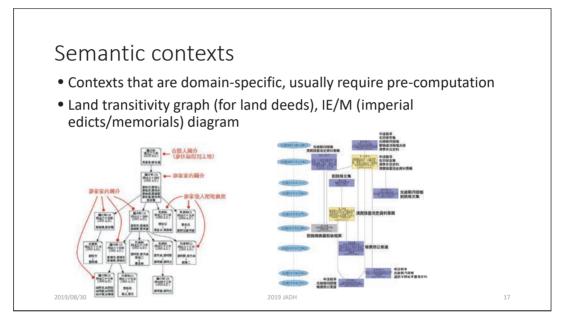
從數位典藏、數位人文到個人化平台的歷程

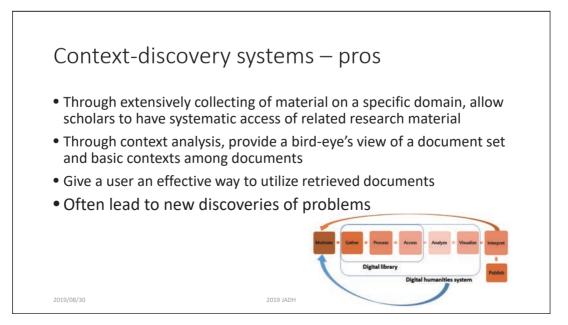




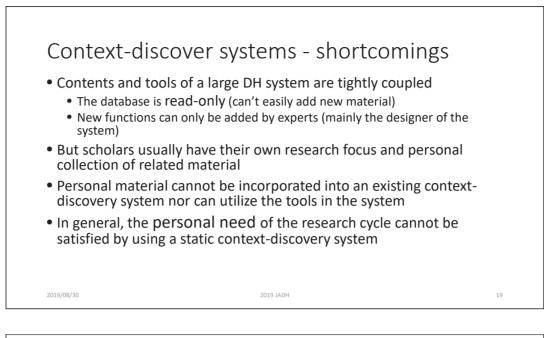


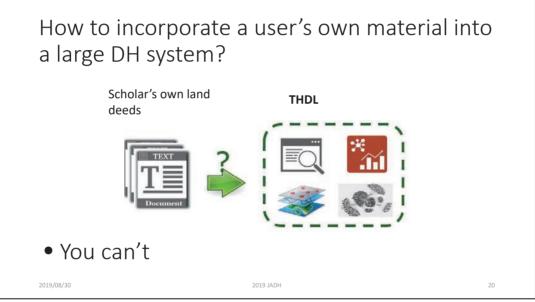


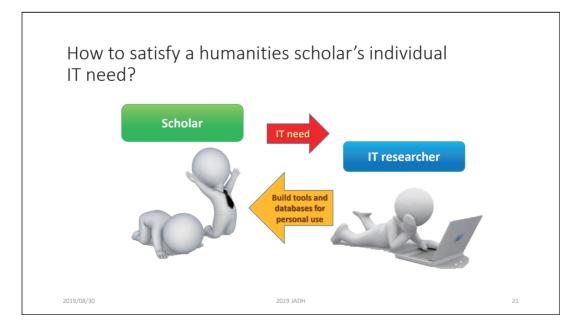




從數位典藏、數位人文到個人化平台的歷程









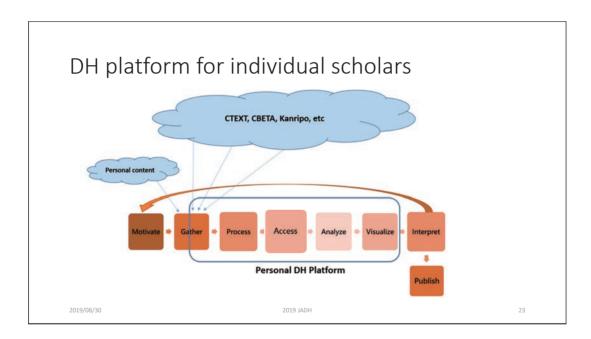
- Tools to process (own and downloaded) material into proper form
- Data access
 Simple way to build a personal DL to provide easy search and access of personal material
 Data analysis

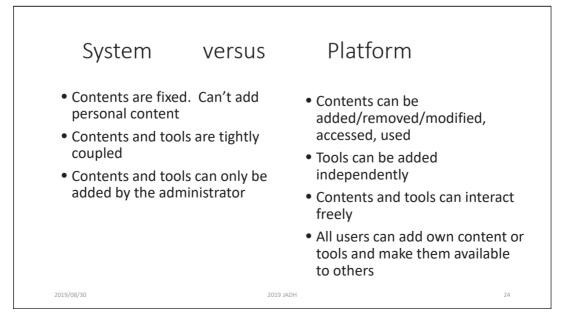
2019 JADH

- Data analysis
 - Tools to analyze material and to find textual context
- Data visualization

2019/08/30

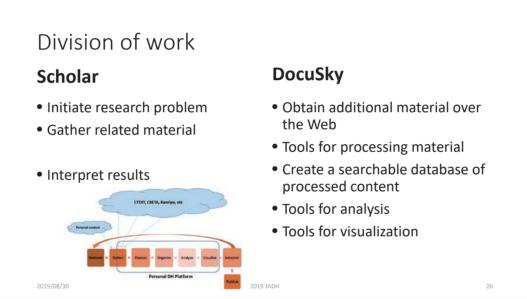
• Tools to provide visual presentations of analytic results

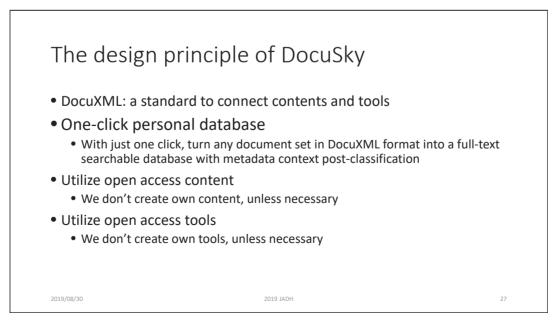


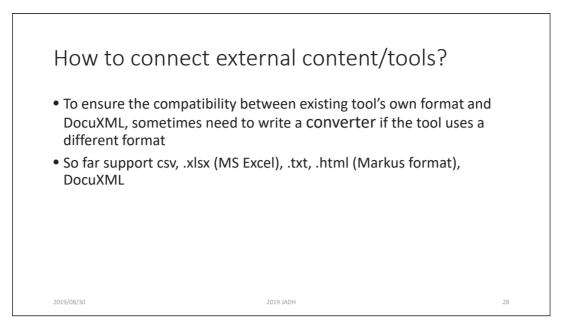


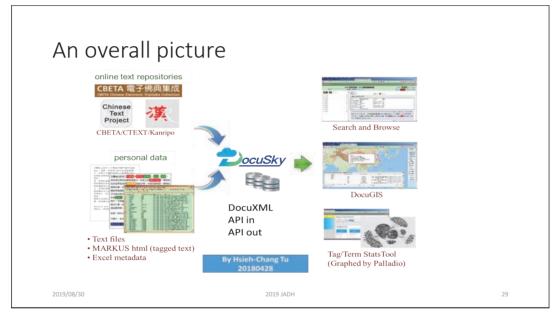
從數位典藏、數位人文到個人化平台的歷程

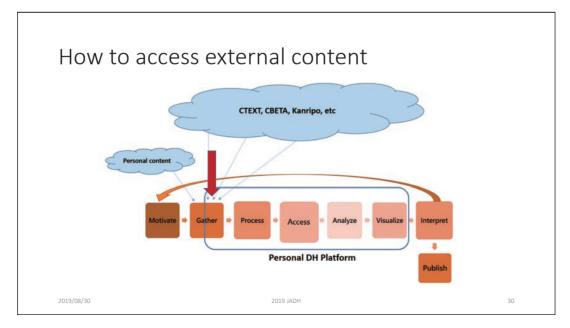




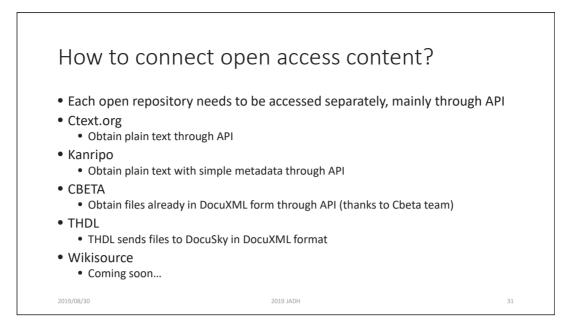


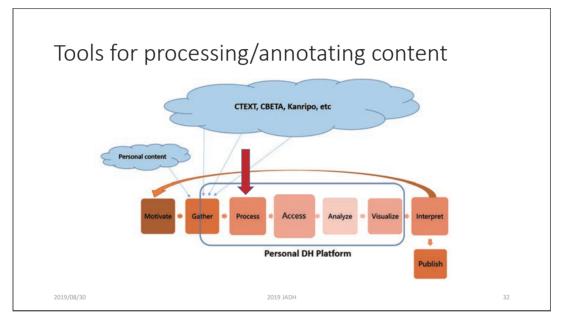


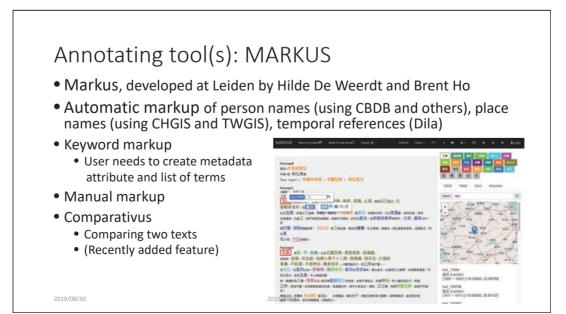


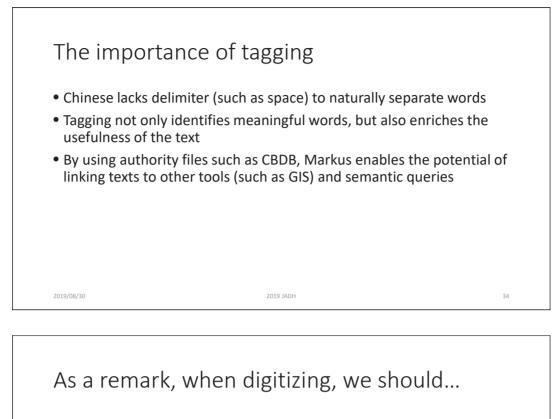


從數位典藏、數位人文到個人化平台的歷程

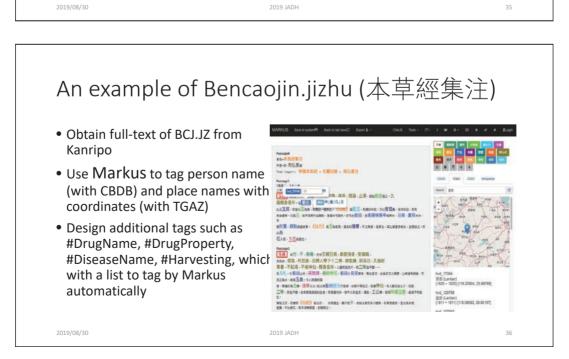




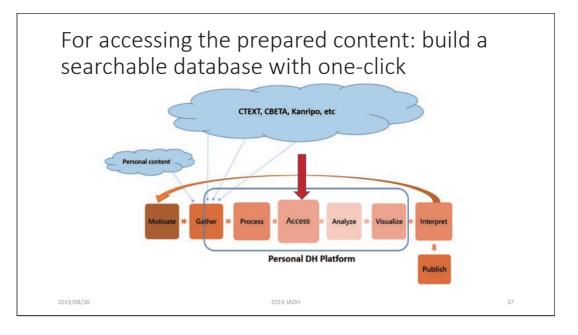


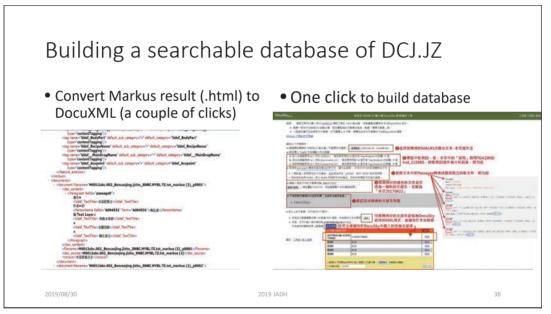


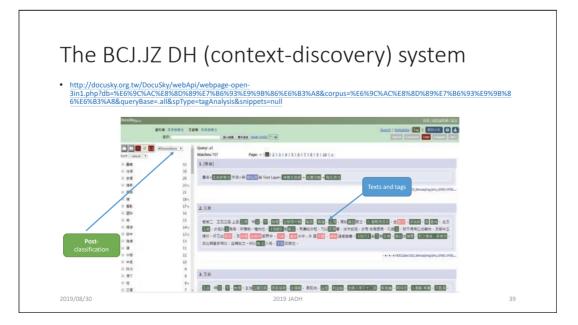
- Build metadata
- Produce full-text when possible
- Include image when possible
- When possible, tag person/place names, normalize dates, tag objects, identify geographic coordinates

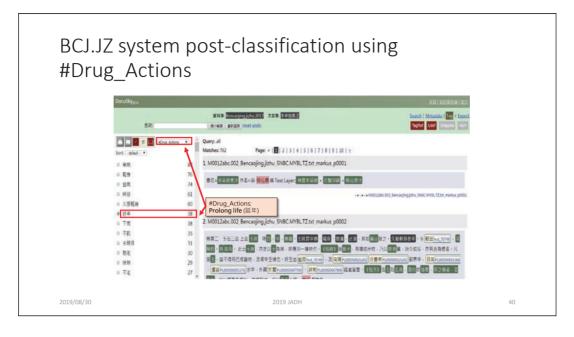


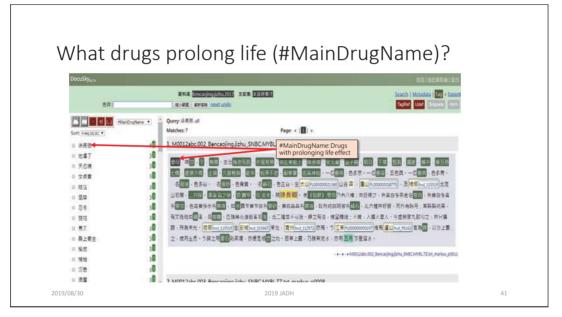
從數位典藏、數位人文到個人化平台的歷程

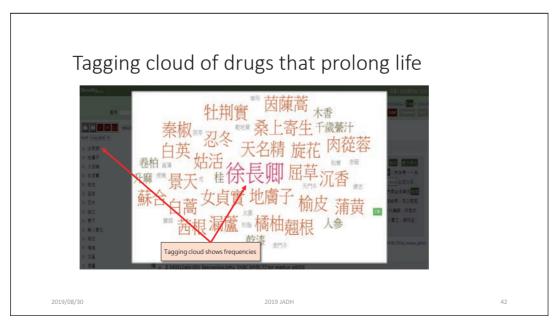






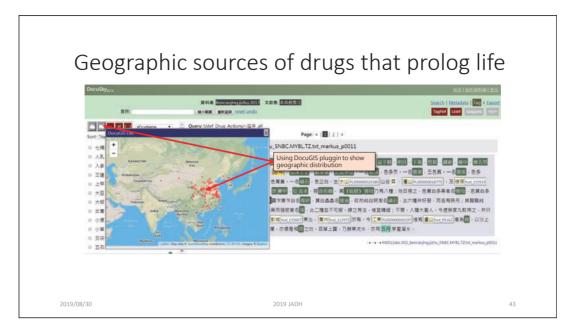


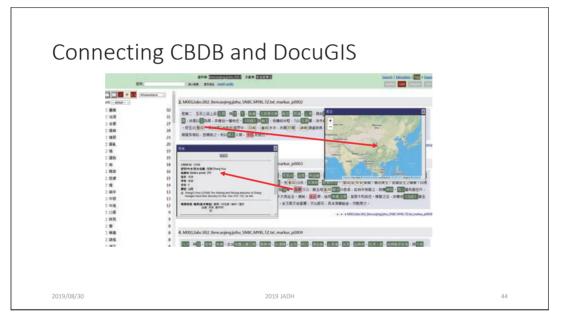


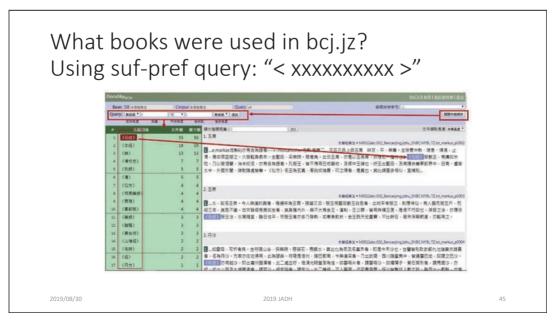


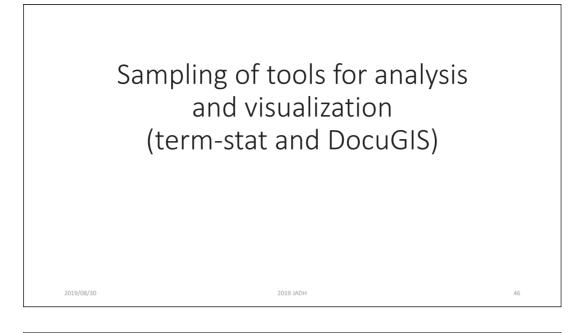
A Journey from Digital Archives to Digital Humanities to Personal Platforms

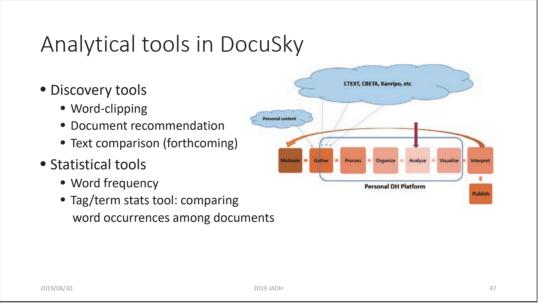
從數位典藏、數位人文到個人化平台的歷程

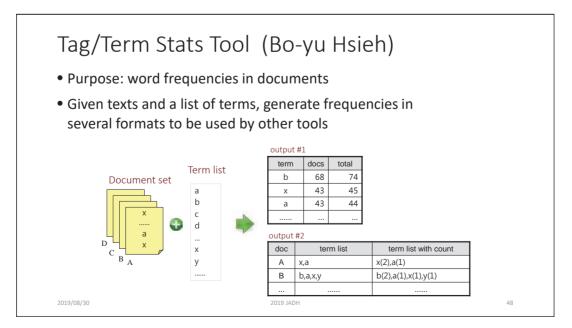




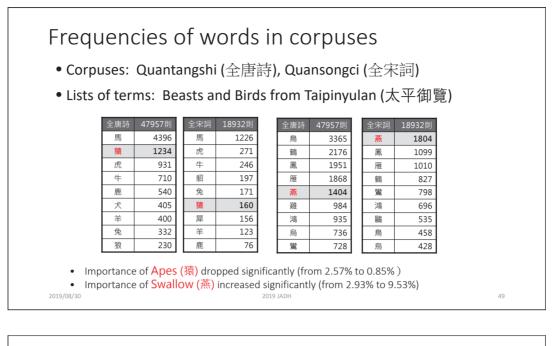


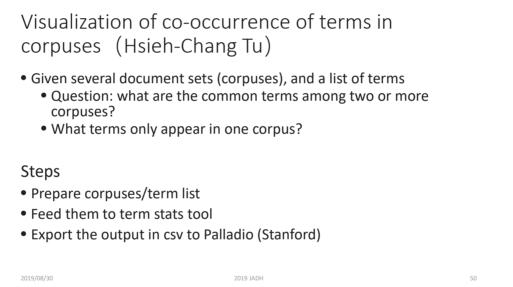


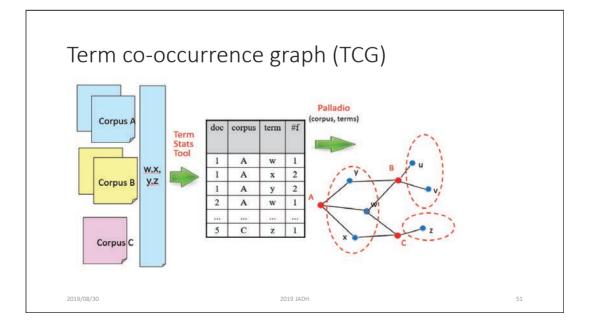


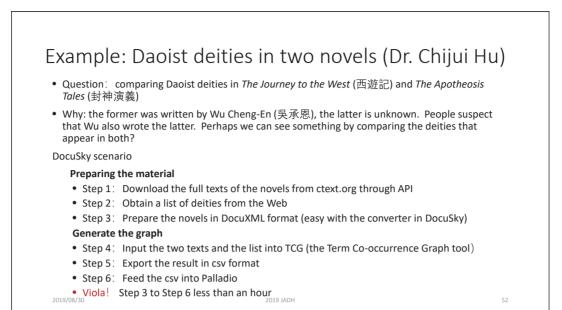


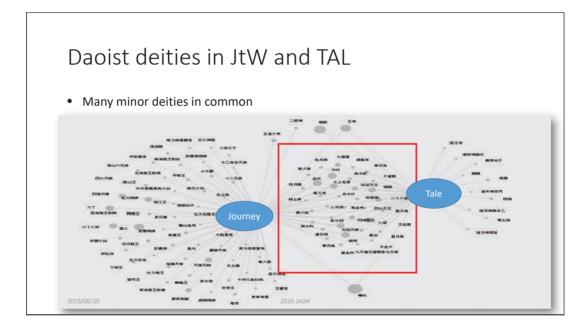
從數位典藏、數位人文到個人化平台的歷程

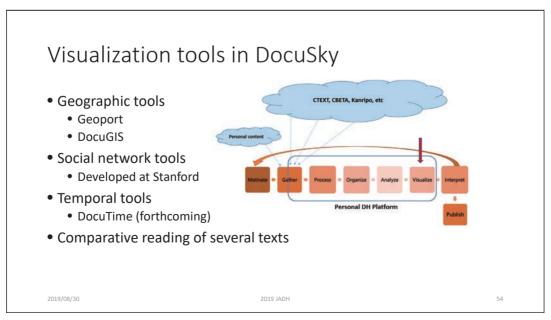






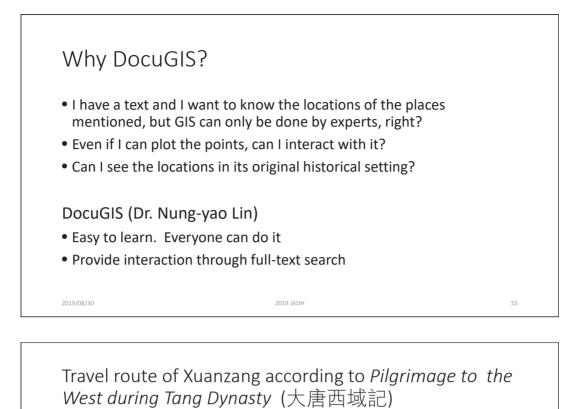






從數位典藏、數位人文到個人化平台的歷程

56

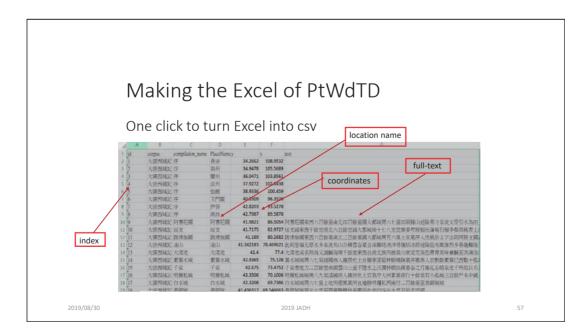


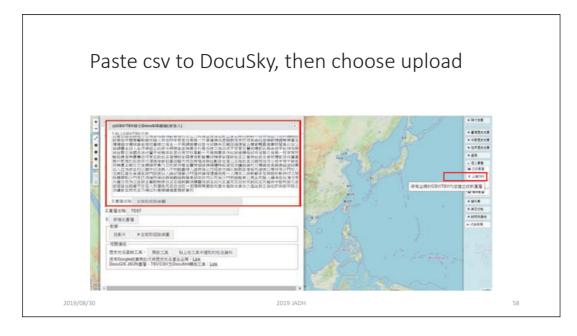
- Full-text of PtWdTD
 - Download from ctext using DocuSky API
 - (PtWdTD describes the same travel as the fiction The Journal to the West 西遊記, but was written by the monk Xuanzang himself)
- Geographic locations
 - Dharma Dram location authority (<u>http://authority.dila.edu.tw/place/</u>)

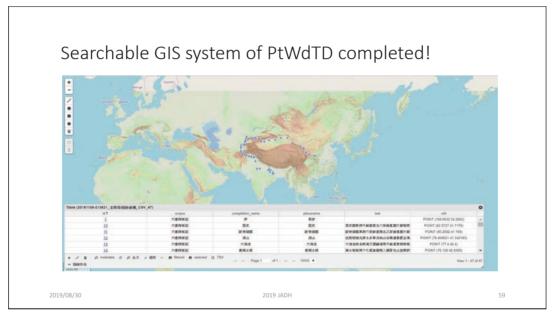
2019 JADH

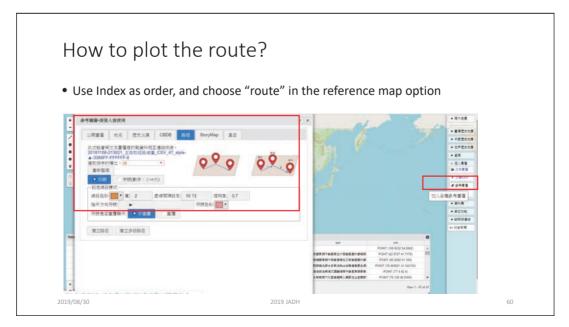
- Build Excel file of names and locations of places and the full-text
- Obtain csv from the Excel file
- Upload to DocuGIS and done!

2019/08/30



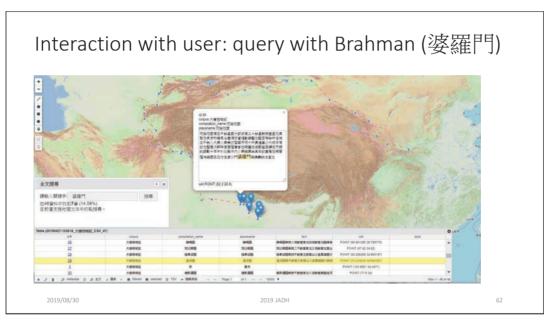




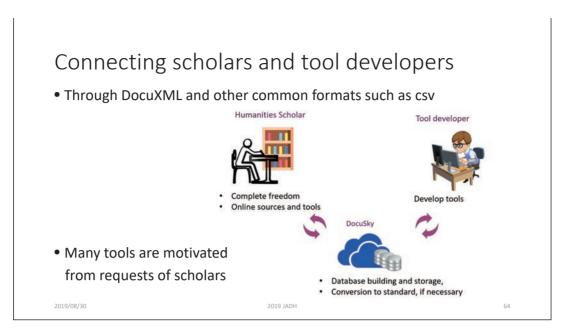


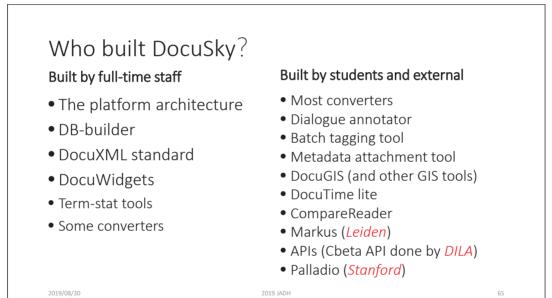
從數位典藏、數位人文到個人化平台的歷程

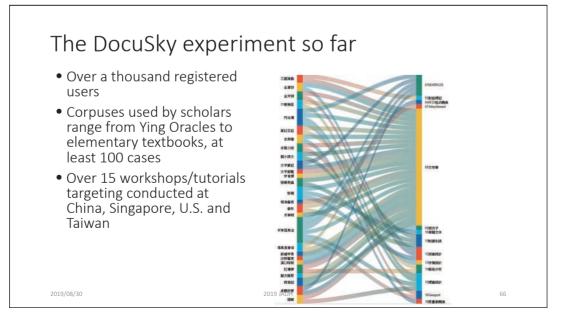






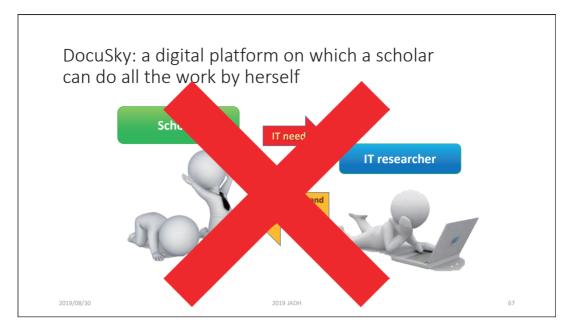






從數位典藏、數位人文到個人化平台的歷程

68



Concluding remarks

- We advocate the division of work between a scholar and digital tools, and that effective utilization of tools is the only way to harness the complexity raised by digital contents
- To do so we propose a new concept of platform with which new contents and new tools can be incorporated easily
- DocuSky aims at becoming such a platform where a scholar can work with digital contents and tools without loosing independence
- DocuXML provides a way for tools and contents (and their developers) to participate and interact

2019 JADH

• A lot more need to be done...

2019/08/30



西雅圖觀點:華盛頓大學的數位人文

Brian Reed

Divisional Dean of the Humanities & Professor of English and Comparative Literature University of Washington, USA

美國華盛頓大學人文學院院長暨英文及比較文學系教授

Abstract 摘要

In the United States, although the Digital Humanities has become a recognized and credentialed area of academic expertise, there remains little consistency across institutions of higher education regarding how and where such scholarship takes place and how it is supported.

In this talk, I will be presenting, from the viewpoint of a dean of humanities at a large public university, an overview of the problems and opportunities that have arisen as my institution has grappled with the emergence of the Digital Humanities as an interdisciplinary field and as a domain for inventive research, teaching, and outreach.

After an overview of how DH has been practiced and institutionalized at a number of peer universities, I will highlight distinctive DH projects at UW—such as Women Who Rock— and I will recount the pivotal role played by the UW Libraries, the Simpson Center for the Humanities, the DX Arts Department, and the I-School in maintaining a dispersed, dynamically changing DH ecology. I will also explain where my own research on "creative practice" in digital poetry has informed my thinking as an administrator.

UW has been successful as an incubator for DH innovation, I will argue, partly because it has not formalized a DH curriculum, department, or center; we have freed faculty and staff to imagine and pursue projects unconstrained by any one definition of DH or by any one disciplinary formation and associated set of practices.

在美國,儘管數位人文已經成為公認的學術專業領域,各高等教育機構對 於數位人文之學術研究應該如何與於何處進行,以及如何獲得相關支持等方 面,仍未有一致性。

在本次演講中,我將從一所大型公立大學的人文學院院長的角度介紹美國 華盛頓大學,在數字人文這樣一個跨學科領域其於發展研究、教學及其相關延 伸而出現問題之際,如何面對並掌握契機。

在概述了數位人文如何在一些相同規模的大學實踐和制度化之後,我將重 點介紹美國華盛頓大學其發展的獨特數位人文計畫 - 例如" Women Who Rock ",並介紹本校的辛普森中心、數位藝術與實驗媒體中心,以及 I-School 等單 位,其在維持分散且持續變動的數位人文生態,所扮演的關鍵角色。此外,也 將以管理者的角度,說明自身目前所進行的"創造性實踐"於數位詩詞的研究,

我認為,華盛頓大學已成功地成為數位人文之創新的孵化器,其中部分的 原因是本校尚未有制式僵化的數位人文課程,部門或中心;相對的,我們給予 教職員工最大的想像和創造空間,不受任何數位人文的定義或任何一個學科領 域範圍和相關執行面的各種限制。

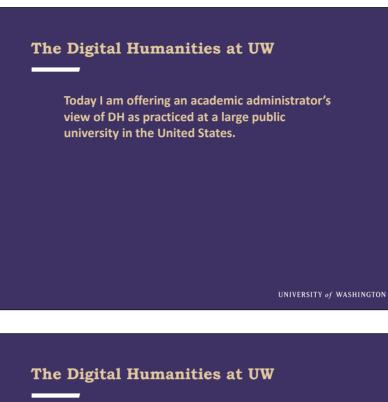
西雅圖觀點:華盛頓大學的數位人文





International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用



Today I am offering an academic administrator's view of DH as practiced at a large public university in the United States.

I hope to give you an example of *systems thinking* about DH within an institution that conceives of itself as a *learning organization*.

UNIVERSITY of WASHINGTON

Peter Senge, The Fifth Discipline (1990)

Systems thinking focuses not on individuals within an organization but instead their interactions within the organization and in between organizations as a whole.

西雅圖觀點:華盛頓大學的數位人文

Peter Senge, The Fifth Discipline (1990)

In *learning organizations* people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning to see the whole together.

UNIVERSITY of WASHINGTON

The Lessons of This Talk

The value of institutional self-assessment in promoting DH in higher education.

UNIVERSITY of WASHINGTON

The Lessons of This Talk

The value of institutional self-assessment in promoting DH in higher education.

The importance, in a resource-limited environment, of cultivating, supporting, and building on what you already have.



The Digital Humanities: What Is It?

Definitional uncertainty has plagued the field from its origins in the mid-1990s.

Equally uncertain has been ascertaining *who* counts as a digital humanist and *where* institutionally DH is being practiced.

西雅圖觀點:華盛頓大學的數位人文

Kim Knight, "A Media Ecological Approach to Digital Humanities" (2011)

The field of digital humanities is made up of the localized practices of practitioners who engage in work that uses digital technologies to analyze the traditional objects of humanistic inquiry, work that approaches the digital technologies and their outputs as objects for analysis, and work that uses digital platforms in humanities pedagogy, as well as digital librarianship and archiving work.

UNIVERSITY of WASHINGTON

Kim Knight, "A Media Ecological Approach to Digital Humanities" (2011)

The field "exists" simultaneously at different scales and in various locations – it is variously treated as a tool set, a sub-set of more established disciplines, or as a discipline in its own right. If you were attempting to locate DH on a map, you might look toward English, Art, History, Communication, or your library, or you may find it as a discrete department, program, or research center. Even better, you would look in all of these places at once.

UNIVERSITY of WASHINGTON

The Digital Humanities: What Is It?

A dynamic *ecology* that involves an institutional and technological infrastructure, curricula, digital projects, and people with diverse skills, specialties, roles, and commitments.

The Digital Humanities: What Is It?

A dynamic *ecology* that involves an institutional and technological infrastructure, curricula, digital projects, and people with diverse skills, specialties, roles, and commitments.

This ecology varies greatly across institutions of higher education in the United States. It is also subject to change and evolution at particular institutions.

UNIVERSITY of WASHINGTO

The Digital Humanities: What Is It?

Understanding what DH is at a university such as UW requires *system thinking*.

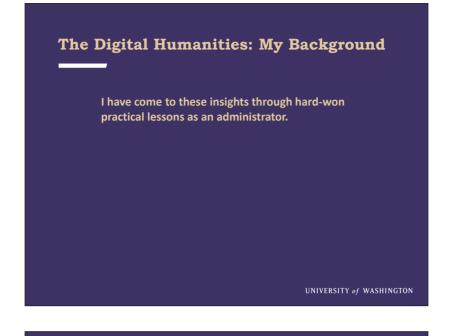
UNIVERSITY of WASHINGTON

The Digital Humanities: What Is It?

Understanding what DH is at a university such as UW requires *system thinking*.

Building DH requires understanding that your university is a *learning organization*, capable of institutional transformation through collaboration, self-assessment, and big-picture thinking.

西雅圖觀點:華盛頓大學的數位人文



The Digital Humanities: My Background

I have come to these insights through hard-won practical lessons as an administrator.

Systems thinking is difficult when you are caught up into the day-to-day work of your institution.

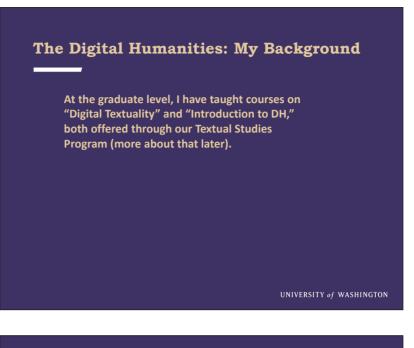
UNIVERSITY of WASHINGTON

The Digital Humanities: My Background

While I do not consider myself a DH expert, I do study born-digital literature and art, including video games, and I have undertaken several digitally based experiments in literary criticism.

International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班: 數位人文的發展與應用



The Digital Humanities: My Background

Prior to becoming Dean of the Humanities, I had administrative involvement in DH at my university. I served on:

- UW Digital Humanities Taskforce (2007-08)
- UW Teaching and Learning Technology Oversight Committee (2011-28)
- Selection Board, DH Summer Fellowships, Simpson Center for the Humanities, 2013-2015 and 2017.

UNIVERSITY of WASHINGTON

The Digital Humanities: My Background

I have served as a peer reviewer for *Digital Humanities Quarterly* and as a member of selection committees for the National Endowment of the Humanities and the Mellon Foundation in the areas of digital preservation and digital scholarship.

西雅圖觀點:華盛頓大學的數位人文

The Digital Humanities: My Background

I have twice participated in the two-week workshop "DH for Chairs and Deans" at the Digital Humanities Summer Institute at the University of Victoria, BC, 2014 and 2016.

UNIVERSITY of WASHINGTON

The Digital Humanities: My Background

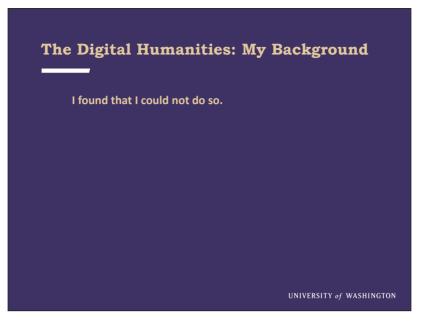
I have twice participated in the two-week workshop "DH for Chairs and Deans" at the Digital Humanities Summer Institute at the University of Victoria, BC, 2014 and 2016.

I gave a keynote lecture at a conference on "World Literature and DH" at Dongguk University, South Korea, in 2017 and participated in a roundtable on "The Present State of the Practical Digital Humanities."

UNIVERSITY of WASHINGTON

The Digital Humanities: My Background

Given these many ways in which my research, teaching, and service have put me into direct contact with DH—locally, regionally, and nationally, and internationally—I should have been prepared, as first Chair of the Department of English and then as Dean of Humanities, to speak authoritatively about "DH at UW."



The Digital Humanities: My Background

I found that I could not do so.

I didn't know enough about the DH infrastructure, programming, curricula, projects, and stakeholders to offer a coherent overview, let alone engage in strategic planning.

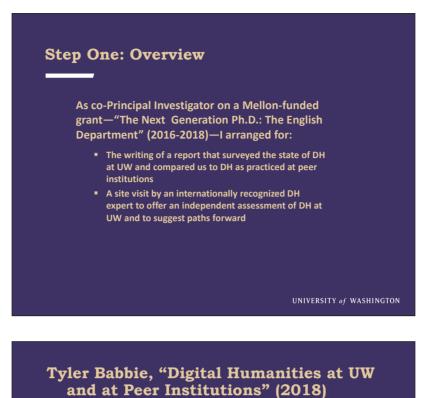
UNIVERSITY of WASHINGTON

Step One: Overview

As co-Principal Investigator on a Mellon-funded grant—"The Next Generation Ph.D.: The English Department" (2016-2018)—I arranged for:

 The writing of a report that surveyed the state of DH at UW and compared us to DH as practiced at peer institutions

西雅圖觀點:華盛頓大學的數位人文



Strengths in DH at UW:

 UW Libraries as a provider of DH-related resources and skill acquisition opportunities

UNIVERSITY of WASHINGTON

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

Strengths in DH at UW:

- UW Libraries as a provider of DH-related resources and skill acquisition opportunities
- The Simpson Center for the Humanities as a funder and backer of DH-related initiatives

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

Strengths in DH at UW:

- UW Libraries as a provider of DH-related resources and skill acquisition opportunities
- The Simpson Center for the Humanities as a funder and backer of DH-related initiatives
- Two graduate certificate programs—in Textual Studies and in Public Scholarship—that have strong DH-related curricula

UNIVERSITY of WASHINGTO

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

Strengths in DH at UW:

- UW Libraries as a provider of DH-related resources and skill acquisition opportunities
- The Simpson Center for the Humanities as a funder and backer of DH-related initiatives
- Two graduate certificate programs—in Textual Studies and in Public Scholarship—that have strong DH-related curricula
- Academic units such as the Linguistics, Human Centered Design, Geography, DXArts, and the Information School are already exceling at DH-related work, such as topic and sentiment analysis, though generally not under the label DH

UNIVERSITY of WASHINGTON

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

Strengths in DH at UW:

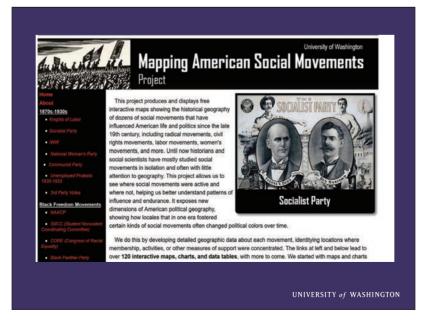
- UW Libraries as a provider of DH-related resources and skill acquisition opportunities
- The Simpson Center for the Humanities as a funder and backer of DH-related initiatives
- Two graduate certificate programs—in Textual Studies and in Public Scholarship—that have strong DH-related curricula
- Academic units such as the Linguistics, Human Centered Design, Geography, and the Information School are already exceling at DH-related work, such as topic and sentiment analysis, though generally not under the label DH
- DH projects pursued by individual faculty

西雅圖觀點:華盛頓大學的數位人文

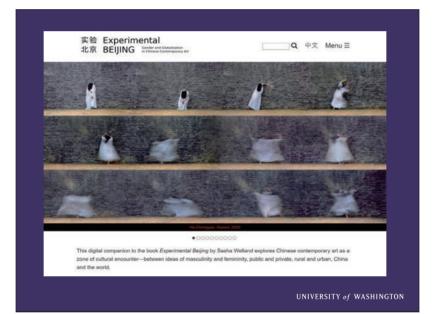
Individual and Collaborative DH Projects

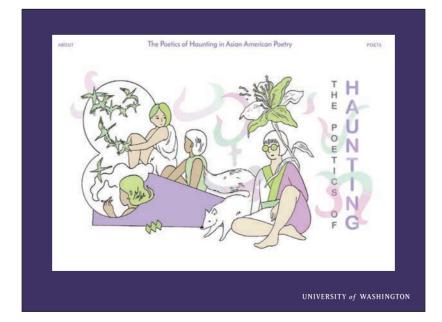
Numerous humanities-based faculty at UW have pursued digitally enabled, innovative projects, singly or in collaboration, sometimes under the heading of DH, other times under the labels *digital history*, *digital archiving*, *textual encoding*, *e-book authoring*, and *digital storytelling*.



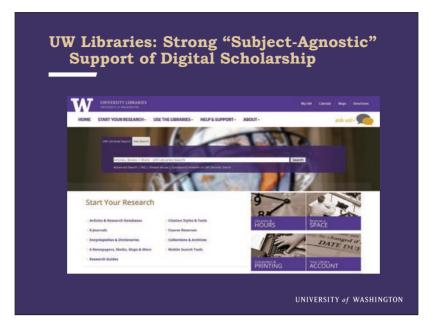


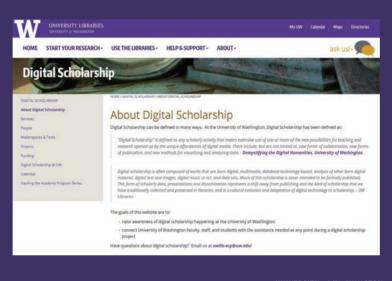






西雅圖觀點:華盛頓大學的數位人文





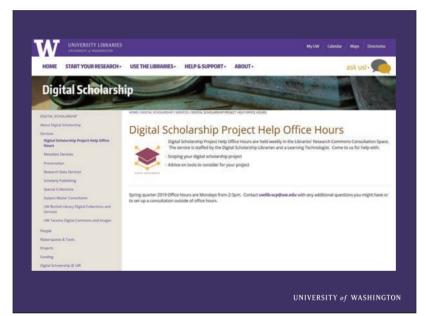
UNIVERSITY of WASHINGTON

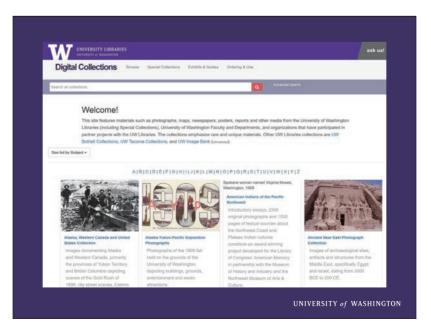
UW Libraries Digital Scholarship

Librarians available for consultation concerning:

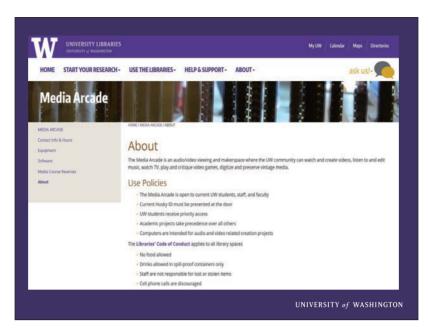
- Metadata development
- Data management, curation, and visualization
- Scholarly communications
- Project management
- Digital preservation
- Digitization
- Information architecture
- Interface design
- Usability
- Open access opportunities and policies

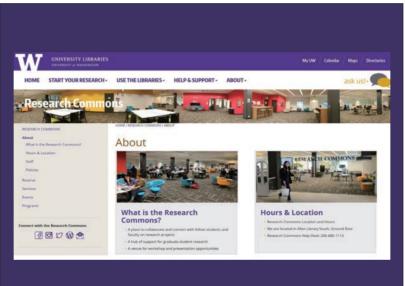


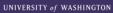




西雅圖觀點:華盛頓大學的數位人文

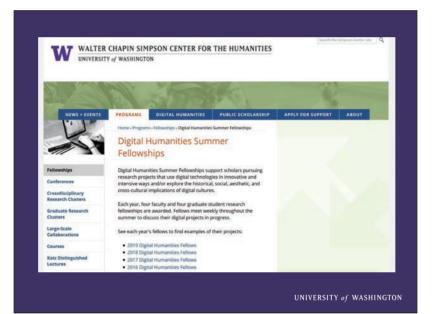


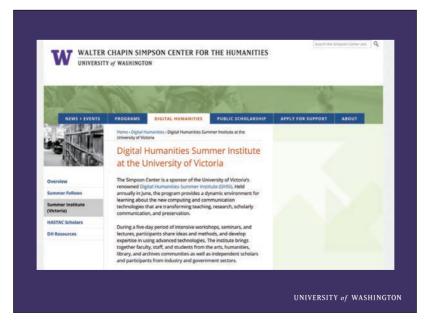












西雅圖觀點:華盛頓大學的數位人文









Certificate in Public Scholarship

The Certificate in Public Scholarship, established in 2010, is a partnership of the School of Interdisciplinary Arts and Sciences at UW Bothell, the UW Graduate School, and the Simpson Center.

Graduate students in any Ph.D. program can earn this certificate.

UNIVERSITY of WASHINGTON

Certificate in Public Scholarship

The Certificate in Public Scholarship, established in 2010, is a partnership of the School of Interdisciplinary Arts and Sciences at UW Bothell, the UW Graduate School, and the Simpson Center.

Graduate students in any Ph.D. program can earn this certificate.

They take core courses, electives, and a practicum. Throughout, they develop a portfolio that documents their community engaged teaching, research, and service.

The View from Seattle: The Digital Humanities at the University of Washington

西雅圖觀點:華盛頓大學的數位人文

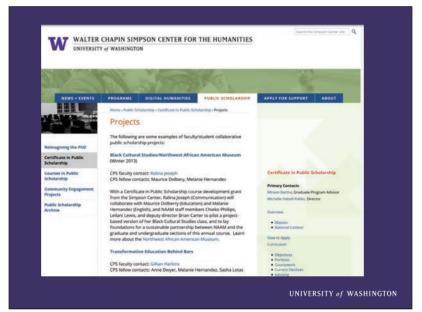
Certificate in Public Scholarship

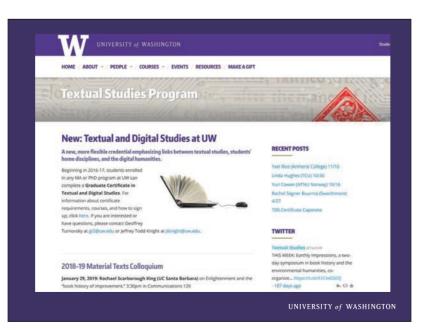
The Certificate in Public Scholarship, established in 2010, is a partnership of the School of Interdisciplinary Arts and Sciences at UW Bothell, the UW Graduate School, and the Simpson Center.

Graduate students in any Ph.D. program can earn this certificate.

They take core courses, electives, and a practicum. Throughout, they develop a portfolio that documents their community engaged teaching, research, and service.

While not explicitly conceived as a DH-related program, many participants have made use of digital tools, platforms, and archives in their public-facing work.







Textual Studies Program

Founded in 1995, it is the oldest graduate program in the US devoted to the "humanistic study of material texts."

It awards a Graduate Certificate in Textual and Digital Studies that can be combined with any doctoral program at UW.

UNIVERSITY of WASHINGTON

Textual Studies Program

Founded in 1995, it is the oldest graduate program in the US devoted to the "humanistic study of material texts."

It awards a Graduate Certificate in Textual and Digital Studies that can be combined with any doctoral program at UW.

It textures a well developed curriculum with seminars and a capstone project, and it attracts students from across and beyond the humanities, including the Information School and the Department of Human-Centered Design in the Engineering School.

西雅圖觀點:華盛頓大學的數位人文

Textual Studies Program

Founded in 1995, it is the oldest graduate program in the US devoted to the "humanistic study of material texts."

It awards a Graduate Certificate in Textual and Digital Studies that can be combined with any doctoral program at UW.

It textures a well developed curriculum with seminars and a capstone project, and it attracts students from across and beyond the humanities, including the Information School and the Department of Human-Centered Design in the Engineering School.

It runs speaker series, hosts conferences, and offers some funding for travel and research.

UNIVERSITY of WASHINGTON

Textual Studies Program

It promotes the study of *all* varieties of text, from epigraphy and holograph manuscripts to multimodal writing on digital platforms.

Volce Text Hypertext, ertextWolce Emerging Practices in Textual Studies EtVoice Text HypertextWolce Text H Edited by XtVolce Text Hyper Contributors Raimoda Modiano & Vinano M. Callewaer Leroy F. Searle xxtHyperte Roger Charle Peter Shillingsburg Hyper Phylis Cultarn ext HypertextWolce Text H Pad Egger Volce Text Hypertext Volce Concr Fah pertextVolce Text Hy John Miles Folg Text Hypertext Volce David Greetharr xtVolce Text Hypertext Susan Hockoy VpertextVolce Text Hype H Jackson e Text HypertextVolce Matin S. Jaffee ext HypertextVolce Text Jacome McGanr co Text HypertextVolce Matin McLooct HypertextVolce Text Jacome McGanr co Text HypertextVolce Rode Rochoect HypertextVolce Text H Soott B. Nosge ce Text HypertextVolce Rode Rochoect HypertextVolce Text H Soott B. Nosge ce Text HypertextVolce Rode Rochoect HypertextVolce Text H Soott B. Nosge ce Text HypertextVolce Rode Rochoect HypertextVolce Text H Soott B. Nosge co Text HypertextVolce Rode Rochoect HypertextVolce Text H Soott B. Nosge co Text HypertextVolce Rode Rochoect HypertextVolce Text H Soott B. Nosge co Text HypertextVolce Text H Soott B. Nosge co Text HypertextVolce Text Hockoect HypertextVolce Text Hypertex

UNIVERSITY of WASHINGTON

Textual Studies Program

It promotes the study of *all* varieties of text, from epigraphy and holograph manuscripts to multimodal writing on digital platforms.

It has, however, from the beginning put front and center the theoretical and methodological problems that digital technologies present for the field.

Voice Text Hypertext Voice Text Hy Voice Text Hypertext pertext Voice Emerging Practices in Textual Studies twoice Text Hypertext Voice Text H Edited by: twoice Text Hyper Controluters Ramonda Modano @ Wrand M Calewaer Leroy F. Sarle wit Hyperte Royer Chaine Peter Shillingsburg Hyper Phylis Guham oxt Hypertext Voice Text H Paul Egger Voice Text Hypertext Voice Conor Fah, pertext Voice Text Hy John Mies Fole Text Hypertext Voice David Greetham xtt Voice Text Hypertext Susan Hocksy pertext Voice Text Hype H Jokks Fole Text Hypertext Voice Main S. Jaffee ext Hypertext Voice Text Jackson eText Hypertext Voice Main S. Jaffee wit Voice Text Hypert Main Mechowa Hypertext Voice Text Jackson e Text Hypertext Voice Main S. Jaffee ext Voice Text Hypertext Moice McGanr ce Text Hypertext Voice Main S. Voice Hypertext Voice Text Jackson co Text Hypertext Voice Ludo Roche text Voice Text Hyperte Lercy F. Sarle Hypertext Voice Text Hypertext Voice Co Text Hypertext Voice Main S. Voice Co Text Hypertext Voice Text Hypertext Voice Co Text Hy

Textual Studies Program

It promotes the study of *all* varieties of text, from epigraphy and holograph manuscripts to multimodal writing on digital platforms.

It has, however, from the beginning put front and center the theoretical and methodological problems that digital technologies present for the field.

It has run workshops on text encoding (TEI) and recently has begun offering undergraduate courses that provide students with an opportunity to learn TEI and work on producing digital critical editions. Voice Text Hypertext Voice Text Hy Voice Text Hypertext pertext Voice Emerging Practices in Textual Studies xtVoice Text HypertextVoice Text H Edited by xtVoice Text Hyper Coninclutors Ramonda Modiano is Winnen M. Calewaer Lercy F. Saarle xxt Hyperte Roper Charle Peter Shillingsburgi Hyper Phylis Cultarr ext HypertextVoice Text H Paul Egger Voice Text HypertextVoice Coror Fam pertextVoice Text Hy John Mes Folg Text HypertextVoice David Greetham xtVoice Text Hypertext Susan Hockey ypartextVoice Text Hypertext Voice Martin S. Jaffee ext HypertextVoice Text H Paul Achove HypertextVoice Text Hypertext Voice Martin S. Jaffee ext Voice Text Hypertext Voice Martin S. Jaffee ext Voice Text Hypertext Voice Randal McLeoc re Text HypertextVoice Text H Souti B. Noege ce Text HypertextVoice Text House Lercy F. Saarle thypertextVoice Text HypertextVoice Text Hypertext

NIVERSITY of WASHINGTON

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

What UW lacks:

 Dedicated space and equipment for the use of scholars and students working on DH projects

UNIVERSITY of WASHINGTON

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

What UW lacks:

- Dedicated space and equipment for the use of
- scholars and students working on DH projects
- A central organizing or planning group to coordinate DH activities

The View from Seattle: The Digital Humanities at the University of Washington

西雅圖觀點:華盛頓大學的數位人文

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

What UW lacks:

- Dedicated space and equipment for the use of scholars and students working on DH projects
- A central organizing or planning group to coordinate DH activities
- Scholars hired outside our departmentalized structure specifically for their expertise in DH methods, topics, and pedagogy

UNIVERSITY of WASHINGTO

Tyler Babbie, "Digital Humanities at UW and at Peer Institutions" (2018)

What UW lacks:

- Dedicated space and equipment for the use of scholars and students working on DH projects
- A central organizing or planning group to coordinate
- DH activities
- Scholars hired outside our departmentalized structure specifically for their expertise in DH methods, topics, and pedagogy

In short: there has been a lack of a *systemic* approach to promoting DH.

UNIVERSITY of WASHINGTON

Jentery Sayers, "Locating Praxis in Digital Studies" (2018)

UW should not let the term DH and associated disciplinary conversations limit its future-thinking.



UW should not let the term DH and associated disciplinary conversations limit its future-thinking.

It should not seek in a top-down manner to import or emulate initiatives already underway elsewhere.

UNIVERSITY of WASHINGTO

Jentery Sayers, "Locating Praxis in Digital Studies" (2018)

UW should not let the term DH and associated disciplinary conversations limit its future-thinking.

It should not seek in a top-down manner to import or emulate initiatives already underway elsewhere.

It should cultivate and support locally originating, digitally based scholarship, teaching, and service that are "collaborative and transdisciplinary, involving a shared problem or concern approached through a combination of different disciplines and distinct forms of expertise."

UNIVERSITY of WASHINGTON

Jentery Sayers, "Locating Praxis in Digital Studies" (2018)

Conclusions:

 What is distinctive about UW are its strengths in Textual Studies and the Public Humanities.

The View from Seattle: The Digital Humanities at the University of Washington

西雅圖觀點:華盛頓大學的數位人文

Jentery Sayers, "Locating Praxis in Digital Studies" (2018)

Conclusions:

- What is distinctive about UW are its strengths in Textual Studies and the Public Humanities.
- UW should create infrastructure and curriculum—and engage in hiring—to build on these existing strengths.

UNIVERSITY of WASHINGTO

Jentery Sayers, "Locating Praxis in Digital Studies" (2018)

Conclusions:

- What is distinctive about UW are its strengths in Textual Studies and the Public Humanities.
- UW should create infrastructure and curriculum—and engage in hiring—to build on these existing strengths.
- In doing so, it must "commit to cultures and people" and give scholars and students the time and support to gain competencies and pursue long-term projects.

UNIVERSITY of WASHINGTON

DH at the University of Washington

What We Have:

UW Libraries

Simpson Center for the Humanities

Graduate Certificate Programs

Innovative Projects and Faculty

International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用



UNIVERSITY of WASHINGTO

DH at the University of Washington

Limitations and obstacles:

No available funding for major equipment and software purchases.

No space readily available for gathering and collaboration.

Few opportunities to hire faculty in the humanities in any given year.

UNIVERSITY of WASHINGTON

DH at UW: Next Steps 2019-202

Better identify stakeholders in DH at UW by setting aside preconceived notions of "what is and isn't DH" and "what belongs or doesn't belong under the label *humanities*" and concentrating instead on what research, teaching, and service faculty do—and how they do so (method, means of dissemination, etc.).

The View from Seattle: The Digital Humanities at the University of Washington

西雅圖觀點:華盛頓大學的數位人文

DH at UW: Next Steps 2019-202

Develop, with department chairs, a five-year Humanities Division hiring plan that includes a series of hires in DH with search committees drawn from across the division.

UNIVERSITY of WASHINGTO

DH at UW: Next Steps 2019-2020

Coordinate with department chairs incipient initiatives to develop undergraduate curricula in the areas of Textual Studies and Public Scholarship.

UNIVERSITY of WASHINGTON

DH at UW: Next Steps 2019-2020

Engage in five-year strategic planning with principal stakeholders that focuses on:

Creating a DH Coordination Committee



DH at UW: Next Steps 2019-2020

Engage in five-year strategic planning with principal stakeholders that focuses on:

Creating a DH Coordination Committee

Increasing connections, collaborations, and resourcesharing between individual faculties and projects

Establishing Textual Studies and Public Scholarship as the core and the "brand" of DH at UW

UNIVERSITY of WASHINGTON

Brian Reed

bmreed@uw.edu http://faculty.washington.edu/bmreed/



DocuSky: Your Powerful Personal Digital Humanities Platform DocuSky:專為您打造的數位人文研究平台

Chi-jui Hu 胡其瑞

Postdoctoral Researcher, Research Center for Digital Humanities, National Taiwan University 臺灣大學數位人文研究中心博士後研究人員

Abstract 摘要

There is nothing new today that humanities researchers using digital technology for their research in the digital age. However, it seems that the various digital tools always do not suitable for their individual needs. The DocuSky Collaboration Platform allows researchers to be freer to use their data and tools, and can easily perform many tasks that must be relied on by professionals in the past.

DocuSky is designed by the Research Center for Digital Humanities (RCDH) and CSIE Digital Humanities Lab of NTU in Taiwan, led by Professor Hsiang and Dr. Tu. It's a digital humanities research platform created specifically for humanities researchers of texts, to suit various needs and different stages of research. It provides tools to build up their own personal database of texts for text processing, tagging, text reorganizing and contextualizing, text-mining, GIS and visualization.

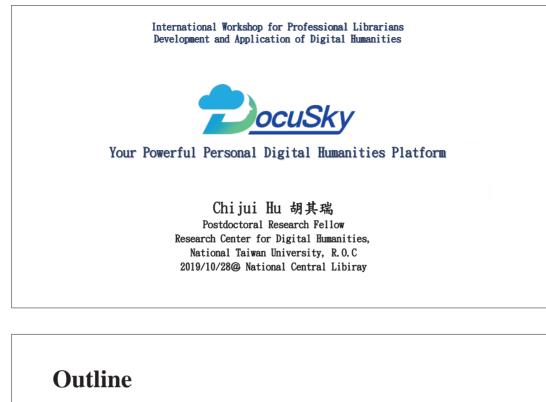
In this workshop, we will present the functions of the DocuSky, and show you how to use it easily and let it become every humanities researchers' powerful digital humanities platform.

Docusky 數位人文學術研究平台由臺灣大學數位人文研究中心項潔教授主 持,資訊工程學系數位典藏與自動推論實驗室杜協昌博士開發,是一個提供研 究者自行上傳個人研究文本,建立個人化資料庫,並且透過數位工具協助研究 者進行分析的數位人文研究平台。

有別於以往跟數位工具緊密結合且量身訂作的分析型數位資料庫,DocuSky 給 予使用者更大的自由度,使用者無論是透過平台取得文本資源,或是將個人所 蒐集的檔案、研究的資料,利用平台中的轉檔機制將文本轉製為符合平台需求 的檔案規格,就能自行建立個人化的資料庫,而不需透過資訊工程人員的協 助。輔以平台所連接的各式數位人文研究工具,對資料庫內的文本進行分析與 視覺化呈現,進而獲得研究發現,甚至發掘更多研究議題。

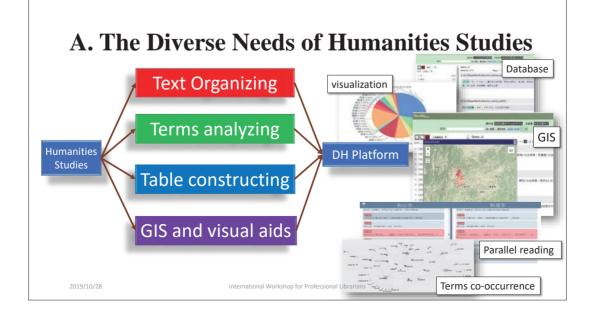
在本課程中,授課教師將引導學員進行數位文本處理與格式轉化、個人化 資料庫建置、數位人文工具分析,以及地理空間資訊呈現的練習與實作。希望 透過本課程的介紹,讓與會學員能夠更為熟悉 DocuSky 的操作,使 DocuSky 更為貼近個人的研究需求,成為個人研究工作中的神兵利器。

DocuSky:專為您打造的數位人文研究平台

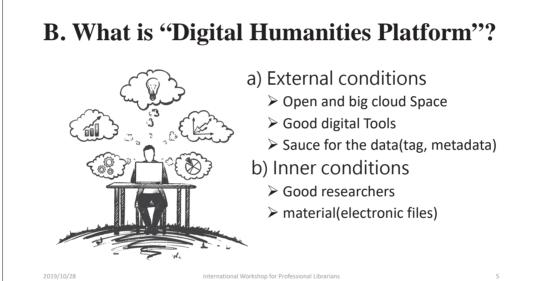


- A. The Diverse Needs of Humanities Studies
- B. What is "Digital Humanities Platform"?
- C. DocuSky and Humanities Studies
 - a) Download texts and build your personal database
 - b) Terms stats and analysis
 - c) Parallel reading and document comparison
 - d) Convert Table content and post-classification
 - e) GIS and fieldwork
- D. A Case Study
- Conclusion

2019/10/28

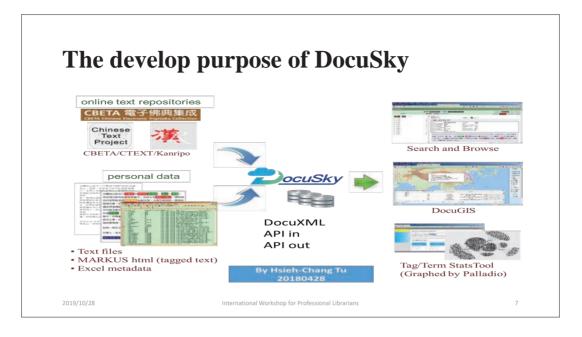


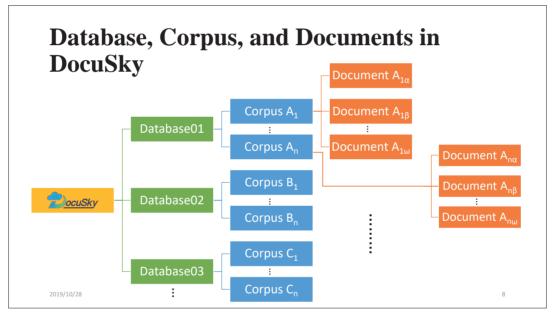


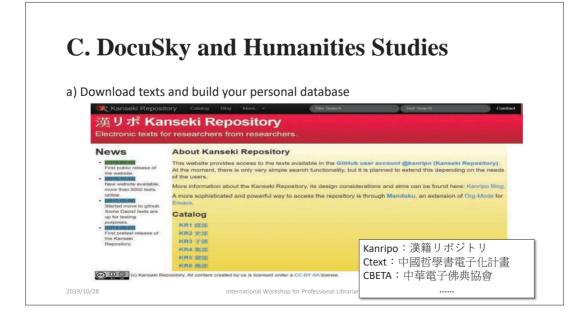


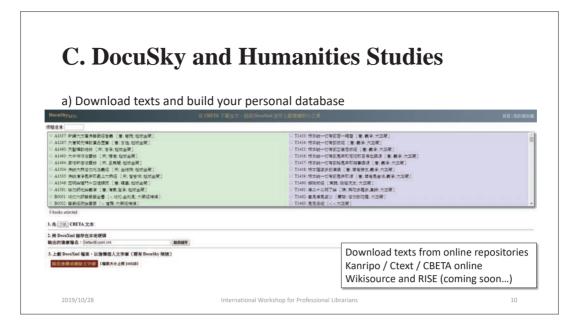


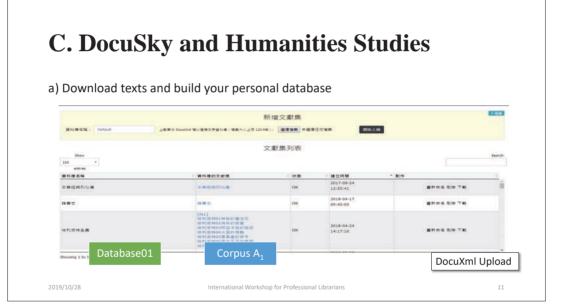
DocuSky:專為您打造的數位人文研究平台









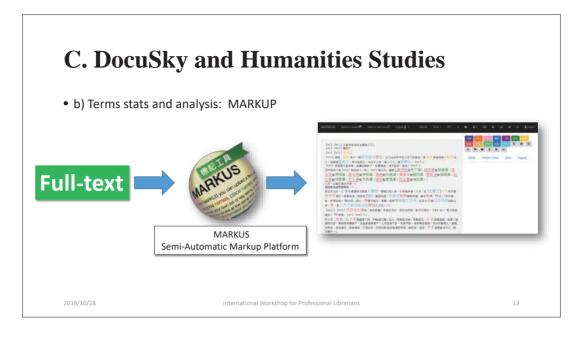


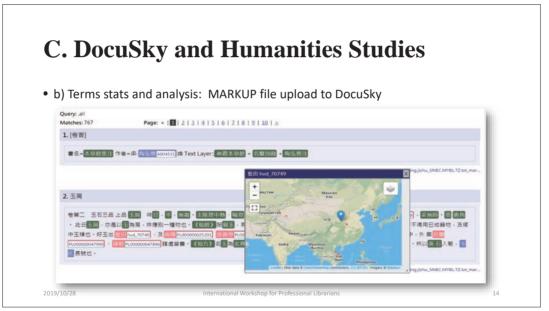


a) Download texts and build your personal database

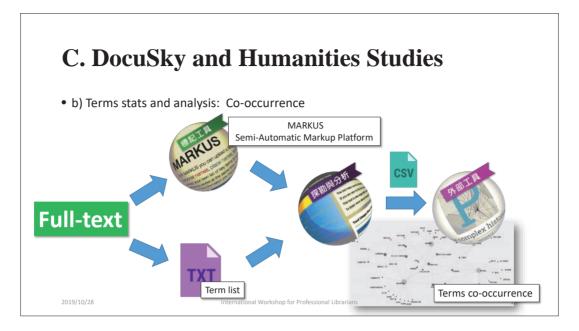
		資料庫: DaoBudMed6D 文獻集: [ALL] ·	Search Metadata Export		
查錄:		#1·納里 童年室時 reset undo 下一篇	DocMeta Snippets 4br/r pb		
		2. 靈寶無量度人上品妙經			
Sort: default * Buddhist Daoist Medical	2869 848 113	2013年1月2日、1月1日、1月1日、1月1日、1月1日、1月1日、1月1日、1月1日、			
		3. 靈寶無量度人上品妙經			
		結 無較衆天真大神上聖高尊妙行真人俱	上品妙經卷之二		
			「開催・三本本・1988・1ー1、大十一巻」用用ビネ文語・・・KR540001_002.txt ist, and Medical in 6 Dynasties in DocuS		
2019/10/28		Prof. Michael Stanley-Bake	, ,		

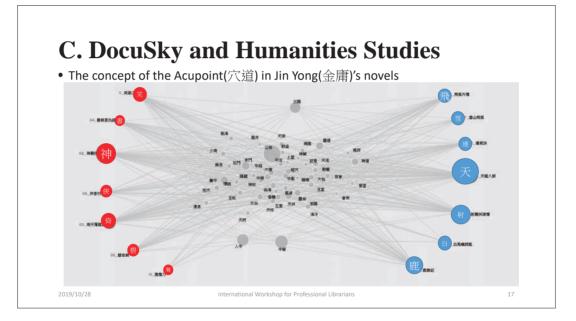
DocuSky:專為您打造的數位人文研究平台

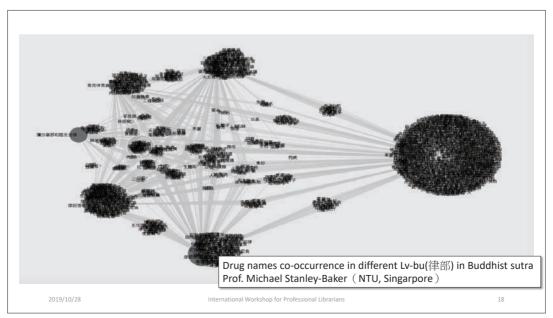






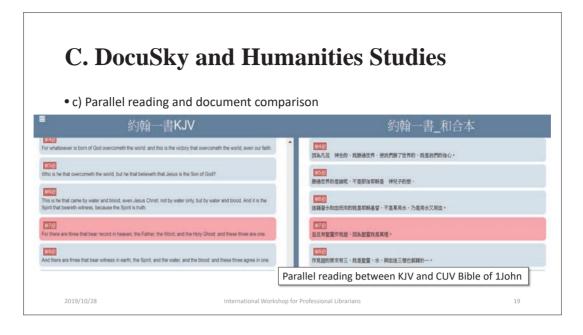


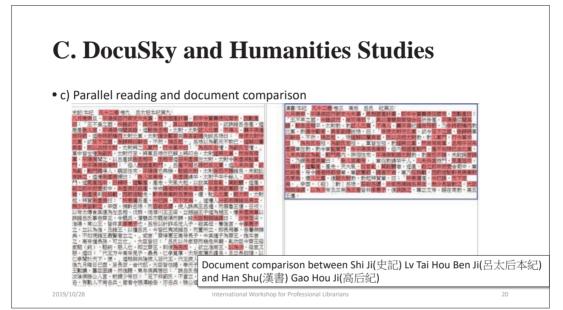


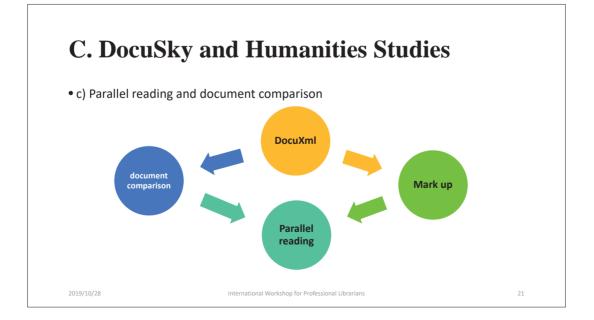


DocuSky: Your Powerful Personal Digital Humanities Platform

DocuSky:專為您打造的數位人文研究平台



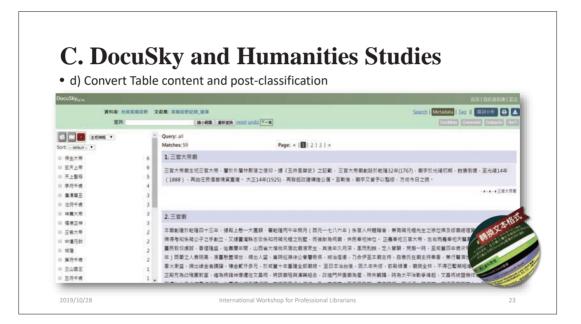


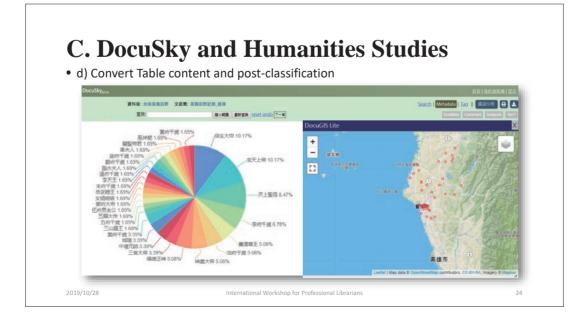


International Workshop for Professional Librarians : Development and Application of Digital Humanities

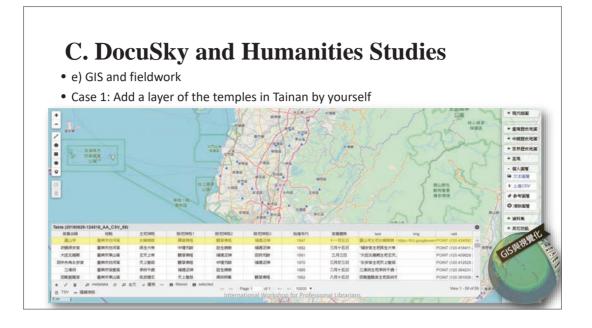
2019 國際圖書館專業館員研習班: 數位人文的發展與應用

-			~					~		
C .	Do	CU	Skv	' an	d H	uma	niti	es Stı	idies	
			•							
• d)	Conve	rt Tab	ole con	tent ar	nd post-	classific	ation			
• ()	Data fro	m Cult	ture Res	ource G	eographi	c Informa	tion Syst	em, Acaden	nia Sinica)	
filename	geo_latitude			doc_topic_11	doc_category_11	doc_category_12	doc_category_13		late_not_before time_norm_year	text
靈山寺	23.401014	120.434592	臺南市白河區	女媧娘娘	釋迦佛祖	觀音佛祖	福德正神	釋迦佛祖;觀音佛祖	1947 十一月五日	靈山寺主祀女媧娘娘,據儒
崁頭保安宮	23.329008	120.418411	臺南市白河區	保生大帝	中壇元帥	註生娘娘	福德正神	中壇元帥;註生娘娘	1882 <mark> 三月十五日</mark>	福安宮主祀保生大帝・據
大庄北極殿	23.310319	120.409628	臺南市東山區	玄天上帝	觀音佛祖	福德正神	田府元帥	觀音佛祖;福德正祥	1891 <mark>三月三日</mark>	大庄北極殿主祀玄天上帝
田中央角永安宮	23.348578	120.412528	臺南市白河區	天上聖母	觀音佛祖	中壇元帥	福德正神	觀音佛祖;中壇元飰	1970 三月廿三日	永安宮主祀天上聖母・據
三清府	23.366028	120.364231	臺南市後壁區	李府千歲	福德正神	註生娘娘		福德正神;註生娘娘	1995 二月十五日	三清府主祀李府千歳・傳
田尾聖龍宮	23.3054	120.361639	臺南市東山區	岳武穆王	天上聖母	周府將軍	觀音佛祖	天上聖母;周府將軍	1892 八月十五日	田尾聖龍宮主祀岳府元帥
天恩堂	23.307058	120.308256	臺南市新營區	謝府千歲	羅府千歲	中壇元帥	福德正神	羅府千歲;中壇元帥	1975 五月二日	天恩堂主祀謝府千歳・創
南鯤鯓代天府	23.285917	120.142731	臺南市北門區	五府千歲	觀音佛祖	地藏王菩薩	註生娘娘	觀音佛祖;地藏王督	1822 四月廿六日	話說明朝末年的一個晚上
惠濟宮	23.244808	120.157342	臺南市學甲區	保生大帝	中壇元帥	福德正神	註生娘娘	中壇元帥;福德正神	1866 三月十五日	中洲「惠濟宮」創建於清
廣安宮	23.220392	120.116772	臺南市將軍區	廣澤尊王	池府千歲	中壇元帥	福德正神	池府千歲:中壇元帥	1866 二月廿二日	郭聖王自道光王寅年,跨
朝天宮	23.190383	120.091186	臺南市將軍區	李府千歲	五府千歲	中壇元帥	中軍府	五府千歲:中壇元帥	1740	溯自前清乾隆五年・清鯤
黄氏崇榮堂	23.163344	120.176358	臺南市佳里區	神農大帝					1705 四月廿六日	本大宗祠创立社、公、192
後營普護宮	23.141292	120.21685	臺南市西港區	保生大帝	中壇元帥	註生娘娘	將軍爺	中壇元帥;註生娘娘	1834 三月十五日	本大宗祠创立的 居在 19
慈聖宮	23.186367	120.338911	臺南市官田區	神農大帝	福德正神	註生娘娘	中壇元帥	福德正神;註生娘娘	1796 二月九日	R SPECIAL
寧安宮	23.160622	120.177908	臺南市佳里區	溫府千歲	丁將軍	観音佛祖	萬善爺	丁將軍;觀音佛祖;會	1788 二月十五日	and a stand
牛庄元興堂	23.139328	120.316481	臺南市善化區	中壇元帥	劉府元帥	観音佛祖	註生娘娘	劉府元帥;觀音佛祖	1768 九月九日	18 ac
竹圍北極殿	23.138531		臺南市玉井區		神農大帝	中壇元帥	福德正神	神農大帝;中壇元帥	1717 三月三日	A CONTRACT CONT
大塩寮保安宮	23.104775		臺南市西港區		天上聖母	謝府千歳	中壇元帥	天上聖母:謝府千歲	1721 八月廿二日	保安
新港西福安宮	23.089281		臺南市新市區		International \	Netkeben for P		『福徳正神:註生娘娘	1755 三月廿三日	福安宮主在海市、被(新

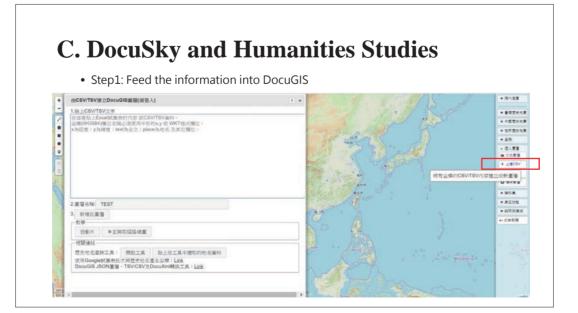




DocuSky:專為您打造的數位人文研究平台

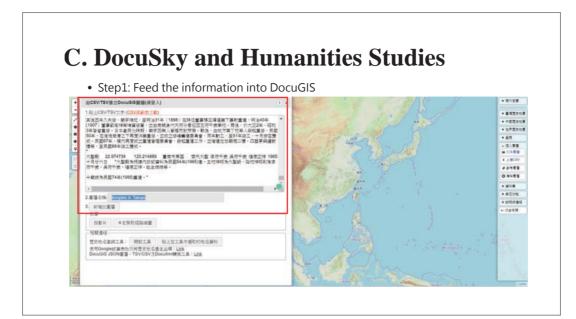


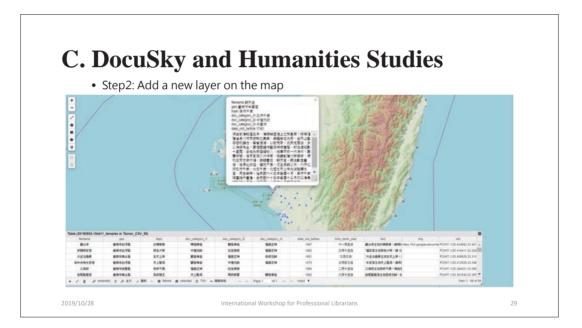
C. DocuSky and Humanities Studies placename text х у doc topic 11 doc category 1 doc category 12 doc category 1 臺南市白河區 女媧娘娘 臺南市白河區 保生大帝 靈山寺主祀女媧娘娘,據傳 福安宫主祀保生大帝,據《 靈山寺 釋迦佛祖 觀音佛祖 福德正神 釋迦佛祖;觀音佛祖 1947 40 崁頭保安宮 福德正神 月十五日 中壇元帥 註生娘娘 中壇元帥;註生娘娠 大庄北極殿 田中央角永安宮 臺南市東山區 玄大上帝 臺南市東山區 王上聖日 觀音佛祖 福德正神 田府元創 觀音佛祖:福德正神 91 三月三日 23.3103 23.348578 120.412 觀音佛祖 1970 三月廿三日 觀音佛祖;中壇元餅 120.364231 臺南市後壁區 字40-1 120.361639 臺南市東山區 岳武穆 正神;註生處處 聖母:厚府將軍 手歲:中壇元帥 1995 二月十五日 1892 八月十五日 三漬府 23,366028 三清府主祀李府千歳・傳説 福德正袖 Placename. 田尾聖龍宮 23.3054 120.36163 田尾聖龍宮主祀岳府元帥・ 1975 五月二日 1822 四月廿六日 天恩堂 23.307058 120.308256 臺南市新營區 謝府千歲 羅府 x (longitude), 天恩堂主祀謝府千歳・創立 南鯤鯓代天府 佛祖:地藏王著 23.285917 120.142731 臺南市北門區 五府千歲 觀音佛祖 話說明朝末年的一個晚上 y (latitude), 惠濟宮 23.244806 120.157342 臺南市學甲區 保生大帝 中壇元帥 1帥;福德正神 1866 三月十五日 中洲「惠濟宮」創建於清同 text 23.220392 120.116772 臺南市將軍區 廣澤尊王 1866 二月廿二日 廣安宮 池府千歳 歳:中壇元師 郭聖王自<u>道光</u>干寅年,跨海 朝天宮 23.190383 120.091186 臺南市將軍區 李府千歲 溯自前清乾隆五年・清餛鯓 千歲;中壇元帥 1740 五府千歲 1703 四月廿六日 **番氏崇榮堂** 23.163344 120.176358 臺南市佳里區 神農大帝 本大宗祠創立於清康熙42年 後營普護宮 23.141292 120.21685 臺南市西港區 保生大帝 1834 三月十五日 本廟所奉主神保生大帝 中壇元帥 註生娘姊 將軍爺 中壇元帥;註生娘娘 23.186367 120.338911 臺南市官田區 神農大帝 23.160622 120.177906 臺南市佳里區 溫府千歲 中壇元帥 萬善爺 1796 二月九日 1788 二月十五日 慈聖宮 福德正神 註生娘娘 福德正神;註生娘娘 慈聖宮主祀神農大帝,創立 丁將軍 觀音佛祖 丁將軍;觀音佛祖;喜 寧安宮為蕭壠開堡古廟・主 寧安宮 23.139328 120.316481 臺南市善仁區 中壇元帥 23.138531 120.479628 臺南市玉井區 玄天上帝 1768 九月九日 1717 三月三日 牛庄元與堂主祀九天中 壇元 竹圍北極殿又稱為鹿陶北極 牛庄元興堂 劉府元帥 觀音佛祖 註生娘娘 劉府元帥;親音佛祖 竹園北極殿 神農大帝 中壇元帥 福德正神 神農大帝:中壇元帥 大塭寮保安宮 23.104775 120.178447 臺南市西港區 廣澤尊王 天上聖母 謝府千崩 中壇元帥 天上聖母;謝府千歲 1721 八月廿二E 保安宮舊廟建於清康熙五十 福德正神 御祖。 120.256647 臺南市新市區 天上聖母 V許毕娘娘 新港西福安宮 3.089281 福德正神:註生娘娘 1755 三月廿三日 福安宮主祀媽祖、據《新市 2019/10/28



International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用



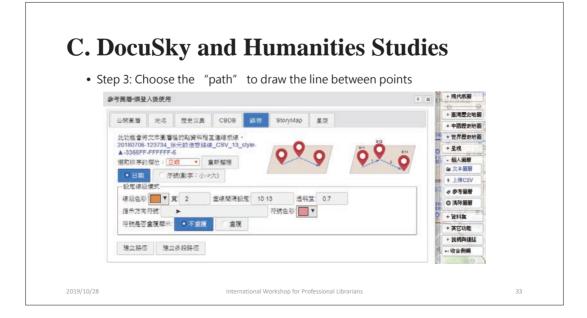


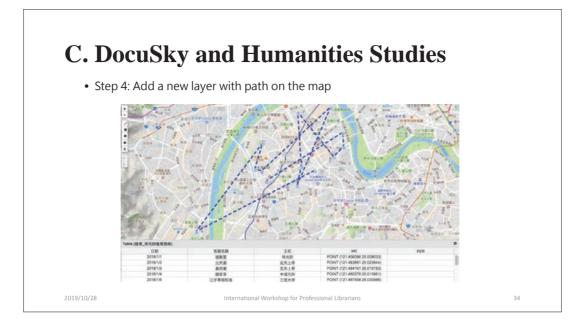
C. Docad Sup and Luna and Line of the religious practice in Taiwar: Gue Main (水) C. De se Sipritual power sharing in the religious practice in Taiwar: Gue Main (水) C. De meaning of Gua Hiunn in Sociology study C. De to use DH way to find something interesting:

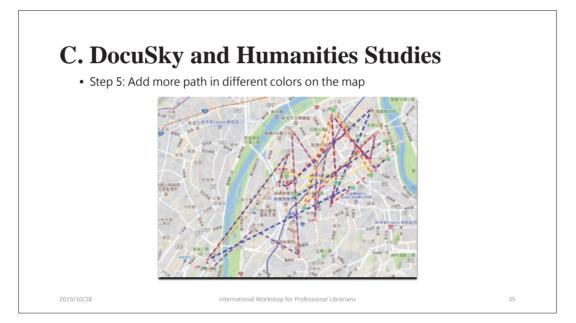
DocuSky:專為您打造的數位人文研究平台

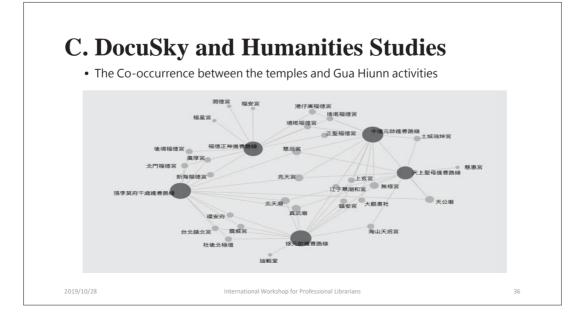
Ľ	. Doc	cuSky	and	Hur	naniti	es Studies
		·				ate column and the coordinate in
		5			included de	the column and the coordinate in
	x and y	columns)				
	A	В	С	D	E	
1	日期	宮廟名稱	7	x	主祀	
2	20180101	迪毅堂	25.008033	121.456086	徐元帥	
3	20180102	北天廟	25.023944	121.483881	玄天上帝	
4	20180103	真武廟	25.019183	121.484161	玄天上帝	
5	20180104	鎮安宮	25.015981	121.480578	中壇元帥	
6	20180105	江子翠潮和	25.030586	121.467558	三官大帝	The second state of the se
7	20180106	無極宮	25.012917	121.468908	西王金母	STATION AND AND AND AND AND AND AND AND AND AN
8	20180107	上玄宮	25.006539	121.467369	包聖青天	I and it that have any star
9	20180108	環安府	25.032281	121.479519	張李莫府千歲	the state of the s
10	20180109	大觀書社	25.011264	121.454456	文昌帝君	the fail of the second
11	20180110	震威宮	25.012025	121.452178	五雷元帥	a ser for the the second se
12	20180111	社後北極壇	25.020467	121.454911		the market of the second se
13	20180112	台北鎮北宮	24.991481	121.430847	五府千歲	The state of the s
14	20180113	海山天后宮	25.011483	121.473808	天上聖母	2000 2007 00 000 000 000 000 00000 000000 000000

		 Step) 2: Fee	ed the	information	into DocuGIS
4	A	В	С	D	Е	由CSV/TSV建立DocuGIS圖層(須登入)
		宮廟名稱	у	x	主祀	1.貼上CSV/TSV文字
2	2018/1/1			121.456086		日期 宮期 名稱 y x 主祀
1	2018/1/2 2018/1/3			121.483881		20180101 油設掌 25.008033 121.456086 徐元帥
	2018/1/3			121.484161		20180102 北天廟 25.023944 121.483881 玄天上帝
-		梨女 旨 江子翠潮和		121.480578		20180103 真武廟 25.019183 121.484161 玄天上帝
	2018/1/5			121.467338		20180104 績安宮 25.015981 121.480578 中壇元帥 20180105 江子翠潮和宮 25.030586 121.467558 三宮大帝
	2018/1/0			121.467369		-20180105 江于梁澍和宮 25.030366 121.467536 三官人帝 -20180106 無極宮 25.012917 121.468908 两王金母
2	2018/1/8				張李莫府千歲	20180107 上玄宮 25.006539 121.467369 包聖青天
0		大觀書社		121.454456		201801/08 現安府 25.032281 121.479519 張李莫府千歲
1	2018/1/10	震威宮	25.012025	121.452178	五雷元帥	20180109 大観書社 25.011264 121.454456 文昌帝君 20/80110 商廠寫 25.012025 121.452178 五面元前
2	2018/1/11	社後北極增	25.020467	121.454911	玄天上帝	20180110
3	2018/1/12	台北鎮北宮	24.991481	121.430847	五府千歲	20180112 台北鎮北宮 24.991481 121.430847 五府千歲
4	2018/1/13	海山天后宫	25.011483	121.473808	天上聖母	00400440 SELETER 05.044400 404.470000 TILES
5						 2.圖層名稱: 徐亓前進督路總









DocuSky:專為您打造的數位人文研究平台

D. A Case Study : Terms analysis about the "Rumors" in the Hmong "rebellion" in the 1950s

- Mass "rebellion" events in southwest China in the 1950s
- Main materials: reports of the Government
- The commonality of events: began with "rumors" and ended with military repression.
- Co-occurrence of the terms in the rumors
- Co-occurrence of the location on the map

2019/10/28

a) Organized data into a table and convert to DocuXml

&仙會事件 &仙會事件 &仙會事件 &仙會事件 &仙會事件 よ仙會事件	跳仙會事件 跳仙會事件 跳仙會事件 跳仙會事件	永綏衛城鄉鹽井寨 鳳凰 乾城 保靖 松桃	28.3724257 28.0188704 28.3042965 28.656493 28.1621519	109.37623 109.270406 109.502324 109.226906	1942 1942	湘西的跳 湘西的跳	(仙會事件) (仙會事件) (仙會事件)
&仙會事件 &仙會事件 &仙會事件 &仙會事件	跳仙會事件 跳仙會事件 跳仙會事件	乾城 保靖 松桃	28.3042965 28.656493	109.502324	1942	湘西的趴	仙會事件
&仙會事件 &仙會事件 &仙會事件	跳仙會事件 跳仙會事件	保靖松桃	28.656493				
14山會事件 日 14山會事件 日	跳仙會事件	松桃		109.226906	1942	湖田市的別	11 A
14個會事件		10.00	29 1621510		1212	10530300	:仙會事件
	跳仙會事件		28.1021319	108.43962	1942	湘西的跳	山會事件
k伯會事件		銅仁	28.1041047	107.488571	1942	湘西的跳	山會事件
	跳仙會事件	臘爾山	28.0814192	109.378777	1942	湘西的跳	山會事件
k仙會事件]	跳仙會事件	麻沖	28.1720642	109.558695	1942	湘西的跳	仙會事件
1 他會事件	跳仙會事件	舊司坪	28.1801981	109.646215	1942	湘西的跳	仙會事件
k仙會事件	跳仙會事件	竿子坪	28.3042965	109.502324	1942	湘西的跳	油會事件
k仙會事件]	跳仙會事件	白岩沖	27.9821402	110.272134	1942	湘西的跳	仙會事件
14個會事件	跳仙會事件	牛岩	28.0131962	109.550676	1942	湘西的跳	仙會事件
k仙會事件	跳仙會事件	亥沖	28.0003102	109.494729	1942	湘西的跳	仙會作件
5場事件 .	馬場事件	納雍縣馬場鄉	26.9079794	105.052417	1956	關於馬	太相正、同問
湯場事件	馬場事件	納雍	26.7981479	104.999005	1956	國方 经	<u>ب</u>
湯場事件	馬場事件	水城	26.4820719	104.351936	1956	關	and an
[場事件]	馬場事件	赫章	27.1170244	104.038612	1956	關	
場車件	E. 揾重件:	鼎 篩	27.065528	104 036943	1056	1251	6
<u>= 및</u> (+)						- Car	and the fi
	仙會事件 仙會事件 仙會事件 仙會事件 仙會事件 仙會事件 場事件 場事件 場事件 場事件 場事件	仙會事件 跳仙會事件 此仙會事件 跳仙會事件 此仙會事件 跳仙會事件 他會事件 跳仙會事件 仙會事件 跳仙會事件 場事件 馬場事件 易事件 馬場事件 易事件 馬場事件 易事件 馬場事件 易事件 馬場事件 局事件 馬場事件	仙會事件 跳仙會事件 職沖 仙會事件 跳仙會事件 舊司坪 仙會事件 跳仙會事件 笞子/平 仙會事件 跳仙會事件 白岩沖 仙會事件 跳仙會事件 白岩沖 仙會事件 跳仙會事件 七岩沖 山會事件 跳仙會事件 大岩 場事件 馬場事件 約雍 場事件 馬場事件 水域 場事件 馬場事件 素章 場事件 馬場事件 素章 場事件 馬場事件 基章	仙會事件 跳仙會事件 麻沖 28.1720642 仙會事件 跳仙會事件 舊司坪 28.1801981 仙會事件 跳仙會事件 笞子坪 28.3042965 仙會事件 跳仙會事件 台岩沖 27.9821402 仙會事件 跳仙會事件 台岩沖 27.9821402 仙會事件 跳仙會事件 キ岩 28.003102 場事件 馬場事件 約雍縣馬場鄉 26.9079794 場事件 馬場事件 約雍 26.7981479 場事件 馬場事件 水城 26.4820719 場事件 馬場事件 赤章 27.1170244 場事件 馬場事件 無章 27.05558	 仙會事件 跳仙會事件 麻沖 28.1720642 109.558695 仙會事件 跳仙會事件 萬司坪 28.1801981 109.646215 仙會事件 跳仙會事件 辛子坪 28.30295 109.50224 仙會事件 跳仙會事件 白岩沖 27.9821402 109.550676 仙會事件 跳仙會事件 大岩沖 28.003102 109.49205 場事件 馬場事件 素場事件 馬場事件 素場事件 馬場事件 本城 26.7981479 104.99005 場事件 馬場事件 素場事件 本城 26.4820719 104.351936 場事件 馬場事件 基章 27.1170244 104.036012 日本 日本 第6 	仙會事件 跳仙會事件 麻沖 28.1720642 109.558695 1942 仙會事件 跳仙會事件 営司坪 28.1801981 109.646215 1942 仙會事件 跳仙會事件 デ子坪 28.3042965 109.502324 1942 仙會事件 跳仙會事件 白岩沖 27.9821402 110.272134 1942 仙會事件 跳仙會事件 牛岩 28.0131962 109.550676 1942 仙會事件 跳仙會事件 支沖 28.0003102 109.494729 1942 場事件 馬場事件 約雍縣馬場鄉 26.0079794 105.052417 1956 場事件 馬場事件 約雍	仙會事件 跳仙會事件 願沖 28.1720642 109.558695 1942 湘西的跳 仙會事件 跳仙會事件 舊司坪 28.1801981 109.646215 1942 湘西的跳 仙會事件 跳仙會事件 管子坪 28.3042965 109.502324 1942 湘西的跳 仙會事件 跳仙會事件 白岩沖 27.9821402 110.272134 1942 湘西的跳 仙會事件 跳仙會事件 白岩沖 27.9821402 109.550676 1942 湘西的跳 仙會事件 跳仙會事件 大岩 28.0003102 109.494729 1942 湘西的跳 場事件 馬場事件 約雍縣馬場總 26.9079794 105.052417 1956 關声 場事件 馬場事件 約雍 26.47981479 104.999005 1956 圖声 場事件 馬場事件 水域 26.4820719 104.351936 1956 圖 場事件 馬場事件 素章 27.1170244 104.038612 1956 圖 場事件 馬島事件 蘇章 27.05528 104.035612 1956 圖

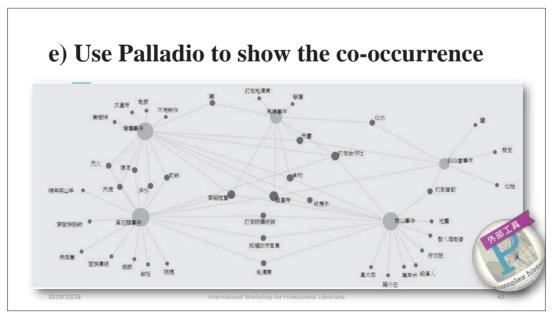
ional Workshop for Professional Librari

2019/10/28

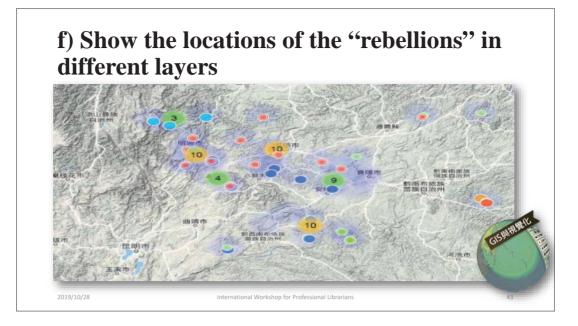
b) Upload the DocuXml to DocuSky 1 3.4 结小範圍 重新直跨 C 🛛 🗹 1. 塘壩事件 回爆對比,從思想上劃清了革命與反革命,民族與反革命宗教信仰與反革命等取我界限,認識到農業合 ○ · 苗族遭没有南市解 不僅漢族內部有反革命。? 民族內也有反革命份 Door 亥沖/-/-粉····
疗教人利用小數厚拢的宗教经信用想做: 第1,以原天的陈微时 保護/-/-論了鎮反政策及憲治反革命條例,對王富才等20 ◎ 礼報刑重大的份子 代表非常深切痛恨,提出:「放走一徑反革命5 專門作了鎮反政策的報告,使代表思想上劃滴了 姆,男女關係問題,苗族開設家要通行證,遷刻 日 王事/-/-· 預利對動門爭 · 會議 安康/-/安康的米賞/-/-拔民族幹部·苗族 E座設會上代表們的 日 昭通垣第/-/-日 昭通大覧/-/-分面家他關及其工作人員提出批評及政進意見。 該對部一種對時,依法處理,總的醫來,會講是 從另方面來看,一部份對民族與反革命認識上最 □ 喧嘩水業/-/-上個人士有些抵標,如來良王正與積極的功場 何。, 实良?縣長吳忠烈(上醫)是王富寸的 昭遺糵津/-/-B STR/-/-王富才 点,一般上層在整個會調中表現是不能構成。 后根/-/-共昭通地委结戦部19570129 貴州少數民族社會 目 水道/-/-

locuSky _{BEDA}			山田(北北市市市) 並出
度時 :		資料車: 新華高台語 文獻集: 新華国政治市局語 個小規模: 東新主時 (teset undo) 丁一里	Starch Metadata Lag & Export Dockless Comment Snippets dars
ti 📻 🔣 🚓 🕅		Query: all Matches: 34 Page: < 1 2	,
「「「「「」」の「「」」の「「」」の「「」」の「「」」の「」」の「「」」の「	34	1. 建语事件	
天田	28	回爆對比,從思想上劃清了革命與反革命,民族與反革命宗教信仰與反革命	8.等影我界观。認識到農業合作化給農民帶來的好處,絕大部分代表表明確了
1 但水	24	不僅須族內部有反革命,?民族內也有反革命份子,反映說:「漢族的惡事	展反革命被打垮了,但我們苗族的惡霸反革命沒有打倒,苗族還沒有做應解
李般柱義	21	放。」「敵人利用少數民族的宗教经信思想進行反革命活動。我們今後要屬	聲傳起來,要歸共產黨,毛主席的話,不要上反革命的篇,」以而天的時間討
打倒合作社	21	論了講反政策及思治反革命條例,對王富才等20名反革命份子分別提出了#	g理應見,要求政府從嚴懲處首惡份子和積極參加暴亂罪刑重大的份子,很多
1 黄钧	20		2的本。」並且在會讓討論中,提供了敵人活動的線索,有利對敵鬥爭,會議
打到業業	18		出處理意見時有了尺度。 P004 另方面,代表隊培養提拔民族幹部,苗族的婚
H	15		敗民族中培養農業水利技術人才的問題提示了建議・在座設會上代表們對一部
但姑	13		5萬,宣傳黨的政策,對少數民族幹部??婦女??近行為遭要遷就,應與漢
	13		大的成效,發揚了民主。同時說明了少數民族人民對爆和政府的續護和支持。
度王	13		民族遺繡認為苗族内出了反革命,給苗族男子娘,見不得人,但人恥笑。信別
打倒毛灌業	11	上屬人士有些結構、如四良王正與積極參加總苗族搬家專件。在討論原專員	451和资时就:"不曾時4351和资何对了,曾時4351和运可可能在 p005 高式 打招做暴亂,殺害我工作幹部仍有同情意見,認為:「王富才是好人不是反革





DocuSky:專為您打造的數位人文研究平台





Conclusion: Division of humanities studies works between Scholars and DocuSky

a. Initiate research problem	DocuSky		
b. Gather related material	a. Obtain additional material over the Web	Scholars	
	b. Tools for procession material	Interpret results	
	c. Create a searchable database of processed content		
	d. Tools for analysis		
	e. Tools for visualization	•By Prof. Jieh Hsiang	2019083



The Development of Information Platform for Humanities Computing: the Case of Center for Southeast Asian Studies (CSEAS) Kyoto University

數位人文資訊平台之建置:以日本京都大學東亞地域研究所為例

Shoichiro HARA 原正一郎

Professor, Center for Southeast Asian Studies (CSEAS), Kyoto University, Japan 日本京都大學東南亞地域中心研究所教授

Abstract 摘要

Since the 1950s' pioneering work initiated by Roberto Busa in a computergenerated concordance to Thomas Aquinas' writings known as the Index Thomisticus, Humanities Computing or Digital Humanities (hereafter HC/DH) has been constantly growing as a field of scholarly activity at the intersection of computational technologies and humanities disciplines. Early CH/DH simply focused on digitizing and analyzing texts in the humanities, but current CH/DH involves collaborative, trans-disciplinary, and computational researches, teaching, and publishing. This means an increasing necessity for the development of information platforms to preserve and provide digital data and application tools, which consequently imply that libraries should become one of the main contributors to the development of HC/DC. As a presentation in "International Workshop for Professional Librarians," I would like to introduce, at first, an overview of HC/DH activities and development in the Asia Pacific area, especially in Japan, followed by a comment about how libraries have contributed to this process.

Though in HC/DH academic conferences, there are many research presentations using attractive tools and methods such as visualization, 3D modeling, virtual reality, network analysis, GIS, topic modeling, machine learning, and high performance computing, it is important to mention on the other hand, that many humanities' researchers have difficulty in accessing the basic tools for constructing data (bases) and sharing data (bases). Thus, in the second part of this presentation, the focus will be on describing an information platform which is applicable to CH/DH. As an example, this presentation will introduce the information system for area studies developed by the Center for Southeast Asian Studies (CSEAS) Kyoto University.

I should remark that developing an information platform has some difficulties. First, high expertise is required to construct database systems, which is an obstacle to database construction by humanities' researchers. The second is linked to the diversity of resources, which enriches the humanities' researches but provides heterogeneous and nonstandard metadata. The third difficulty is that this heterogeneous metadata makes it difficult to share and use different data simultaneously.

The CSEAS information platform has been developed to solve these problems. It is a complex information system that includes as main components the MyDatabase (MyDB), Resource Sharing System (RSS) and Basic Information Service (BIS). MyDB is a database builder, allowing humanities researchers to construct databases without requiring former expertise. MyDB can accept any metadata vocabulary including nonstandard ones. This enables humanities' researchers to use their own metadata vocabulary according to their own purpose.

On the other hand, these heterogeneous metadata makes the data sharing process difficult. The RSS has been developed to share heterogeneous metadata on the Internet and to provide users with a uniform interface to retrieve databases seamlessly in one operation. Humanities' data which comprises small fragments of heterogeneous metadata have a possibility to become big data through associations among each other. This is a different feature from ordinary big data which have simple data structures and are collected from sensors. Place, time, people are expected to be essential knowledge that will allow for associating heterogeneous metadata, and the current BIS has been developed as the digital ontology dictionaries of Japanese historical place names and world chronological names, which will be used as tools to link different data sets. Thus, MyDB, RSS, and BIS will contribute to accelerate humanities' open data. Finally, I would like to discuss about the future of the CSEAS information platform and international collaborations on humanities' data sharing.

自 20 世紀 50 年代 Roberto Busa 針對 Thomas Aquinas 的著作,利用電腦自 動產出索引並完成了具先驅性的著作:《Index Thomisticus》開始,電腦運算與 人文研究領域始有了交集,人文運算或數位人文(以下簡稱 HC / DH)的學術領 域一不斷發展壯大。早期的 CH / DH 只關注資料數位化,或是分析人文科學中的 文本,但目前的 CH / DH 涉及的層面包含協作,跨學科,計算研究,教學和出版。 這意味著開發資訊平台以保護和提供數位化數據和應用工具的必要性日益增加, 也代表著圖書館應成為 HC / DC 發展的主要貢獻者之一。故本次演講,我首先要 介紹亞太地區 HC / DH 活動和發展的概況特別是在日本的發展現況,並說明圖書 館如何在這之中能有所貢獻。

雖然在 HC / DH 學術會議上,有許多研究報告使用了具吸引力的工具和方法,如視覺化工具,3D 建構,虛擬實境,網絡分析,GIS,主題建構,機器學習和高性能計算等。然而,其實許多人文科學研究人員難以獲得構建數據(基礎)和共享數據(基礎)的基本工具。因此,在本次演講的第二部分中,重點將是描述適用於 CH / DH 的資訊平台。本演講將介紹由東京亞洲研究中心 (CSEAS)京都大學開發的區域研究資訊系統為例。

開發資訊平台的過程中會遭遇一些困難。首先,構建資料庫系統需要很高 的專業知識,這是人文科學研究人員構建置資料庫常面臨障礙;再者,是資源 的多樣性,它豐富了人文科學的研究,但提供了異質性和非標準的詮釋資料; 最後,則是這種異質性數據使得研究人員難以同時共享和使用不同的數據資 料。

而CSEAS 資訊平台則可解決這些問題。它是一個複雜的資訊系統,包括 MyDatabase (MyDB),資源共享系統 (RSS)和基本信息服務 (BIS) 作為主要 組件。 MyDB 是一個資料庫建置器,允許人文科學研究人員無需相關的電腦資 訊的專業知識即可建置資料庫。而 MyDB 亦可以接受任何詮釋資料詞彙表,包括 非標準詞彙表。這使人文科學研究人員能夠根據自己的目的使用自己的詮釋資 料詞彙。

另一方面,這些異質性的詮釋資料使得資料共享過程變得困難。RSS的開發是為了在網路上共享異質性詮釋資料,並為使用者提供統一的介面,以便在每一次的操作中可完整的檢索資料庫。包含小片段的異質性詮釋資料的人文數據,有可能通過彼此之間的關聯而成為大數據。這是與普通的僅包含單純資料結構的大數據有明顯不。時間、地點以及人,應該作為允許異質性詮釋資料相互連結的基礎資訊,而目前BIS已經被做為為日本歷史地名和世界時間名稱的

數位本體詞典,它們將被用於連結不同資料集的工具。因此,MyDB,RSS和BIS 將有助於加速人文科學的開放數據。最後,我想討論一下 CSEAS 資訊平台的未 來以及人文數據共享的國際合作。 數位人文資訊平台之建置:以日本京都大學東亞地域研究所為例



Short Bibliography

Name Shoichiro HARA Affiliation: The Center for Southeast Asian Studies (CSEAS), Kyoto University Title: Professor Humanities Computing, Geographic Information Systems (GIS), Machine Learning, Bigdata, Database, Health **Research:** Informatics Work History: 1987: Graduated from Department of Medical Science (Doctor Course), University of Tokyo. Received Doctor's Degree in Medical Science (Ph.D.): Artificial intelligence, Medical expert systems, Health information systems 1989: Joined the National Center for Science Information Systems (NACSIS), and assigned to the Research and Development Department as a Research Associate: Full text databases, Health record system, and Humanities computing 1991: Moved to the National Institute of Japanese Literature (NIJL), and assigned to the Research Information Department as an Associate Professor: Humanities computing, Database integration, OCR, and GIS 2006: Moved to the Center for Integrated Area Studies (CIAS), Kyoto University, and assigned to the Division of Area Informatics as a Professor: Humanities computing, Spatiotemporal systems, and Academic data management 2014: Director, the Center for Integrated Area Studies (to 2016) (On January 1st, 2017, the Center for Southeast Asian Studies (CSEAS) started as a newly integrated institution,

(On January 1st, 2017, the Center for Southeast Asian Studies (CSEAS) started as a newly integrated institution, between the former Center for Southeast Asian Studies and the Center for Integrated Area Studies) 2017: Deputy Director, the Center for Southeast Asian Studies (CSEAS) (to 2018)

Structure of the Presentation

1. Digital Humanities or Humanities Computing

Short history

• Digital humanities or humanities computing in Japan

2. My Experiences in Humanities Computing

- Transcription of Japanese classical texts
- Transcription of historical earthquake documents

3. Humanities Computing and Area Studies

- Area studies and area informatics
- MyDatabase (a database builder)
- Resource sharing system (database integration system)

4. Latest Challenges

- Linked open database
- Spatiotemporal infrastructure
- Web big data

Digital Humanities or Humanities Computing



Digital Humanities or Humanities Computing - Definitions -

Since this field has been constantly growing and changing, specific definitions are difficult or may not be appropriate.

Susan Hockey: We are concerned with <u>the applications of computing to research and</u> <u>teaching within subjects that are loosely defined as "the humanities,"</u> or in British English "the arts." * *susan Hockey* was Professor of Library and Information Studies and Director of the School of Library, Archive, and Information Studies at University College London until July 31, 2004. She was Chair of the Association for Literary and Linguistic Computing (1984-1997) and a member, twice Chair, of the Steering Committee for the Text Encoding Initiative.

Wikipedia: Digital humanities (DH) is an area of scholarly activity at <u>the intersection of</u> <u>computing or digital technologies and the disciplines of the humanities</u>. It includes the systematic use of digital resources in the humanities, as well as the reflection on their application.

Historically, the digital humanities developed out of humanities computing and have become associated with other fields, such as humanistic computing, social computing, and media studies. In concrete terms, the digital humanities embrace a variety of topics, from curating online collections of primary sources to the data mining of large cultural data sets to topic modeling.

Digital Humanities or Humanities Computing - Short History: 1949 to early 1970s -

Father Roberto Busa: made an index verborum of all the words in the works of St. Thomas Aquinas and related authors, totaling some 11 million words of medieval Latin. He wanted to produce a "lemmatized" concordance where words are listed under their dictionary headings.
First volume was published in 1974



from Wikipedia

• A CD-ROM of the Aquinas material appeared in 1992







數位人文資訊平台之建置:以日本京都大學東亞地域研究所為例

FEDERALIST

from Wikipedia

Digital Humanities or Humanities Computing - Short History: 1949 to early 1970s -

By the 1960s: Many digital concordances

Use of Computers in a Disputed Authorship Study: quantitative approaches to style and authorship studies

• **Mosteller and Wallace**: an attempt to identify the authorship of the twelve disputed Federalist papers. * *The Federalist Papers is a collection of 85 articles and essays written by Alexander Hamilton, James Madison, and John Jay under the pseudonym "Publius" to promote the ratification of the United States Constitution.*

Limitations of the Technology

- Mainframe Computers: IBM System/360 (1964)
- Data input: by hand on punched cards or paper tape
- Character set: uppercase letters only, Japanese, Chinese?
- Data media: magnetic tape => serialization => use a kind of markup
- Presentation of document structure: Markup
 - ✓ Intrinsic markup system
 - COCOA (Count and Concordance Generation on Atlas): the first influential markup system used in OCP)

Digital Humanities or Humanities Computing - Short History: 1970 to mid 1980s -

Academic Societies and Humanities Computing Centers

- 1966: Computers and Humanities (Journal)
- 1973: Association for Literary and Linguistic Computing (ALLC: UK)
- 1976: Oxford Text Archive (OTA)
- 1978: Association for Computers and the Humanities (ACH: USA)
- 1988: IPSJ SIG Computers and the Humanities: じんもんこん
 - * IPSJ: Information Processing Society of Japan: 情報処理学会

Technical Development

- **Computer:** mainframe computers (Available to use by researchers but expensive to use and register was necessary to use computing facilities)
- KANJI (漢字) code: JIS X 0208 (JIS C 6226 or JIS 78) in 1978, GB 2312 in 1980, KS C 5619 in 1982, BIG 5 in
- 1984, CNS 11643 in 1986 \rightarrow ISO/IEC 10646 (or Unicode) in 1993
- Storage: tape to disk
- Database: relational database prevailed => good for metadata
- · Many papers based on vocabulary studies by concordance programs
 - ✓ however, there were little new methodologies ?

Digital Humanities or Humanities Computing - Short History: mid 1980s to Early 1990s -

Personal Computers

• IBM

- Apple Macintoshes (Graphic User Interface (GUI) and HyperCard)
- * HyperCard provided a model of file cards with ways of linking between them

PC Based Applications

- Word-Cruncher: A free eBook reader with research tools to help students and scholars' study important texts
- TACT: Text Analysis Computing Tools (TACT) are designed to do text-retrieval and analysis on literary
 works (to retrieve occurrences of a word, word pattern, or word combination)
- Micro-OCP: PC version of Oxford Concordance Program. Though it was possible to run Micro-OCP on PC, but it was still a batch program

HyperCard (Hyper Text)

• Beowulf Workstation (Patrick Conner): This presents a text to students with links to a modern English version and linguistic and contextual annotations of various kinds

E-mail

- ListServ: a software for electronic discussion lists
- Humanist (Willard McCarty): became central to the maintenance and development of a community and it has made a significant contribution to the definition of humanities computing



Digital Humanities or Humanities Computing - Short History: mid 1980s to Early 1990s (cont.) -

Creating a <u>Standard Encoding Scheme</u> for Humanities Electronic Texts

- Chaos?: There had been many different and conflicting encoding schemes
- Frustration?: Tired of wasting time reformatting texts to suit particular software with the inadequacies of existing schemes
- Nancy Ide (1987): Organized an invitational meeting at Vassar College, Poughkeepsie
 ✓ Poughkeepsie Principles
- SGML (Standard Generalized Markup Language): Published by ISO, offered a mechanism for defining a
 markup scheme for many different types of text

Text Encoding Initiative (TEI)

- Based on Poughkeepsie Principles *Poughkeepsie [pakipsi] in New York State, USA
- Steering committee: representatives from ACH, ALLC, and the Association for Computational Linguistics
- Much of the TEI work was done by e-mail using private and public discussion lists
- 1990: First draft version
- 1994: First full version of the TEI Guidelines
- Publication of the TEI Guidelines coincided with full-text digital library developments
 - ✓ use TEI rather than inventing a markup scheme from scratch

Digital Humanities or Humanities Computing - Short History: Early 1990s to Present -

The Internet and the World Wide Web

Mosaic (1993)

- Web: A tool as finding some kinds of information but not really as a serious tool for humanities research?
- Hyper Text Markup Language (HTML): A weak markup system that perpetuated all the problems with word processors and appearance-based markup?
- New paradigm: Web is a superb means of publication, not only for the results of scholarly work but also for promoting their activities among a larger community of users
- Anyone can be a publisher on the Web: delivery of scholarly material over the Internet (no limit on size, hyper links are a useful way of dealing with annotations, incremental publications, etc.)
- Collaborative projects using the Internet
- A large Number of Project

ISO/IEC 10646 (or Unicode) :1993

Digital Libraries

• TEI in digital libraries: raised issues about the philosophy of the TEI (TEI was designed flexible definition of tags, but libraries define tags as standards and followed closely)

Digital Humanities or Humanities Computing - Recent Topics -

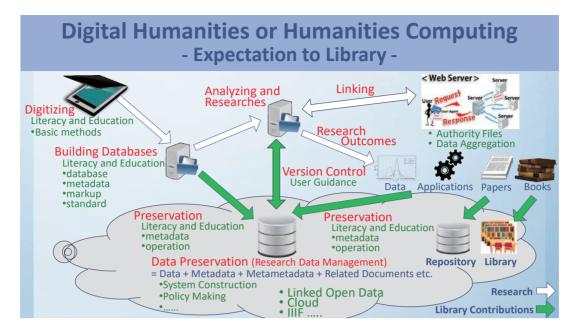
Technologies

- Big data
- Cloud computing
- Databases (relational databases, object-oriented databases, NoSQL, etc.)
- Statistics
- Geographic and chronological information systems (Spatiotemporal sciences and tools)
- Machine learning (Artificial intelligence, Deep learning)
- Visualization (Real time, 3D etc.)
- IIIF (International Image Interoperability Framework)
- 5G cellular network, and more

Digital Humanities: the applications of computing to research and teaching within subjects that are loosely defined as "the humanities," **then we should Focus on**

- Digitizing and building databases ?
- Authoring data ?
- Analyzing data and/or discovering new knowledge ?
- Visualizing research results ?
- Disseminating research results ?
- *

What Paradigm Shift did Digital Humanities or Humanities Computing give to Humanities?



Digital Humanities or Humanities Computing in Japan

Academic Societies

IPSJ SIG-Computers and the Humanities: じんもんこん * IPSJ: Information Processing Society of Japan: 情報処理学会 Established: 1988 URI: <u>http://www.jinmoncom.ip/</u> Members: about 240 Council for Databases in the Humanities, Japan: 人文系データベース協議会 Established: 1999 URI: <u>http://www.jinbun-db.com/</u> Members: about 40 Japanese Association for Digital Humanities Established: 2011 URI: <u>https://www.jadh.org/</u>







Digital Humanities or Humanities Computing in Japan

The Kyoto University Library Network

- URI: <u>https://www.kulib.kyoto-u.ac.jp/?lang=en</u>
- Founded: 1899
- Total number of books: 990,778 (Main Library) 7,093,273 (Total)
- Topic: On the Katsura Campus, five separate libraries of the Graduate School of Engineering have been operated since the opening of the campus in October 2003. The university has recently decided to consolidate these five libraries and build a new library that will serve as an Area-Focus Hub Library with a university-wide functionality. The building site is located in the B Cluster situated in the center of the campus, with a view of Kyoto's city center. Necessary preparations and arrangements are currently underway towards the opening of the library in the academic year 2019 (tentative).



Digital Humanities or Humanities Computing in Japan

Kyoto University Research Information Repository (KURENAI) contains peer-reviewed journal articles, dissertations, departmental bulletin papers and any kind of scholarly works of Kyoto University.

- URI: <u>https://repository.kulib.kyoto-u.ac.jp/dspace/?locale=en</u>
- Over 172,000 Articles (2018.3)
 - ✓ 240 Journals (2017)
 - ✓ over 5,000,000 fulltext downloads (2017)



• Kyoto University Open Access Policy: This policy mandates faculty members to publicize in principle their academic articles on the Internet by depositing them in "Kyoto University Research Information Repository KURENAI". Kyoto University Open Access Policy is to be implemented in academic year 2015. Faculty members will be notified about the detailed procedure in due course.

Digital Humanities or Humanities Computing in Japan

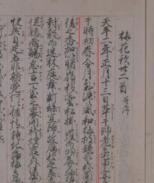
Kyoto University Rare Materials Digital Archive

- URI: https://rmda.kulib.kyoto-u.ac.jp/en
- In order to contribute to academic and cultural development, Kyoto University Library Network is engaged in digitizing rare materials our libraries hold and making them available online. We now have one of the largest digital collection in Japan.
- The system has viewers enabling the smooth display of high-definition images and is compliant with the International Image Interoperability Framework (IIIF). We will be continuously transferring the data from the previous system to this new system.
- 1,164,364 images of 13,575 titles are available in this digital archive as of July 30, 2019.



Digital Humanities or Humanities Computing in Japan

Japan's new era name Reiwa (令和) is derived from a poem in "Manyoshu (万葉集)", the oldest existing compilation of Japanese poetry.



[曼朱院本] 萬葉集 20巻 (manshuinbon man'youshuu):**京都大学附属図書館** Main Library, Kyoto University

Digital Humanities or Humanities Computing in Japan

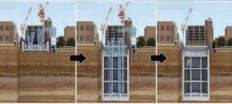
The University of Tokyo Library System

- URI: https://www.lib.u-tokyo.ac.jp/en
- Founded: 1877

• Total number of volumes: 1,285,461 (General Library) 673,355 (Komaba Library) 432,333 (Kashiwa Library) 9,745,659 (Total)

• Topic: The General library is now renewing (2020). An entirely new building has constructed (2017) beneath the area in front of the current library entrance, whilst a full-scale renovation of the main building's interior will preserve the original exterior of the historically significant structure. Compact shelves underground is at the depth of 46m, holding about 3 million books.





2019 國際圖書館專業館員研習班:數位人文的發展與應用

Digital Humanities or Humanities Computing in Japan

U-PARL: UEHIRO Project for the Asian Research Library in University of Tokyo General Library

• URI: <u>http://u-parl.lib.u-tokyo.ac.jp/en</u>

• U-PARL: Among the extensive collection of materials at the University of Tokyo, the extent and quality of Asia-related holdings are particularly noteworthy. Based on this collection, the university is planning to create world-class Asian research. For the conceptualization and realization of the "Asian Research Library," the new research project division "U-PARL was established in 2014 through an endowment by the Uehiro Foundation on Ethics and Education of Japan. U-PARL provides resources such as databases and online materials that will aid researchers, and develop collections that will be available at the Asian Research Library.



The am of UPR4, is to provide resources such as distances and usere networks that will assochers, and severage discloses that will be available at the Awain Reveals (Josoy Yee arcommty in the process of organizing, and expension by the strange of the available at the Awain Reveals (Josoy Yee arcommty in the process of organizing, and expension by the strange of the available at the Awain Reveals (Josoy Yee arcommty in the process of organizing, and expension by the strange of the Awain Reveals (Josoy Yee arcommty in the process of organizing, and the strange of the strange of the Awain Reveals (Josoy Yee arcommty in the process of organizing, and the strange of the strange of the Awain Reveals (Josoy Yee arcommty in the process of organizing, and the strange of the strange of the Awain Reveals (Josoy Yee arcommty in the process of the strange of the Awain Reveals of the Strange of the Awain Reveals are arrowed of the strange of the Awain Reveals (Josoy Yee arcommty in the Strange of the Awain Reveals are arrowed of the strange of the Awain Reveals (Josoy Yee arcommty in the Strange of the Awain Reveals are arrowed of the strange of the Awain Reveals (Josoy Yee arcommty in the Strange of the Awain Reveals (Josoy Yee arcommty in the Strange of the Awain Reveals (Josoy Yee arcommty in the Strange of the Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommty in the Strange of Awain Reveals (Josoy Yee arcommany Jee arcommany Jee arcommty in the Strange of Awain Reve

Digital Humanities or Humanities Computing in Japan

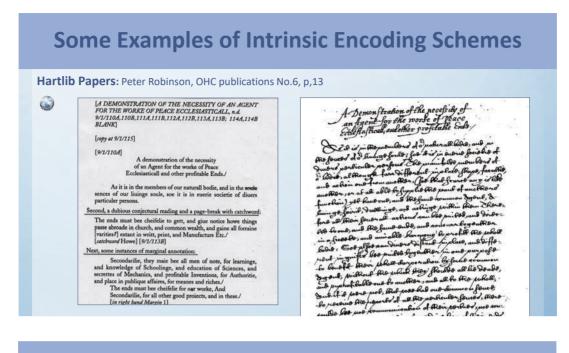
Some Collections of Chinese Books held by the University of Tokyo

- 皇明條法事類纂 Digital Archive: https://iiif.dl.itc.u-tokyo.ac.jp/repo/s/koumin/page/home
- 明 · 弘治十八年八月二十日勅命 Database: https://iiif.dl.itc.u-tokyo.ac.jp/repo/s/chokumei/page/home
- U-PARL漢籍・碑帖拓本資 (this collection includes 古今歴代十 八旗氏族通譜, etc.)

https://iiif.dl.itc.u-tokyo.ac.jp/repo/s/uparl/page/home

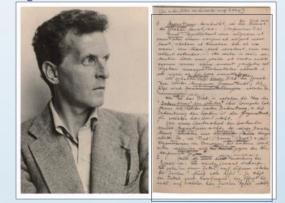
My Experiences in Humanities Computing - Transcription of Japanese Classical Texts -





Some Examples of Intrinsic Encoding Schemes (cont.)

Wittgenstein Archive



- </p

- from Wikipedia https://wittgenstein-initiative.com/category/articles/

<q/{usq/susq}heute{usq/eusq}/q>{p/es}/s> /sec <R/332/R><sec/<indl/7/indl>

My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts -

Project Purposes

- Design of data model and appropriate encoding schema for Japanese classical texts
- Reconstruction of original materials in digital form
- Some quantitative analysis (e.g., authorship)

Project Period: 1988 – 2004 (led by Hisashi YASUNAGA and Shoichiro HARA)

- 1988: Beginning of the project: Before SGML (1986) and TEI (1994) ✓ Grant in Aid by the Ministry of Education, Science and Culture
- 1993: Discontinuation of the project
- 1998: Restart of the project
 - ✓ proofreading and construction of the database system
- 1999: Experimental release
- 2002: Second version (XML)

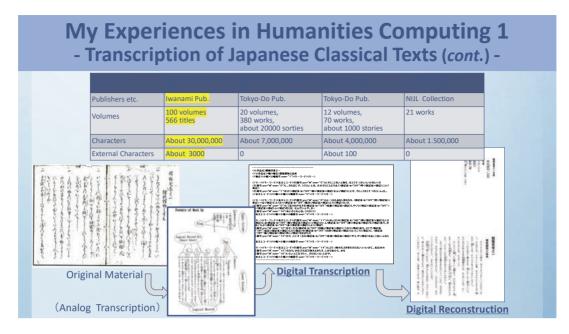
Current Database: Free for universities, academic institutes, libraries etc.

Copy Right: NIJL and publishing companies



へいや、275。 2260 L8/(廿三)一△清寧天皇△五年△元年/庚申(かの えぎる)△卅/九(一小)即位或卅七。





My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

Encoding Schema

KOKIN Rules: An Intrinsic Schema (before SGML and TEI)

- KOKubungaku (国文学) INformation Rules
- Markup rules for digital transcription of Japanese classical texts

Objectives

- KOKIN Rule 1: Logical structure markup rules
- KOKIN Rule 2: Serializing rules
- KOKIN Rule 3: Lexical annotation rules

Design

• Easy use by researchers

My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

KOKIN Rule 1: Structure Markup Rules

Texts contain various elements (e.g., titles, chapters, sections) that create hierarchical (or nested) structure

-¥Y ____ Ln or Mn

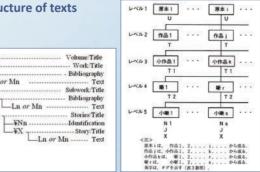
-¥1 -Ln or Mn

WNn

-¥T1 -

KOKIN Rule 1 is introduced to describe the structure of texts

- Record: string, line, title, etc.
- Record Set: chapter, work, etc. ¥U —
- File: Book etc.



My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (*cont.*) -

Schema of KOKIN Rule 1

<record> <tag begin=""></tag></record>	<pre>::= <tag begin=""><tag><tag end=""> <tag begi<br="">::= ' ¥ (Japanese-Yen-Mark)'</tag></tag></tag></tag></pre>	n> <rag></rag>	> <data><1</data>	ag End>
<tag end=""></tag>	::= '★(Star-Mark)'			
<tag></tag>	::= <tag symbol=""> <tag symbol=""><tag attril<="" th=""><th>oute></th><th></th><th></th></tag></tag></tag>	oute>		
<data></data>	::= <line> <original data=""></original></line>	Tag	Attribute (n)	Role
	<repeating symbol=""><original data=""></original></repeating>	UTn	None	Logical Recordset (Volume and its Title) Logical Recordset (Work and its Title)
<line></line>	::= <original data=""> </original>	100.00	1 2	Logical Recordset (Subwork and its Title) Logical Recordset (Group of Stories and its Title)
	<serial number=""><original data=""></original></serial>	Yn	None 1	Logical Recordset (Bibliography of Work) Logical Recordset (Bibliography of Subwork)
<serial number=""></serial>	::= see Table	Pn	serial number 1	Logical Recordset (Pages) Logical Recordset (Bibliography of Subwork)
<repeating symbol=""></repeating>	::= ':'	Nn A B J	serial number	Logical recordset (Serial Number of Story) Logical Recordset (Author of Story) Logical Recordset (Supplement about Story)
	,	JX		Logical Recordset (Keyword) Logical Recordset (Story/Tifle)
<original data=""></original>	::= see flag Rule	Ln	serial number	Logical Record (Upper Column)
<tag symbol=""></tag>	::= see Table	Mn Qn	serial number None	Logical Record (Lower Column) Logical Recordset (Postscript of Work)
<tag attribute=""></tag>	::= see Table	G	1	Logical Recordset (Postscript of Subwork) Logical Recordset (Picture)
		gn	None serial number	Logical Record (Title of Picture) Logical Recordset (Test in Picture)
		H	None	Logical Record (Table) Logical Record (Title of Table)
		nn	serial number	Logical Record (little of lable) Logical Recordset (Text in Table)

My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

KOKIN Rule 2: Serializing Rules

Layout of Japanese classical texts is two dimensional

- Main sentences and supplementary sentences (e.g., annotations, side notes)
 Many supplementary sentences are
- parallel to the corresponding main sentences
- KOKIN Rule 2 is introduced to serialize two-dimensional text layout

Original Text	KOKIN Markup
戸越	/戸越/ (トゴシ)
, 戸越里村	/戸越/ (ノ) 里村
サガミノカミ 相模守 和象守トモ	/相模守/(サガミノカミ 和泉守トモ)
^{泰時ノチ} タカトキ 平高時	/高時/(泰時ノ子#タカトキ)

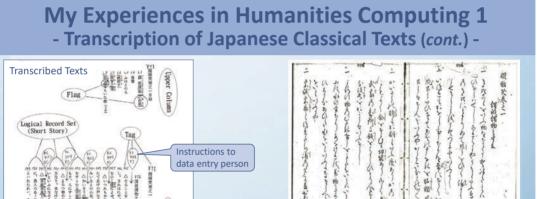
Example of KOKIN markup(Flag Rule)

My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

Schema of KOKIN Rule 2

<original data=""></original>	::= <flag begin=""><data element=""><flag end=""><supp> <dataelement><spaceflag> <supp><data element=""> <data element=""></data></data></supp></spaceflag></dataelement></supp></flag></data></flag>
<data element=""></data>	::= <string></string>
<flag begin=""> <flag en<="" td=""><td>d> <space flag=""> ::= '/'</space></td></flag></flag>	d> <space flag=""> ::= '/'</space>
<supp></supp>	::= <right supp=""> <left supp=""> <bi-supp></bi-supp></left></right>
<right supp=""></right>	::= <supp begin=""><supp element=""><supp end=""></supp></supp></supp>
<left supp=""></left>	::= <left begin="" supp=""><supp element=""><supp end=""></supp></supp></left>
<bi-supp></bi-supp>	::= <supp begin=""><supp element=""> ' ' <supp element=""><supp end=""></supp></supp></supp></supp>
<supp element=""></supp>	::= <single supp=""> <double supp=""></double></single>
<single supp=""></single>	::= <supp element=""></supp>
<double supp=""></double>	::= <supp element=""><supp separator=""><supp element=""></supp></supp></supp>
<supp begin=""></supp>	::='('
<left begin="" supp=""></left>	::= "("
<supp end=""></supp>	::= ')'
<supp separator=""></supp>	::= '#'
<supp element=""></supp>	::= <string> <string><string separator=""><string></string></string></string></string>
<string separator=""></string>	
<string></string>	::= see KOKIN Rule 3





うらけんだちね

らくれろれ

Original Texts

んちし

ノ浦ろり

ちちく

Target source for digital transcription わう

142

MA X (inthe

やいう

i

My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

Example of Digital Transcription

H Clother and Clot

11 BATHA

Lower

Column

A construction of the second s

actinings of

Logical Record

ころうちのう か

山田町町

方々に立出あた うからす。しかるあり

```
10
20
30
40
50
60
70
```

Instructions to

data entry person

- ★T最管抄 ※P41 ★T1.60官抄参第一 ★T2 (2) 第2年代 L7.20官式 (試んこ) 本天地人定/(ル) 後之首音也。 L8/三皇 (さんくわう) ムム/天連子 (てんくわうじ) ム/地皇子 (ちくわうじ) ム/人皇子 (じんくわうじ) L9又三皇ムム/伏徹 (ふくざ) ム/神皇 (しんのう) ム/黄帝 (くわうてい) L10五帝△ム/小曼 (ぜうかう) ム/窓m043600.8m043357; (せんざよく) ム/高辛 (かうしん) ム/勇 (げう) ム/勇 (しゆん) L11/三王 (さんわう) ムム/夏 (か) ム/殷 (いん) ム/周 (しう) ム夏ム十七王ム四百三十二年。殷日南ム三十 [王] ム大百十八年。 田人三十 80 90
- 100 110
- 周ムニー L12ムムムムムムムムムム七王ム八百六十七年。 L13十二諸役ム/第(〈テイ〉)ム/章(〈サウ〉)ム/束(〈ソウ〉)ム/音(〈シン〉)ム/香(〈エイ〉)ム/養(しん)ム/ 斉(〈セイ〉)ム/素(〈エン〉)ム/書(〈ロ〉)ム/葉(〈ソイ〉)ム/差(〈ソン〉)ム/酸(〈チム〉)ム/奥(〈ゴ〉)不入猶佚 L14/六圓(りつこく〉ム元王/(ノ)時六圓/溢(〈コハシ〉)。/義(〈カン〉)ム/酸(〈グキ〉)ム/燃(〈テウ〉)ム芳点素公差
- 130
- L15/秦(しん) △大帝△五十六年。
 L16/漢(かん) △十二帝△二百十四年。/王莽(わう〈マウ〉) △十四年。/更始(〈カウシ〉) △三年。
- ¥P42

- 140 150 160 170 180 200 210 220 230 240 250 260
- ¥P 4 2 L 1 / 後濃 (ごかん) ム十二帝公百九十五年。 L 2 三雄ムム/観(ぎ) ム/具(ご) ム/蜀(しよく) L 3 ム観ム五帝公田十五年。 L 4 ム具ム四帝公田十九年。 L 6 / 番(しん) ム十五帝。 L 6 / 番(しん) ム十五帝。 L 7 ム/個位(ぎゐ) ム贏帝以後。 L 8 ム帝燕公後派公後秦公後蜀公国美公南涼△/前(女) 秦△前涼△前燕△/後燈(こうてう) L 9 ム代為紙-北淡乙西派北派会後秦 L 1 0 南朝人北朝・・・・・・・



My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

Processing External Characters (外字)

KOKIN Description

天ノ/&m027888;(カトリ)ヲ以テ太子ヲ 糸曾

SGML/XML Description

天ノ<ruby fg="OFF"><rb><gaiji set="mojikyo" code="027888" /></rb><rt>カトリ </rt></ruby>ヲ以テ太子ヲ

HTML Description

天ノ<ruby><rb></rb><rt>カトリ</rt></ruby>ヲ以テ太子ヲ

My Experiences in Humanities Computing 1 - Transcription of Japanese Classical Texts (cont.) -

Example of Japanese Classical Texts

- ・愚管抄 (GUKAN-SHO) by KOKIN Description 🗍
- ・愚管抄 (GUKAN-SHO) by SGML/XML Description 🗍
- 愚管抄 (GUKAN-SHO) by PDF: SGML/XML to PDF by programs
- 愚管抄 (GUKAN-SHO) by HTML (Web): SGML/XML to HTML by XSLT 🔬

Gukanshō (最管控) is a historical and literary work about the history of Japan. Seven volumes in length, it was composed by Buddhist priest Jien of the Tendai sect c. 1220 (from Wikipedia).

My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents -

Historical Earthquake Documents in the Ancient and Medieval Ages in Japan (2003 -, Led by Professor Emeritus ISHIBASHI)



Construct a New Text Database

• Resources were printed materials and difficult to access (e.g., Out of print)

Resources were difficult for researchers to retrieve information (No databases)

High Quality Comprehensive Text: Critically examine all the collected documents, select

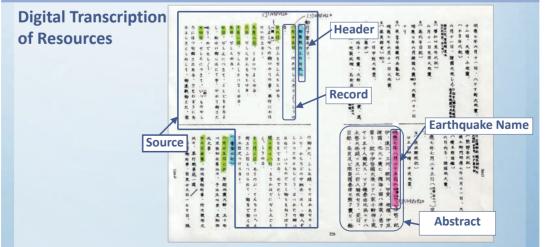
- appropriate ones, revise them correctly and reorganize as new documents in digital form • Each article was collected without precise examinations and text critiques (second and third sources,
- thread and thrum ...)
 Seismologists are difficult to critically read historical documents as they didn't take any pieces of training for historical materials

• Low reliability will lead to wrong seismologic conclusions as seismologists tend to believe written issues Compose a seismic intensity database compatible with the Intensity Data Points (IDP) prevailing in Europe

My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents (cont.) -







My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents (cont.) -

XML Markup Guidelines

XAll information needed for text critique and revision works must be preserved

- · Logical elements such as titles, records, sections, etc. are marked-up
- · Editorial notes and revision notes are marked-up
- Descriptive parts such as rubi and gaiji (external characters) are marked-up
- Layout structures such as the number of lines in a page are NOT marked-up
- Why XML?
 - ✓ Data Portability
 - ✓ Coping with Hypermedia
- Not use TEI
 - ✓ Simple Structure

Clas

✓ Tools (schema and programs) in the previous project (Transcription of Japanese Classical Texts) can be available

My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents (cont.) -

Basic Structure of the Historical Earthquake Documents

ssicEarthquake
Volume+
- Header?
- Earthquake+
- Header?
- E.ID
- J.Date
- S.Date
- E.Description
- E.Sources+
- Header?
- E.ID
- Source.ID
- Source.Record+
- Record.ID
- E.ID
- Source.ID
- J.Date
- S.Date
- Record.Descri

Root of the document A volume of historical earthquakes Bibliographic information of the volume An earthquake Biographic information of the earthquake ID of the earthquake (Must be unique) Date of the earthquake by Japanese calendar Date of the earthquake by European calendar General description of the earthquake Bibliographic information of the material ID of the earthquake (Must be unique) ID of the material (Must be unique Minimum unit on the earthquake information ID of the record (Must be unique) ID of the earthquake (Must be unique) ID of the material (Must be unique) Date of the event by Japanese calendar Date of the event by European calendar scription Contents of the record

My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents (cont.) -

An Example of XML Text

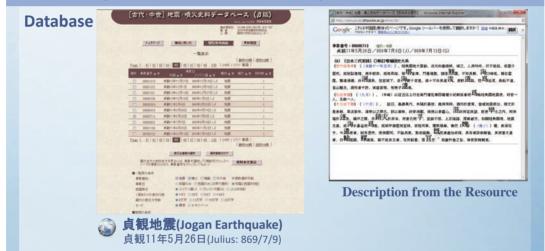
<?xml version="1.0" encoding="Shift_JIS"?> <?xml-stylesheet type="text/xsl" href="./ClassicEarthquake-Ext.xsl"?> <!DOCTYPE ClassicEarthquake SYSTEM

"/classicFarthquakeSimple_ver3.dtd"[]><ClassicFarthquake> <Volume vol="ZOTE!"><Header><titleStmt>增訂大日本地震史料</titleStmt>/Header>

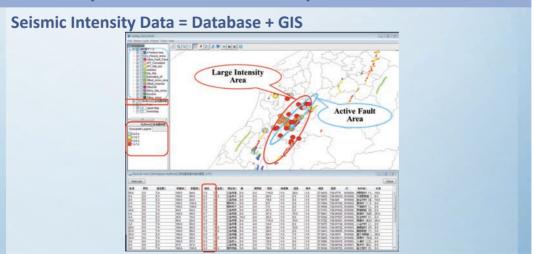
<source.Record page="228">
<Record.ID>00003733</Record.ID><E.ID>14980920</E.ID><Source.ID>00000348</Source.ID>
<I.Date>明應七年八月廿五日</I.Date><S.Date type="Gregorian"/>
<Record.Description><section><ruby><rb>廿五</rb>// ひうゆる、しよゆへの御いのりの事、ふ行におほせいたさるゝ、</section></Record.Description>
</Source.Record>
<

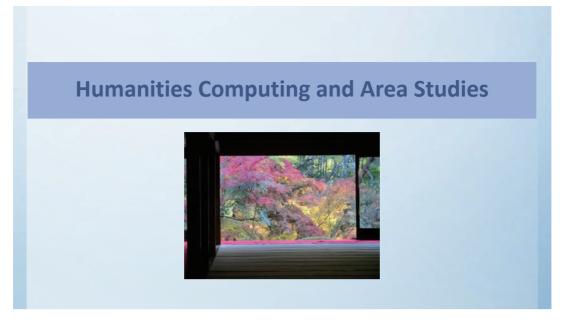
<i>.</i> .	ces in Humanities Computing 2 Historical Earthquake Documents (co	
Proofreading Document Outputs		Revenue Revenue 来 (續本明通鑑) * (續本明通鑑) * (續本明通鑑)

My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents (cont.) -

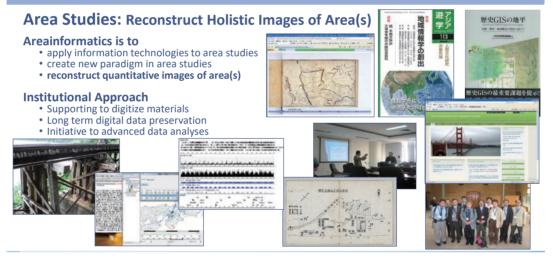


My Experiences in Humanities Computing 2 - Transcription of Historical Earthquake Documents (cont.) -





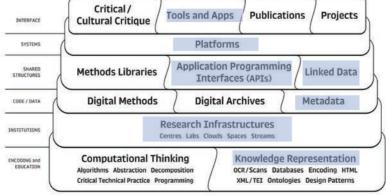
Humanities Computing and Area Studies - Area Studies and Area Informatic -

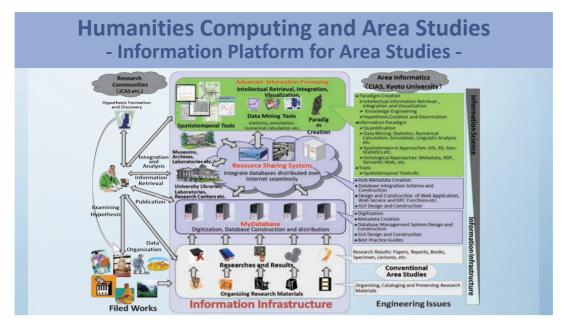


Humanities Computing and Area Studies - Area Informatic and Tools -

Digital Humanities Stack







Humanities Computing and Area Studies - Road Map of Our Development -

Phase1: Who and What

- Search for Resources
 - ✓ Bibliographic data
 - ✓ Search by limited vocabularies
- Resource Sharing Systems

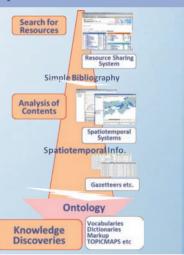
Phase 2: When and Where

Analysis of contents by spatiotemporal attributes

- Spatiotemporal model and tools
 - ✓ Overlay variety of maps, images, calendars etc.
 - ✓ Visualization, simulation, data mining etc.

Phase 3: Ontology

- Linking/integrating everything
- Organization of knowledge
- Knowledge discoveries



Humanities Computing and Area Studies

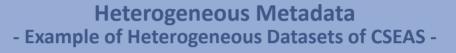


	leterogeneous Me es are the Basis of Re	
	Libraries, Archives	Researches
Target	Public	Individual / Research Group
Object	Public / General	Research /Specific
Collection Organization	Institutional	Individual / Research Group
Variety	Large	Large
Collection Policy	Consistent	Inconsistent / Changeable
Collection	Whole	Parts
Size	Large	Small
Metadata	Standard(general) / Large / Complex	Heterogeneous(Specific) / Small / Simple
Usage	Simple	Complex / Inconsistent
Durability (life time)	Long	Short

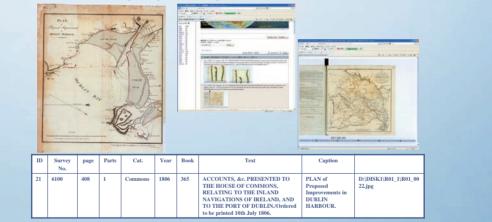
Our Challenges

Durable , interoperable and flexible repository of heterogeneous datasets
Key technologies: metadata + XML + HTTP + ontology

- 1. MyDatabase supports to build databases
- 2. Resource Sharing System links heterogeneous databases
- 3. Web API realizes flexible database integration and usage

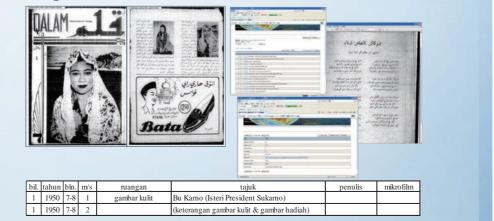


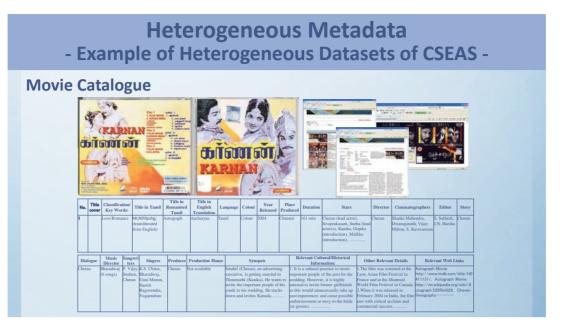
British Parliamentary Paper (Maps)



Heterogeneous Metadata - Example of Heterogeneous Datasets of CSEAS -

Islamic Magazines





MyDatabase - Coping with Heterogeneous Metadata -

MyDatabase: Server Function for Users (Researchers) to Build Heterogeneous Databases

Durable System

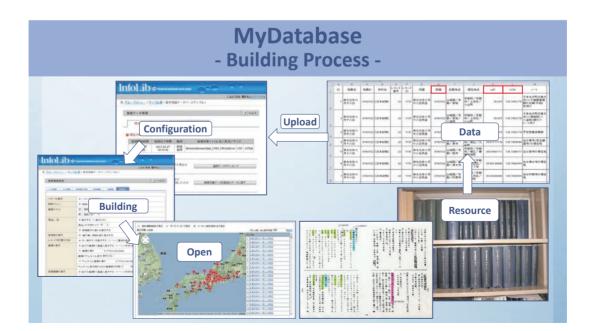
- Simple Functions ⇒ Minimum operations and constrains
 - ✓ Data portability (XML)
 - ✓ Basic retrieval functions

✓ Simple GUI

- **Simple Operation**
 - Simple Configuration (Minimum parameters)
 GUI

Minimum Constraints on Data Structure

- Simple data type (String)
- ✓ Key field or element
- Free from DD/DTD(Schema)
 - CSV/TSV data: first normal form (relational data model)
 - ✓ XML data: well-formed XML document



A	в	0	D	E	F	G	н	I	J	К	L	M	
ID	地震名	地震ロ	史料名	レコード 番号	レコード ID	和暦	西曆	記載地名	現在地名	LAT	LON	ノート	
	1 貞元元年六 月十八日	9760722	〔日本記略〕	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/宮城	京都府/京都 市/上京区/ 小山町	35.015	135.7452	千本丸太町交差点 近くに大極験基壇 跡の石碑(平成5 年)あり	
	2 貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/宮城/ 八省時	京都府/京都 市/上京区/ 小山町	35.015	135.7452	1千本丸太町交差点 近くに朝望 八省院) 対 レートあり	
	3 貞元元年六 月十八日	9760722	(日本紀略)	22	1151	Tabl	e (CS	V, TSV)	都府/京都 /中京区/ 樂週西町	35.015	135.7452	2778 平安宫囊:	
	4 貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/東寺	京都府/京都 市/南区/九 条町	34 97777778	135.7494	4444 左は東寺(国寺)の明	
	5 貞元元年六 月十八日	9760722	[日本記略]	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/西寺	京都府/京都 市/南区/唐 橋平垣町	34,97861111	135.7386	6111 左は西寺(
e 日本元石・ Proor22 (日本記報) 22 1151 「八江雪太二、 Proor22 (副の書示・ 「小ノ南区・海・ 「小ノ南区・海・ 「八田」 version=「1.0" encoding="Shift_JIS"?) (7ml version=「1.0" encoding="Shift_JIS"?) (7ml version="Listication type="Listication") (7ml version="Listication") (7ml version="Listi								大二震 千人 段二及					

							MyD - Me						
		ID		Te	erms		Fie Attrik			Гime		PI	ace
4	A	в	0 4	10	Ε	F	G	н	1/	J	К	1/	M
	ID	地震名	地贯ID	史料名	レコード 番号	1/1D	和層	西原	人教地名	現在地名	LAT	/ON	ノート
2	SID	ENAME	EID	RNAME	RID	RECID	JDATE	NDATE/	ONAME	PNAME	LAT	LON	NOTE
3.	ID	BASIC	PRIME	BASIC	BASIC	BASIC	BASIC	BASIC	BASIC	BASIC	BASIC	BASIC	BASIC
8	,	貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/宮城	京都府/京都 市/上京区/ 小山町	35.015	135.7452778	千本丸太町交差点 近くに大極殿基壇 跡の石碑(平成6 年)あり
	2	貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/宮城/ 八省院	京都府/京都 市/上京区/ 小山町	35.015	135.7452778	千本丸太町交差点 近くに朝堂院(=八 省院)跡のブレート あり
6	3	貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/宮城/ 豊楽院	京都府/京都 市/中京区/ 聚来遍西町	35.015	135.7452778	平安宫盘未散路
,	4	貞元元年六 月十八日	9760722	(日本記略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/東寺	京都府/京都 市/南区/九 条町	34,97777778	135.7494444	左は東寺(教王護 国寺)の現在地
3	5	貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/西寺	京都府/京都 市/南区/唐 橋平垣町	34.97861111	135,7386111	左は西寺の現在地
	6	貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/清水寺	京都府/京都 市/東山区/ 清水一丁目	34.99138889	135.7869444	左は清水寺の現在 地
0	7	貞元元年六 月十八日	9760722	(日本紀略)	22	1151	貞元元年六月 十八日癸丑	9760722	山城国/京 都/円覚寺	京都府/京都 市/上京区/ 柳風呂町	35 02638889	135.7694444	左は円覚寺の現在 地

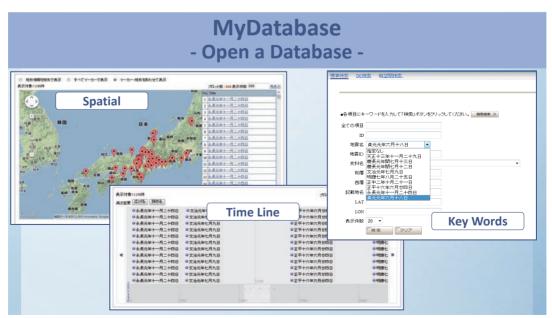
MyDatabase - Data Upload and Configurations -



2019 國際圖書館專業館員研習班:數位人文的發展與應用

		-		yDataba ute Defin		-		
NO	1	2	3	4	5	6	7	
日本語見出し	3D	1286	SEID .	史料备	V0-569	V3-F10	10	28
発調発出し	SID	ENAME	EID	RNAME	RID	RECID	JDATE	NDA
	48 10	NER SEARCH	ALL PONE	NER SEARCH			RE BASIC	-
MIT	フィールド連加	NER nout_order=asc 7x-1LFillS	フィールド通知	構築 input_order=ast フィールド連切	フィールド追加	フィールド連加	フィールド連加	2
単性 (プレビュー)	ID	SEARCH&input_type=dropdow &input_order=asc	PRIME	SEARCH&input_type=dropdown &input_order=asc	DATA	DATA	BASIC	BAS
1	1	真元元年六月十八日	9760722	(日本紀報)	22	1151	▲元元年六月十八日教丑	976
2	2	奠元元年六月十八日	9760722	(日本紀時)	22	1151	编元元年六月十八日癸丑	976
2	2	貞 元元年六月十八日	9760722	(日本紀編)	22	1151	桌 元元年六月十八旦癸丑	976
‡ ‡ ‡	1 1 2	* * *	4	4	4 14 15	4 .3 .4		1 1 1
1228	1233	要具元年間七月十二日	15960904	(本邦大地區极限)	501	4821	交缘五丙申年(彭子慶長元年)	159
1229	1234	要員元年壁七月十二日	15960904	(本邦大地靈樹寮)	501	4621	文排五河半年(彭子慶長元年)	199
1230	1235	慶長元年壁七月十二日	15960904	(本邦大地豪俊表)	501	4821	文埼五丙申年(彭子慶長元年)	159
(•





The Development of Information Platform for Humanities Computing: the Case of Center for Southeast Asian Studies (CSEAS) Kyoto University

數位人文資訊平台之建置:以日本京都大學東亞地域研究所為例



MyDatabase Web API - Overview -

MyDatabase

A large and complex system and hides internal mechanism and complexities from users by

- Limiting the functions available to users
- No data types
- ✓ metadata is easy to define and modify
- Restricting database accesses via API (Application User Interface)
 - \checkmark no need expertise knowledge and skills concerning database systems
 - ✓ easy-to-build user interfaces and applications.

MyDatabase Web API for User Applications

MyDatabase Web API is specifications and access points for user applications

- Hide internal mechanism of MyDatabase
- Web services should be built by users
- REST (Representational State Transfer) -like Web Service
 - CGI programing
 - MyDatabase API returns HTML/XML/JSON

MyDatabase Web API - Query Example -

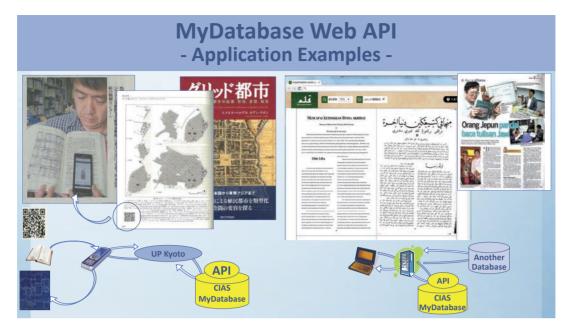
Search for records whose temple name (c3) is 相国寺 (Shoukoku-ji) (%E7%9B%B8%E5%9B%BD%E5%AF%BA) and get resumes in simple JSON form

http://some_uri/simplejson/G0000005geoname?operation=searchRetrieve&version=1.2&query=(c3="%E7%9B %B8%E5%9B%BD%E5%AF%BA")&recordSchema=original

{"c1":"10026682",	"c2":"10026682",
"c3":"相国寺",	"c4":"ソウコクジ",
"c5":"35.03333333",	"c6":"135.7627778",
"c16":"山城", "c18":"上京区", "c22":"京都市上京区////", "c25":"大日本地名辞書", ••••••}	"c19":"为ミキョウ", ····· "c23":"京都市上京区////", "c27":"30048501:相国寺(

•••• 生重複"••••

2019 國際圖書館專業館員研習班:數位人文的發展與應用



Resource Sharing System - Overview -

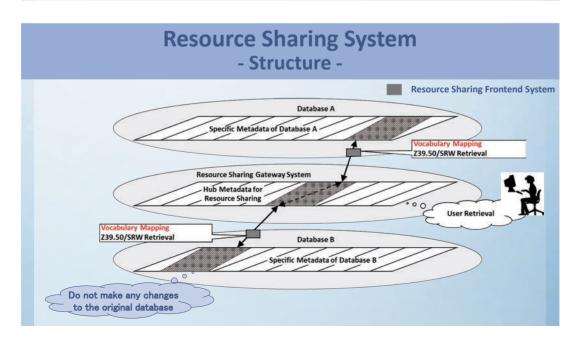
Metadata of Research Databases are Heterogeneous

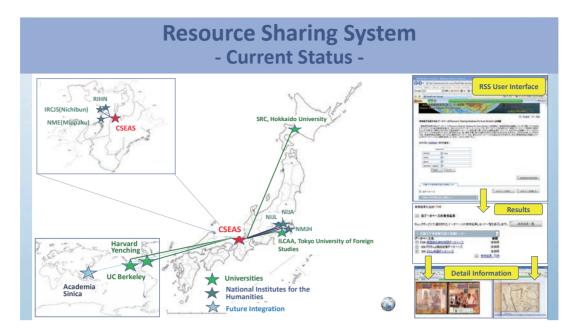
Resource Sharing System (RSS)

- Resource Sharing System: is a framework to retrieve various databases on the Internet seamlessly
- Each Database: has an intrinsic metadata in accordance with a researcher's research purpose
- Seamless: means that users can retrieve all database shared by Resource Sharing System without conscious of record structures, retrieval operations, database locations, and medias

Applying Some Standards

- Database: Portability (XML)
- Data structure: Standard "HUB" Metadata independent of shared databases (Dublin Core, MODS)
- Retrieval: Standard Information Retrieval Methods (Z39.50, SRW)



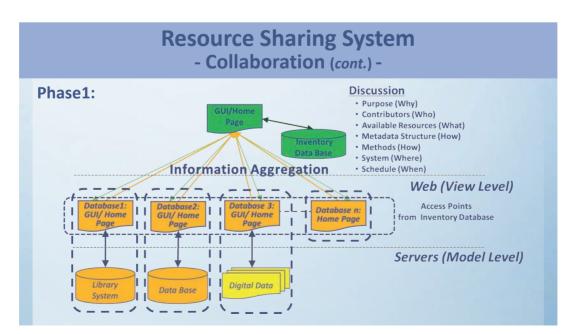


Resource Sharing System - Collaboration -

Academic Asset Preservations and Sharing in Southeast Asia (RsDA)

- The workshop intends to share local needs and seeds for digital archiving and preservation of scholarly assets and to discuss ideas to develop a network to promote collaboration in the region. The main aims of the workshop are to:
 - 1. understand and share the current circumstances about academic digital assets in the leading ASEAN academic memory institutions
 - 2. provide opportunities for the participants to exchange and share thoughts, views and ideas about how we should deal with resource sharing and digital asset preservation, what kind of technology will be available and who should be contributors to this issue
 - 3. discuss practical approaches to build a network fo collaboration rather than presentation of papers
- Venues: 2017 (Kyoto, Japan), 2018 (Shah Alam, Malaysia), 2019 (Khon Kaen, Thailand)
- Participant: Malaysia, Philippines, Singapore, Taiwan, Thailand, Japan







Latest Challenges 1 - Linked Open Database (Motivation) -

Reconstruction of the Platform

Current Resource Sharing System is not Flexible to Integrate/Link Databases

- Flexible links between university databases and cyberspace will create large-scale knowledge databases
 - ✓ Basis: data model, linked data, URI, ontology etc.
 - ✓ Applications: text mining, natural language processing, artificial intelligence, machine learning etc.

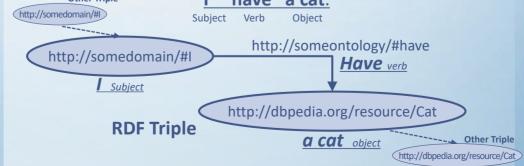
Current Resource Sharing System is Impossible to Automatically Develop Links

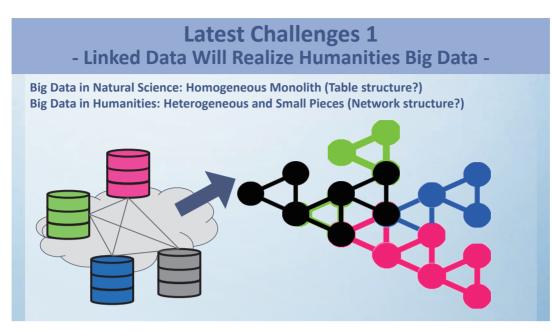
- Development of applications to discover useful hints/knowledge from large-scale databases
 - ✓ intelligent search engine, ontology etc.

Latest Challenges 1 - What is Linked Open Data? -

Linked Open Data: structured data which is interlinked with other data built on Web technologies such as HTTP, URI and RDF (Web of Data) • RDF (Resource Description Framework): way of knowledge representation

Other Triple I have a cat.





Latest Challenges 1 - Development of Linked Open Data Platform -

So far

MyDatabase

- Easy-to-use & schema-free database builder
- Humanities researchers can store their data as they want
- MyDatabase API enable a user to build own user interfaces and applications

Resource Sharing System

- Metadata mapping
- Standardized API (SRU)

Next (under construction)

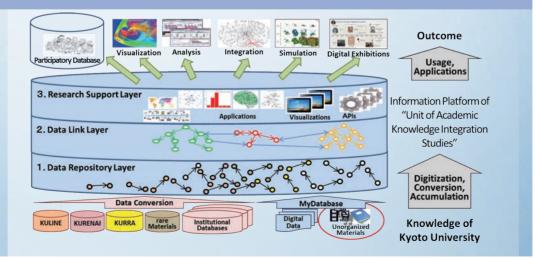
MyDatabase-LOD

- Follows MyDatabase user interfaceAutomatically transform table structure
- to RDF and assign URLs
- SPARQL endpoint and the resemble API as MyDatabase
- RDF Editor for editing RDF data and semantic annotations

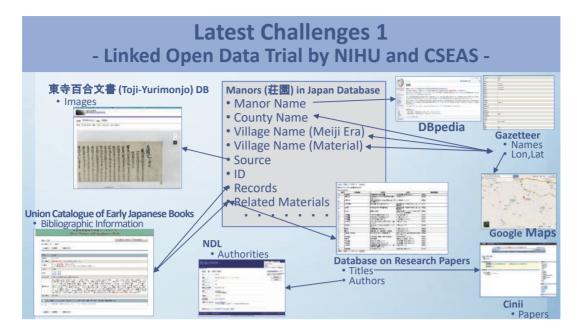
Resource Sharing System

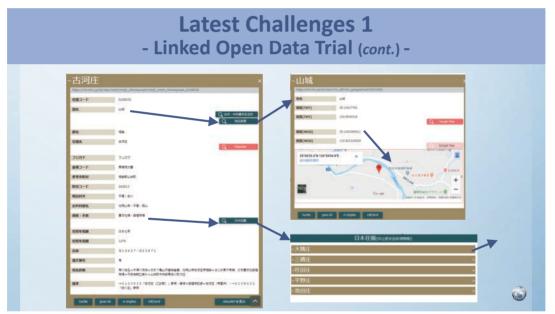
• Linking data by URIs and vocabulary mapping using ontology

Latest Challenges 1 - Schema of New Platform -



2019 國際圖書館專業館員研習班:數位人文的發展與應用





Latest Challenges 2 - Spatiotemporal Infrastructure (Motivation) -

Digital Gazetteers

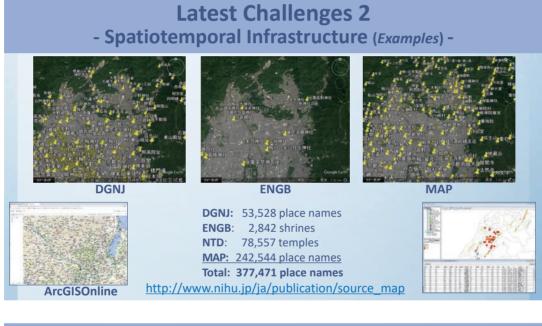
Essential Knowledgebase to Associate Heterogeneous Data Sets in the Context of Geographical Proximity

- Quantitative approaches to humanities data
 - ✓ paleo-earthquakes researches etc.

There Are Some Free Digital Gazetteer Databases but not Comprehensive

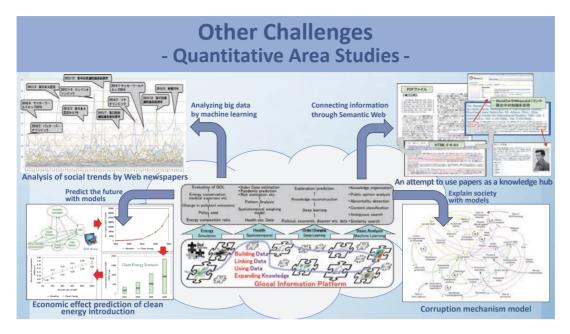
- Getty Thesaurus of Geographic Names
- Address Matching Service by CSIS University of Tokyo
- But not historical ones

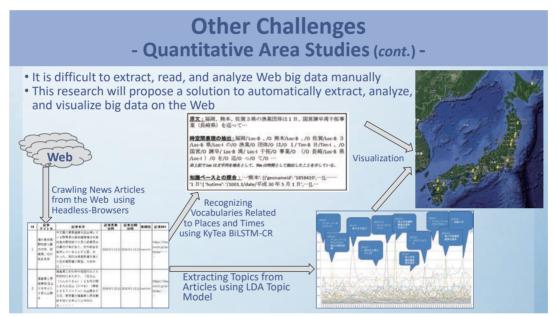






2019 國際圖書館專業館員研習班:數位人文的發展與應用





Thank you and Questions?

- 1. Digital Humanities or Humanities Computing
 - Short history
 - Digital humanities or humanities computing in Japan
- 2. My Experiences in Humanities Computing
 - Transcription of Japanese classical texts
 - Transcription of historical earthquake documents
- 3. Humanities Computing and Area Studies
 - Area studies and area informatics
 - MyDatabase
 - Resource sharing system
- 4. Latest Challenges
 - Linked open database
 - Spatiotemporal infrastructure
 - Web big data



An introduction of social analytical approach to digital humanities 社會網路分析應用於數位人文

Muh-chyun Tang 唐牧群

Associate University Librarian & Professor, Library & Information Science, National Taiwan University 國立臺灣大學圖書資訊學系教授兼圖書館副館長

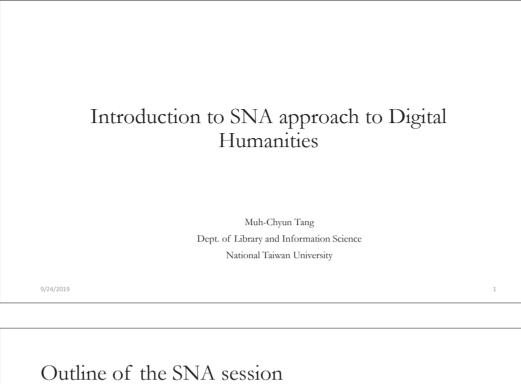
Abstract 摘要

Originated in sociology to study the dynamics among social actors, Social Network Analysis (SNA) has been widely adopted in a variety of fields. For humanities scholars, SNA affords a variety of visualization methods to measurement to reveal relationships and structure among entities such as historical figures and words. Aimed for humanities scholars who are interested in how SNA might assist their research, the talk will introduce three basic groups of SNA measurements, namely, "cohesion", "centrality", "community-detection." Rationale behind these measurements and their applications on previous cases will be provided to help beginners of SNA to adopt a SNA perspective to elaborate and develop research questions in their domains.

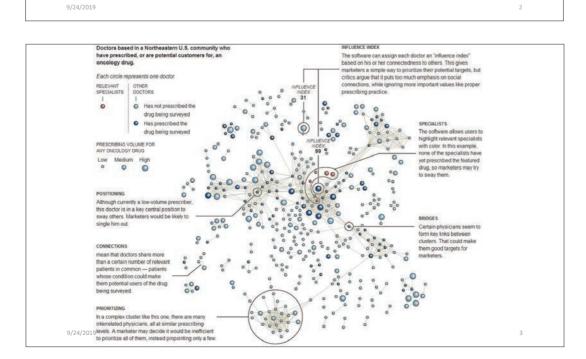
社會網絡分析(SNA)起源於社會學中其研究社會行動者之間的動態,而目 前已被廣泛應用於各個領域。對於人文學者而言,SNA提供了多種視覺化的方法 來檢視資料,以呈現實體之間的關係和結構,如歷史人物及文字。對於有興趣了 解 SNA 如何協助人文領域研究進行之人文學者,本講座將介紹三個基本的 SNA 測 量方式,包含「凝聚性」、「中心性」以及「社群檢測」。這些方法背後的基本原 理將提供測量及其在先前案例中的應用,以幫助社會網絡分析的初學者採用 SNA 視角來闡述和發展其領域的相關研究問題。

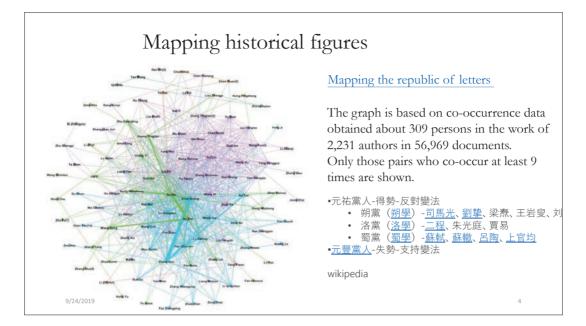
An introduction of social analytical approach to digital humanities

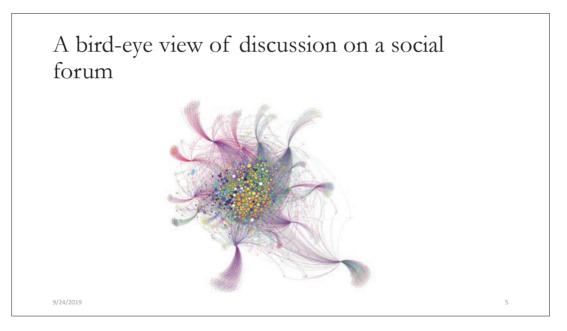
社會網路分析應用於數位人文

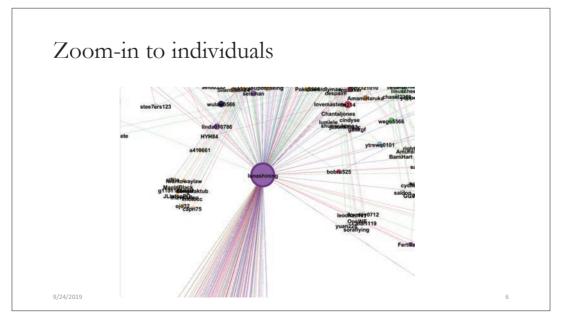


- What is SNA?
- Relational data
- What can SNA do?
- A domain analysis of digital humanities

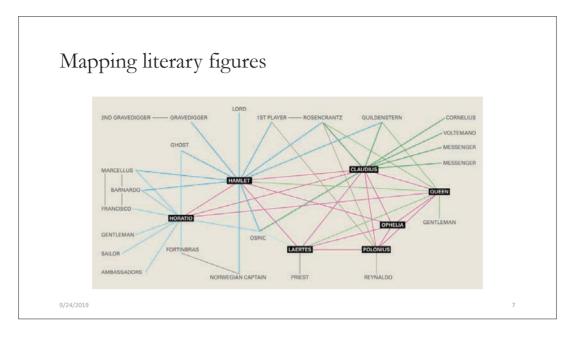


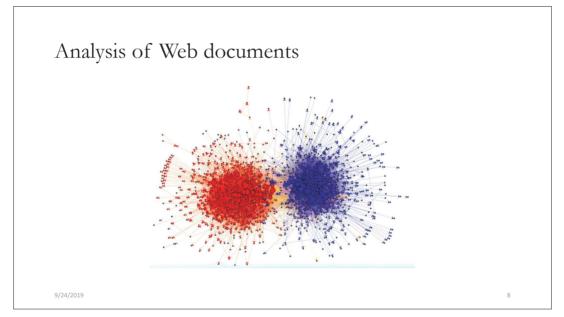


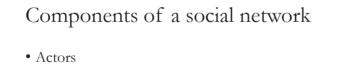




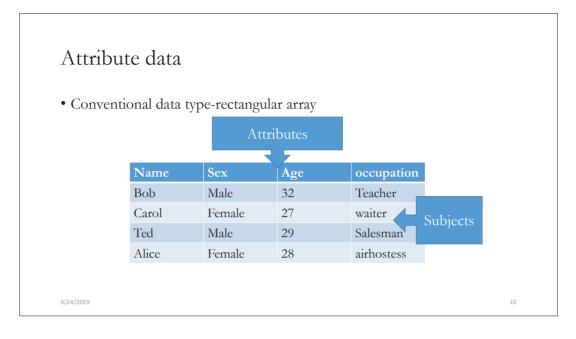
社會網路分析應用於數位人文

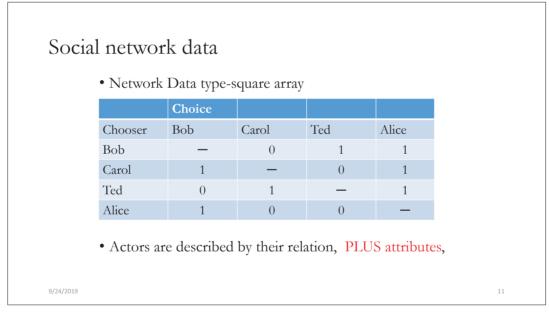


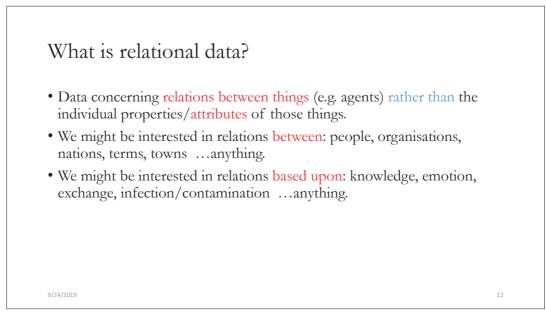




- Attributes?
 - E.g. demographic, psychological/personality traits
- Relationships
 - What are the attributes of relationship?







社會網路分析應用於數位人文

Relational data

- Property of two or more units,
- · Relations among units
- Example:
 - Distance between locations
 - Correlation between variables
 - Similarity/occurrence between objects
 - Interaction among individual
 - Diffusion of innovation between organizations

9/24/2019

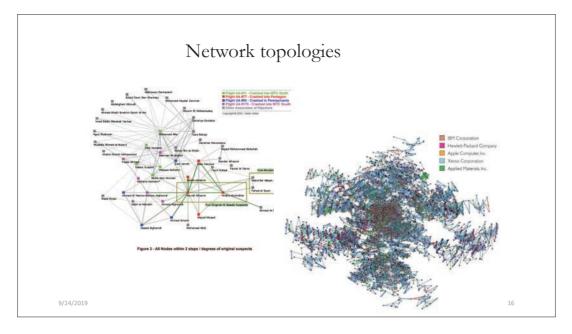
What can SNA do? • Identify structures – Display and analyze the **forest** given data on trees E.g. Visualizing knowledge domainOpen or close ego network etc. • Identify subsets of networks – Derive meaningful subgroups Cohesive Structural equivalent Identify actors who occupy strategic positions · Bridge and boundary spanners • Identify dynamic - flow of information, influences, diseases etc. • Establish "causal" relationship Attributes and relationshipPerformance and relationship · the number of separate social connections between any two actors 9/24/2019 14

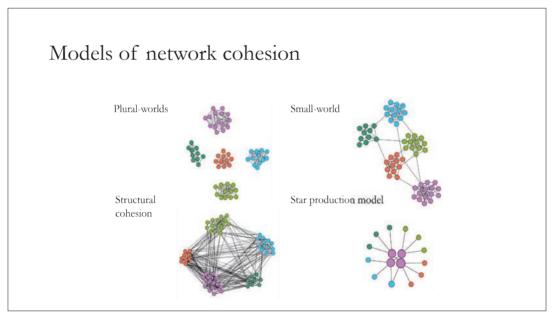
Network analytical measures

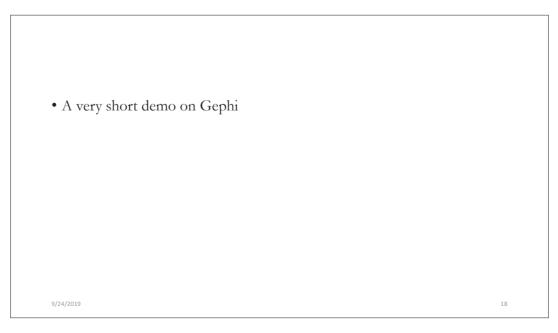
- Cohesion
 - Is the network held together or fragment?
- Centrality
 - Who is the central figure in the network?
- Community detection
 - Are there subgroups within the network?
 - Community detection, group partition

9/24/2019

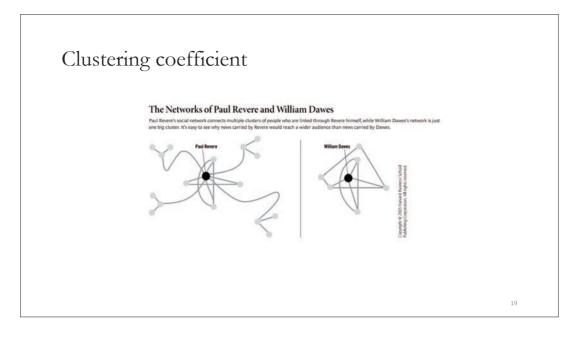
2019 國際圖書館專業館員研習班:數位人文的發展與應用

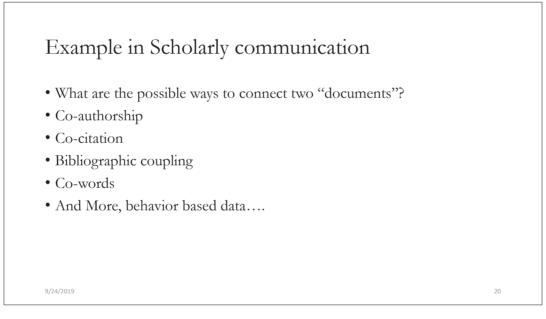


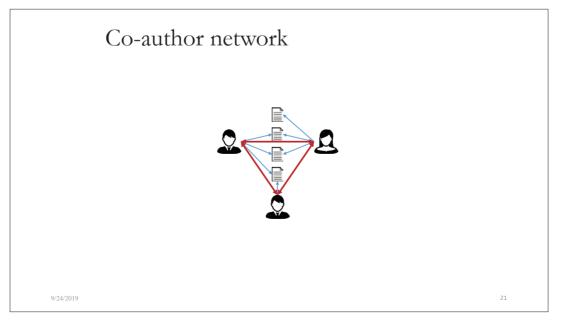




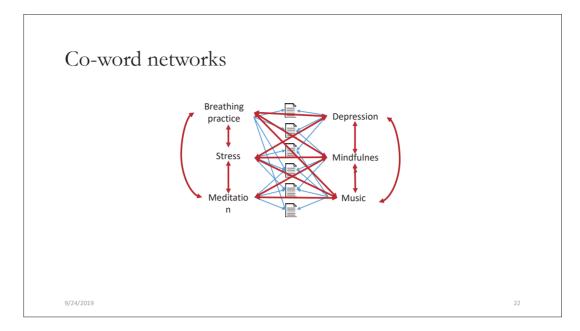
社會網路分析應用於數位人文

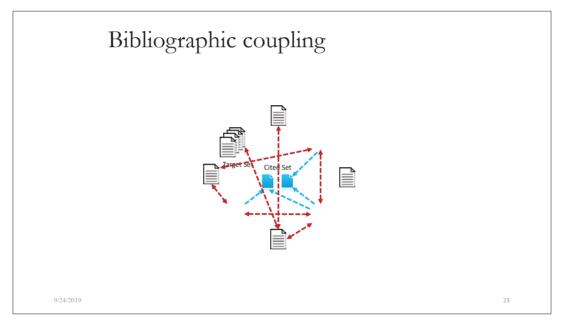


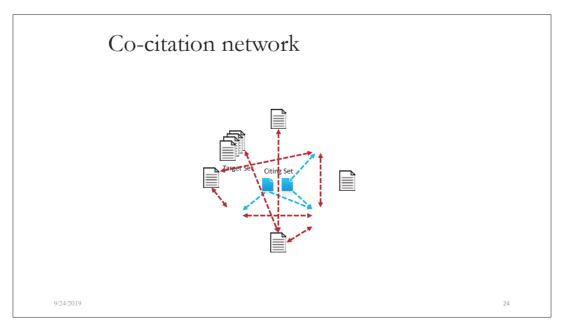




2019 國際圖書館專業館員研習班:數位人文的發展與應用







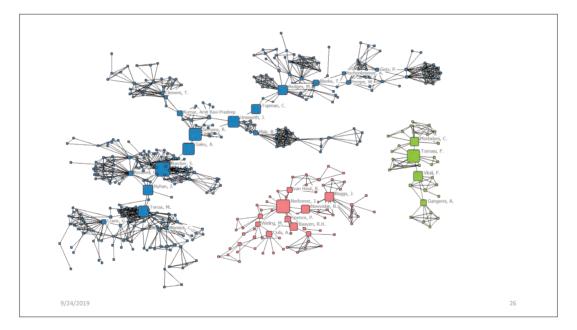
An introduction of social analytical approach to digital humanities

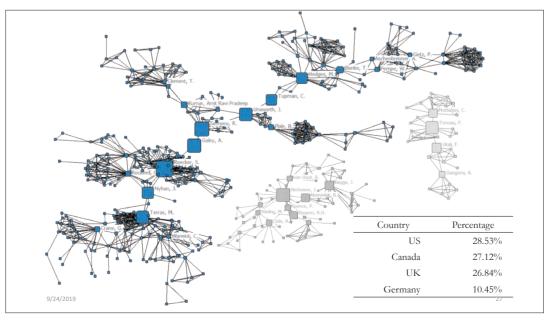
社會網路分析應用於數位人文

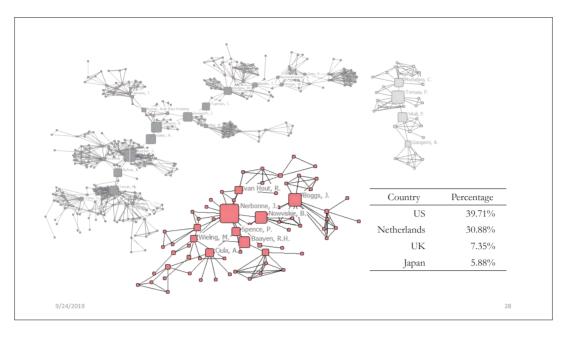
25

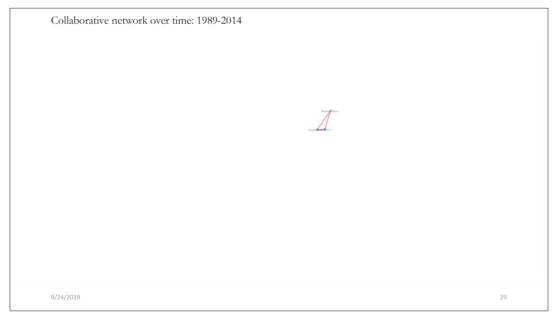
• Tang, M. C., Cheng, Y. J., & Chen, K. H. (2017). A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses. *Scientometrics*, *113*(2), 985-1008.

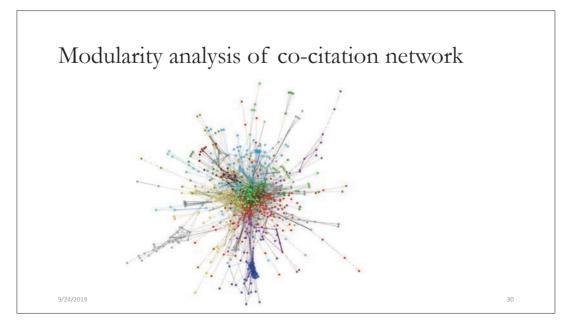
9/24/2019





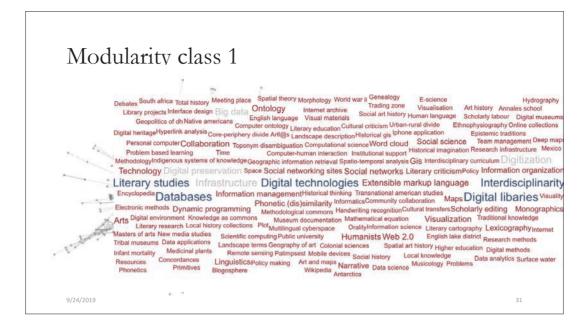


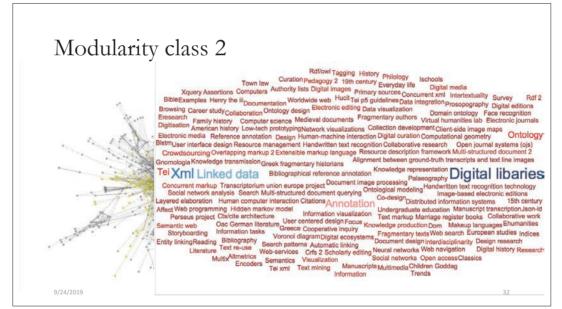


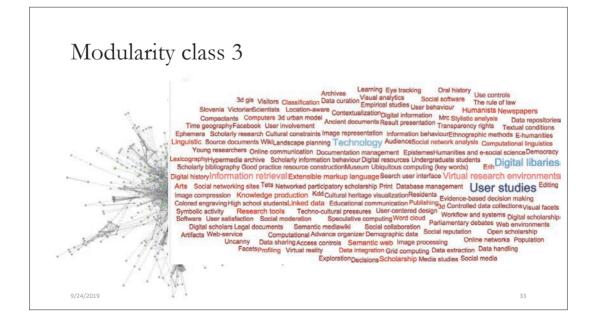


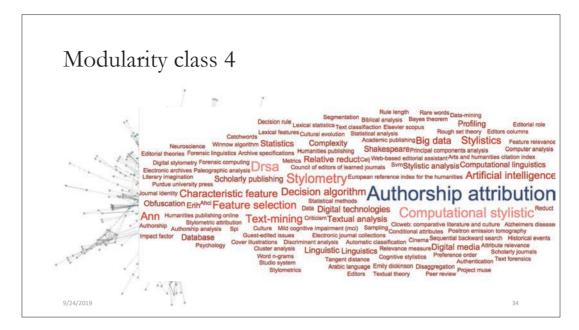
An introduction of social analytical approach to digital humanities

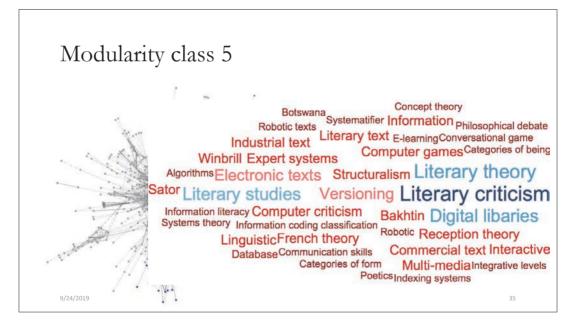
社會網路分析應用於數位人文

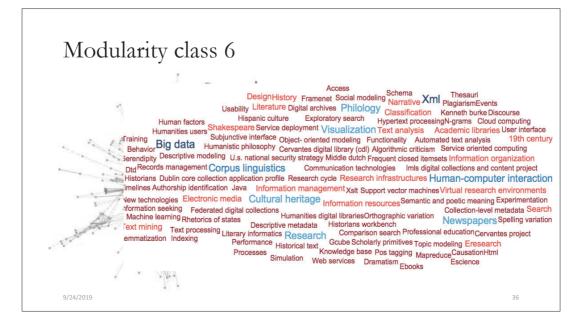






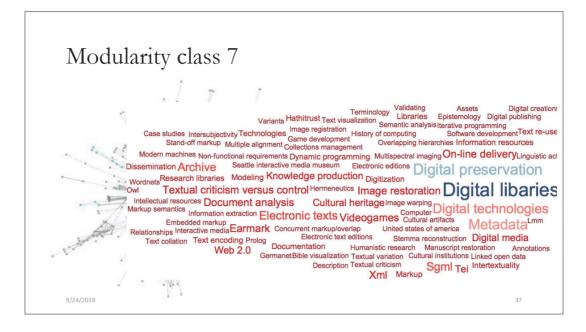


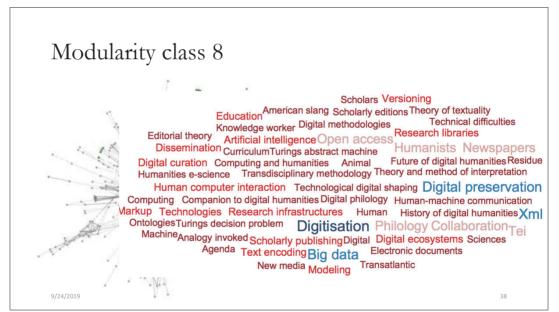


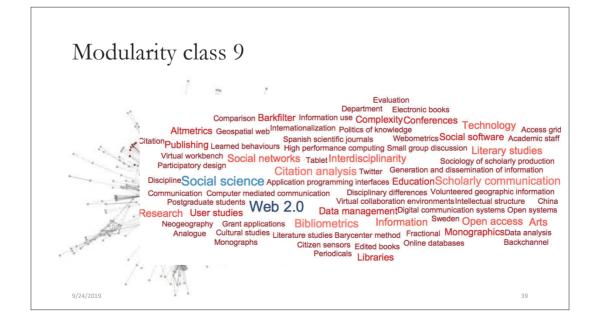


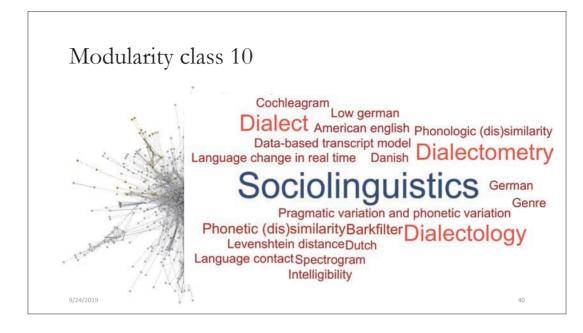
An introduction of social analytical approach to digital humanities

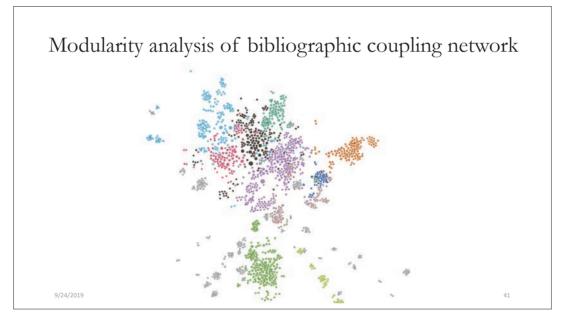
社會網路分析應用於數位人文

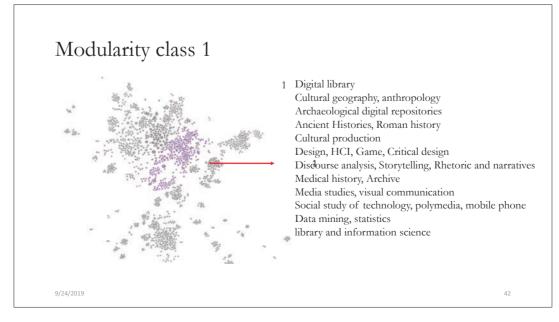




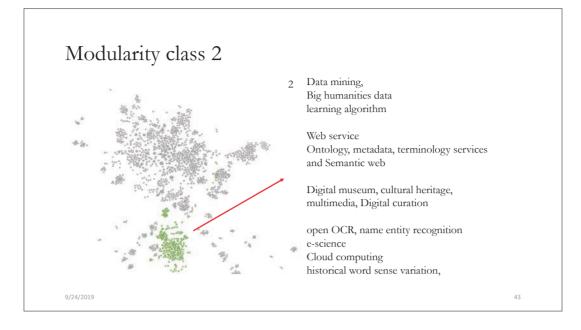


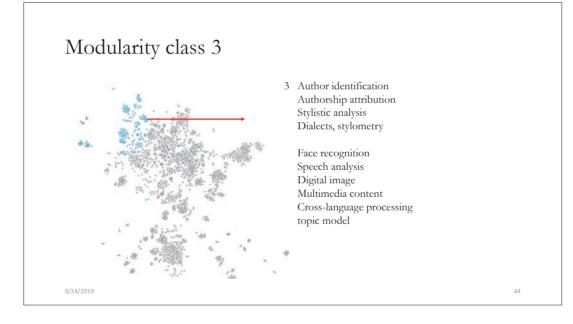


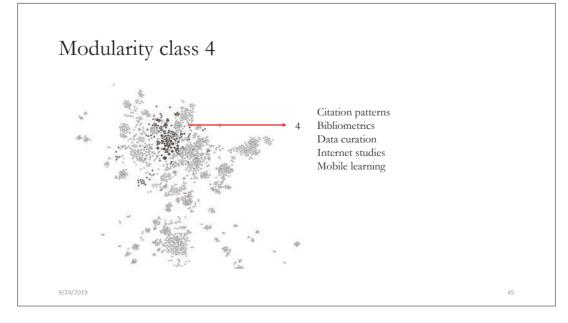


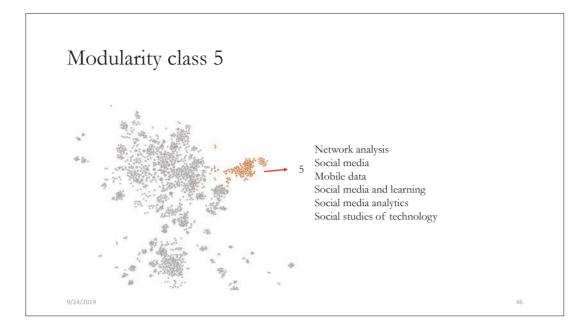


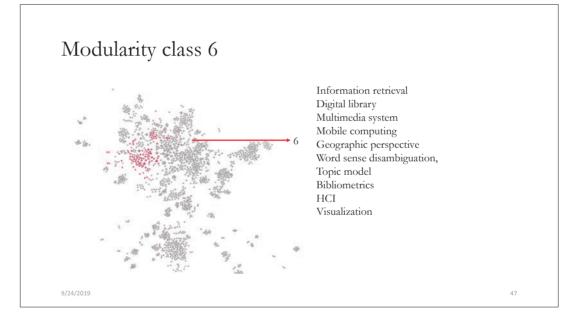
社會網路分析應用於數位人文

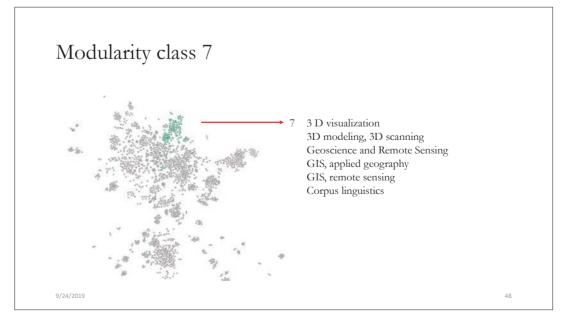






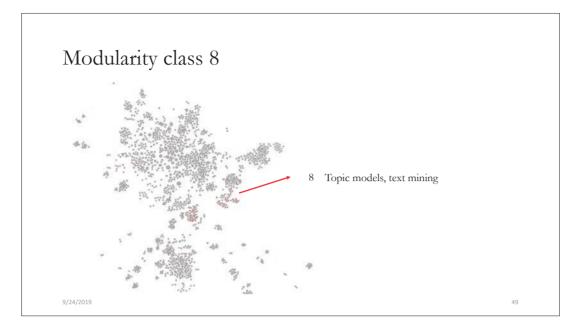


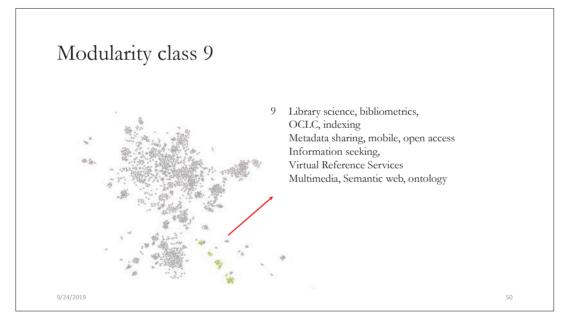


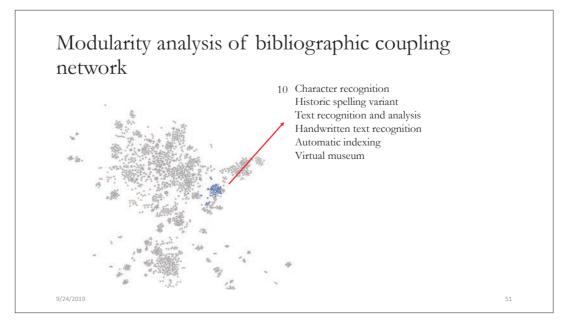


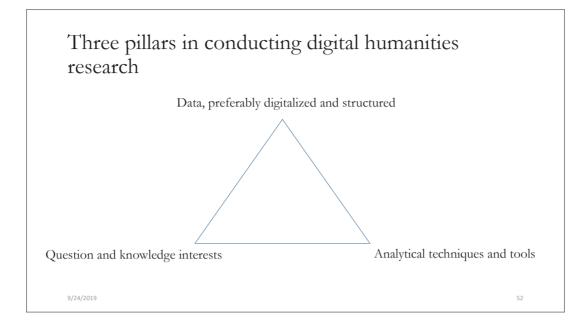
An introduction of social analytical approach to digital humanities

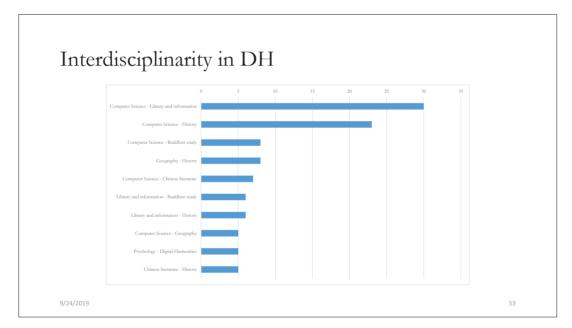
社會網路分析應用於數位人文











Digital Infrastructure, Bibliographic Data, and Chinese Print History 數位基礎建設、書目資料以及中國印刷史

Paul Vierthaler

Assistant Professor, Chinese Studies at the College of William & Mary, USA 美國瑪麗與威廉學院漢學研究 助理教授

Abstract 摘要

In this talk, Paul Vierthaler will discuss recent trends in the digital humanities in Europe as well as his own digital research into "quasi-historical" narratives written in late Ming and early Qing China. He will begin by discussing recent developments in digital humanities infrastructure in Europe. He will introduce database and toolbuilding projects at several universities and libraries across Europe, focusing in particular on projects in Asian Studies. These efforts are now fostering wide-spread adoption of the digital humanities in Europe and across the world. After discussing these general trends, he will introduce his own research, which directly applies numerous recent developments in the digital humanities. In this part of his talk, he will discuss the work he has done developing and utilizing large bibliographic datasets that he has derived from premodern bibliographic compendia, public domain digital library catalog records, and modern bibliographic research. Vierthaler uses these datasets to study trends in late imperial publishing and his goal is to develop proxies that measure the historical attention span of late imperial authors. Vierthaler combines this with research he has done text mining late imperial documents to study how historical information moves through a lattice of fictional, semi-historical, and unofficial historical documents in the late Ming and early Qing.

在本次演講中,Paul Vierthaler 教授將介紹數位人文研究於歐洲的最新發展 趨勢以及他自身進行關於明末清初的「準歷史性」敘事的數位研究。首先 Vierthaler 教授將為大家說明歐洲數位人文相關基礎建設的發展現況,並將介紹 在歐洲的幾所大學和圖書館其相關資料庫庫和工具開發專案,尤其著重於亞洲

相關的研究項目。而這些努力正在促進歐洲和世界各地的數位人文領域的廣泛 應用。除了概括性的總體趨勢介紹,Vierthaler 教授亦將介紹他自身的研究,該 研究直接應用了數位人文研究許多近期的相關發展。在這一部分的演講, Vierthaler 教授將介紹其開發並利用了大量的書目資料集,這些書目資料是他從 現代化之前的書目概要、數位圖書館書目記錄和現代書目研究中所獲得的。 Vierthaler 教授利用這些數據資料來研究晚清中國出版的趨勢,他的目標是開發 可衡量晚期清中國作家歷史意識範圍的工具。Vierthaler 教授將這與他進行的晚 清文獻資料探勘的相關研究結合,以研究歷史資訊如何藉由明末清初的虛構, 半歷史和非官方的歷史文獻進行傳播。

Digital infrastructure, bibliographic data, and Chinese print history

Professor Paul Vierthaler, College of William & Mary

Outline

Part 1: European Infrastructure for the Digital Humanities

Part 2: Digital Bibliographic Data as Research Data

Part 3: Digital Analysis and Chinese Print History

Part 1: European Infrastructures for Digital Humanities

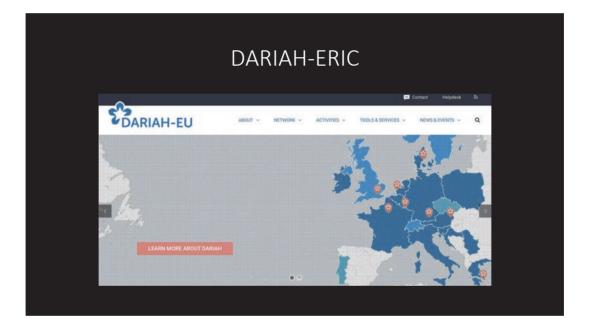
2019 國際圖書館專業館員研習班:數位人文的發展與應用

The "Big Tent" of Digital Humanities

Digital Access

Digital Presentation

Digital Analysis

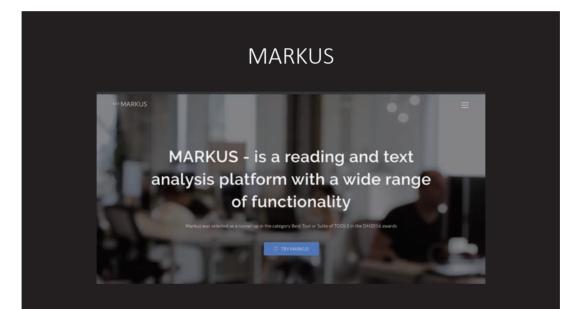




Digital Infrastructure, Bibliographic Data, and Chinese Print History

數位基礎建設、書目資料以及中國印刷史

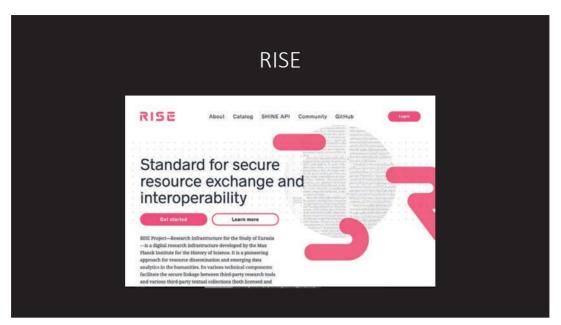




Chinese Text Project

928 ZA	🐖 Chinese Text Project 🛛 🖪 🙆 🕵 🗵
C[About the site] Introduction Fort Inst page Distorts Creations Creations Creations Creations Creations Creations	Welcome Welcomes Test Project hampage. The Olineer Test Project is an online upon access digital library that makes per-molers Olinear toos handled to makes and executives all anound the work. The site attempts to make upon access digital relation to equive new ways of heteractions with there tests that are not possible on port, With over Test thousand Stille and more than fire billion characters, the Olinear Test Project is also the largest distance of the moleculary. The calaxies.
C.com Statistics System Statistics Digital humanities Pre-Sin and Hen Confucientian Disblant Depailer Disblant dis Mannes Disblant dis Mannes	You may with to <u>test more about the protect</u> , were the <u>even on and this <u>particular</u> or this table of cancers, or <u>conset</u> the <u>instructions</u>, <u>HO</u>, or is of <u>HO</u>, <u>H</u>VWO isolation for a particular threater to the viscal particular balance of the state. BUENO - R2SEA - CRX - BEEL - TRANSREZ_PARTICULAR</u>
Another State Medicary Mathematica Misselfameous Schools Another Classics Another Classics Another Classics Chronics Medicare Catomate	This the is provided as an open scores resource. If you use this site is your teaching or research, please ask your university thraw to <u>subtoch</u>
Post-Hen OWel, Jin, and North-	Latest additions
Cost-Inco Disco-Ming Data Ting Disco-Ming Disco-Ming Disco-Ming Discussion Discussion Library	Context Encoded Harvard Versching Library Chinese materials added Tanks to the sport of Harvard Versching Library, cert 3 million space of skanned materials from the Yanching Library collection has 2016-10-10 Internet of the fact days and the fact fact days and

2019 國際圖書館專業館員研習班:**數位人文的發展與應用**





Centres for Digital Humanities and Scholarship

Centres for Digital Humanities

UK: King's College Exeter University College London Netherlands: Amsterdam Utrecht

Germany: Leipzig Göttigen Cologne Trier Max Planck Belgium: Liege Leuven

Digital Humanities at Leiden University

Centre for Digital Humanities

Centre for Digital Scholarship

Centre for Data Science

Part 2: Digital Bibliographic Data as Research Data

I: Structured Bibliographic Data

I: Structured Bibliographic Data

MARC Records as Research Data

I: Structured Bibliographic Data

WorldCAT

2019 國際圖書館專業館員研習班:數位人文的發展與應用

I: Structured Bibliographic Data

Digital Public Library of America

I: Structured Bibliographic Data

The Chinese Rare Book Project

II: Semi and Unstructured Bibliographic Data

II: Semi and Unstructured Bibliographic Data

Dynastic Histories (e.g. 明史藝文志)

《<u>日曆</u>》一百卷、洪武中,詹同等編,具載 太祖徵討平定之績,禮樂治道之詳。《<u>寶訓</u>》 十五卷《<u>日曆</u>》即成,詹同等又請分類更輯 聖政為書,凡五卷。其後史官隨類增至十五 卷。

II: Semi and Unstructured Bibliographic Data

Late Imperial Bibliographies (e.g. 四庫全書總目提要)

·百三十卷〈江蘇巡撫採進本 史記集 冟解∙ ,河東聞<u>喜人,官至南</u> 駰字龍駒 傳。馺 □阆(,殊恨省略 乃採 發明 ,別撰 E 及衆書之 經諸史 並漢書音義 此書...

II: Semi and Unstructured Bibliographic Data

Modern Bibliography (e.g. 中國通俗小說總目)

III: Unifying Disparate Data

III: Unifying Disparate Data

Disentangling the Work, Edition, Item, and Supplementary Information

III: Unifying Disparate Data

Building a Flexible, API-drive BiblioGRAPHic Database

III: Unifying Disparate Data

Neo4J and a Graph-based view of Chinese Bibliographic Information

2019 國際圖書館專業館員研習班:數位人文的發展與應用

III: Unifying Disparate Data The Magpie Project

Part 3: Digital Analysis and Chinese Print History

Extracting Relevant Data

Extracting Relevant Data

Identifying Pertinent Information

Does it exist in a Structured Field?

Can I get it from a natural language statement?

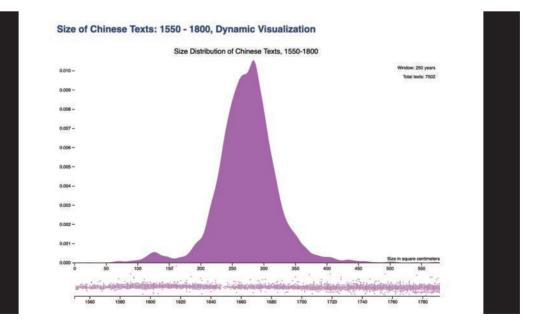
245-01/\$a史記鈔:\$b九一卷/\$c茅坤編選. 260-02/\$a[China]:\$b吳興閔氏,\$c明泰昌元年[1620] 500-03/\$a朱墨套印本. 500-04/\$a九行十九字,左右雙邊,白口,無魚尾,書眉上刻評,框高 20.9 x 14.4. 500-05/\$aSeals: "而慎"[...et al.].

510-06/\$a著錄:內閣文庫漢籍分類目錄,中國古籍善本書目,國立中央圖書館善本書目.

600-07/\$a司馬遷,\$dapproximately 145 B.C.-approximately 86 B.C.\$t史記\$xAbstracts

九行十九字 左右雙邊 白口 無魚尾 書眉上刻評 框高 20.9 x 14.4

Regular Expressions (d+(d+)?) s?x s?(d+(d+)?)Aggregating the Data



Linking Textual Content With Bilbiographic Data

Tracing Historical Narratives in Untrustworthy Genres

The Historical Focal Length of Late Imperial Narratives

Digital Humanities Infrastructures Enable Digital Research

Thank you!

Leiden University Centre for Digital Humanities NVIDIA College of William & Mary OCLC National Central Library Shannon Stewart, Hilde De Weerdt, Tina Lu

Questions?

Twitter: @pvierth

Email: pavierthaler@wm.edu

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

玩庫博:人文社會學者的中文文本探勘工具

Ho-chia Chueh 闕河嘉

Associate Professor, Department of Bio-Industry Communication and Development National Taiwan University 國立臺灣大學生物產業傳播暨發展學系副教授

Abstract 摘要

This presentation will introduce CORPRO, a Chinese language corpus-based textmining software for humanities and social science scholars. CORPRO is designed to incorporate self-complied corpus and self-setup corpus analysis condition, including self-defined dictionary, stop-word, words-grouping. Corpus analysis functions include term frequencies, collocations, keywords in corpus and concordance, all with related statistics. CORPRO provides humanities and social researchers (without programming skills) to conduct textual-mining independently. Its interface allows researchers to easily explore interesting features of corpus, and to repeatedly analyze corpus by setting different analysis conditions. That is, humanities and social science researchers can continuously analyze corpus on their own while keeping their subject domain knowledge in mind. It is in this way to increase dialogues between corpus analysis and subject theory. This presentation will use Cao Xueqin's the Dream of Red Chamber as corpus to demonstrate CORPRO, including two topics: significant difference between two sub-corpora and network of characters.

此次演講將介紹「庫博中文語料庫分析工具」。「庫博」是以語料庫語言學 為基礎的電腦輔助文本分析軟體工具,功能包括詞彙編輯(包括自建詞典、停 用詞、同類詞)、詞頻統計、詞彙之間的搭配關係、文本語料庫比較之正負顯著 詞彙、關鍵詞索引。「庫博」設計提供人文研究可以個別設定研究所需之斷詞、 類目建構、語助詞篩選等功能。庫博介面設計有利於沒有電腦程式語言能力之 人文社會研究者單獨操作,使研究者便捷發掘文本語料有趣的特徵,並反覆試 驗文本語料。目的在於提供人文學者檢視文本語料時與本身學科領域知識對話 的機會。本次演講將以紅樓夢為語料庫說明庫博的應用價值,包括兩個主題: 紅樓夢前後文本之差異與人物網絡關係。

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

玩庫博:人文社會學者的中文文本探勘工具

玩庫博:人文社會學者的中文文本探勘工具 PLAYING WITH CORPRO

A CHINESE TEXT-MINING TOOL FOR HUMANITIES

闕河嘉 HO-CHIA CHUEH

國立臺灣大學生物產業傳播暨發展學系 BICD, NATIONAL TAIWAN UNIV. 陳光華 KUANGHWA CHEN

國立臺灣大學圖書館 NATIONAL TAIWAN UNIVERSITY LIBRARY

國立臺灣大學圖書資訊學系 LIS, NATIONAL TAIWAN UNIVERSITY

林川傑 CHUAN-JIE LIN

國立臺灣海洋大學資訊工程學系 CSIE, NATIONAL TAIWAN OCEAN UNIV.



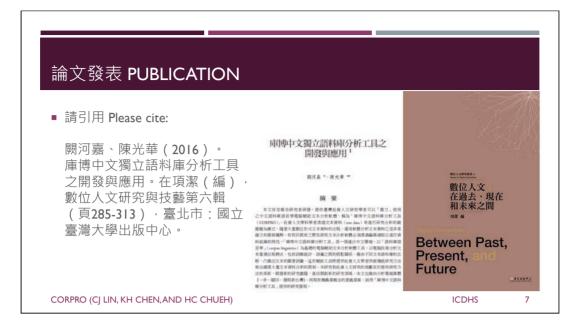
<section-header><list-item><list-item><list-item><list-item><list-item><list-item><list-item><list-item>







玩庫博:人文社會學者的中文文本探勘工具







CORPRO Alpha D:\CORPRO Workspace\test	、	乳始主畫 百	面 Main Inter	face
系統 設定	_			
標的語料庫 詞頻列表 關鍵詞檢索 搭配詞 關鍵詞分布 共現頻率 共現分析 参照語料庫 詞語顯著性 字串頻	率 詞語;	詞典 停用詞 [同類詞 文檔檢視 詮釋	資料欄位 更
匯人 編輯詮釋 重新斷詞 移	多除選擇			新增副除
# 詞 段落 檔案路徑		詞	來源	
就緒 移用	除所有文件	团人	匯出新增	匯出完整



中文可進行斷詞或事先斷詞 Word S	egmenta	ation Nov	v or Beforeha	nd
回日回日 回日 回日 <thli> <thli> <thl></thl></thli></thli>	<u>積率</u>]詞語:,, 移除選擇	調機停用詞	□頻詞 文懽檢視 詮釋寶 來源	料欄位 更.,, 新雄 朝除
就緒	修除所有文件	(正人)	匯出新增	匯出完整

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

玩庫博:人文社會學者的中文文本探勘工具

						CC	DRPRO Alpha D:\CORPRO Wor	斷詞完	戓 Word Se	gmentation	ОК
永統 副	安定										
標的話	香料庫	詞頻列表 關鍵語	同檢索 搭配詞	關鍵詞分?	15 共現	頻率	共現分析 参照語料庫 詞語顯著性	字串頻率 詞語;	調典 停用詞 同	類詞 文檔檢視 詮釋	資料欄位 更
Fe	i A	编辑詮釋	重新断詞					移除選擇			新增量
		PHINE IT THIN	386.071 881 675		823	CD.94	436 000 Ddr /277	15' 00' 20 14	-	where here	101 140
	題			0.0203 - 20 - 1	詞	段落	檔案路徑		詞	來源	
1 00		士隱夢幻識通靈		冒秀.docx	6072		1 F:\research\CORPRO\corpus\紅		5		
2 00		夫人仙逝揚州城		國府.docx	4502		1 F:\research\CORPRO\corpus\紅				
3 00		陵城起復貫用村		黛玉.docx	6272		1 F:\research\CORPRO\corpus\紅				
4 00		前又偏连得前即		蘆案.docx	4348		1 F:\research\CORPRO\corpus\紅				
6 00		生田罗洪社使罗		境情.docx	5639 5296		1 F:\research\CORPRO\corpus\紅	使罗王肃\社侵			
7 00		實工的訊券附值	制始始一進現	國府.docx 寶玉.docx	5290		1 F:\research\COF 1 F:\research\COF		×		
8 00		百亿川油峡火堤 賽銀小羊利乐院	新林市先建加 曹操工十款位	資玉.docx 芸軒 docx	4945		1 F:\research\CO				
9 00		員欽介志未首阮	却進好預會問	是常.docx	4120		1 F:\research\CO				
10 01		實總會利權受長	建大醫論環細	窮源.docx	3667		1 F:\research\COF i Wor	rd segment	ation OK!!		
11 01		素厚富府排安室	日配属曹珠把	注心.docx	4186		1 F:\research\CO	-			
12 01		配圖書設相思局	曹天祥正昭周	月鑒.docx	3148		1 F:\research\CO				
13 01		可即死封體禁尉	干配圖協理窗	國府.docx	3678		1 F:\research\COF		確定		
14 01		如海损館揚州城	寶寶玉路謁北	#∓.docx	4065		1 F:\research\COF		FIL AL		
15 01		鳳姐弄櫂鐵檻寺	秦鲸卿得趣鳗	頭鹰.docx	3693	i.	1 F:\research\COR	IN IN ALL IN ADA IN ALL			
16 01	16 賈	元春才選風蕩宮	秦鲸卿夭逝黃	泉路.docx	5432		1 F:\research\CORPRO\corpus\#1	樓夢全集\紅樓			
17 01	17 大	觀園試才題對額	賈寶玉機敏動	諸賓.docx	5941	-	1 F:\research\CORPRO\corpus\紅	樓夢全集\紅樓			
18 01	18 林	黛玉誤剪香囊袋	賈元春歸省慶	元宵.docx	6378		1 F:\research\CORPRO\corpus\#1	樓夢全集\紅樓			
19 01	19 情	切切良實花解語	意綿綿靜日玉	生香.docx	6899	6 1	1 F:\research\CORPRO\corpus\紅	樓夢全集\紅樓			
20 02		風正言彈妒意	林黛玉俏語謔嫌	音.docx	3976		1 F:\research\CORPRO\corpus\紅				
21 02		暴人嬌嗔箴寶玉	倘平兒軟語救	賈璉.docx	4535		1 F:\research\CORPRO\corpus\紅				
22 02		曲文寶玉悟揮機	制燈迷賣政悲	識語.docx	5204		1 F:\research\CORPRO\corpus\紅	樓夢全集\紅樓			
23 02		廂記妙詞通戲語	牡丹亭艶曲警	芳心.docx	3928		1 F:\research\CORPRO\corpus\紅				
24 02		787 MAT 4/2 1-2 494 15.2	痴女兒遺帕惹	相思.docx	6288		1 F:\research\CORPRO\corpus\紅				
25 02		THE / ALL MAY MAY ARE LALL / ME	紅樓夢通靈遇	雙真.docx	5812		1 F:\research\CORPRO\corpus\紅				
26 02		稜橋設言傳心事	蒲湘館春田發	器情.docx			1 F:\research\CORPRO\corpus\紅				
27 02		翠亭楊妃戲彩蝶	埋香家飛燕泣	殘紅.docx	4706		1 F:\research\CORPRO\corpus\紅				
28 02		土面情赠茜香羅	醉寶釵羞瞳紅	麝串.docx	7070		1 F:\research\CORPRO\corpus\紅				
29 02	29 享	福人福深還禱福	痴情女情重愈	目情.docx	6218		1 F:\research\CORPRO\corpus\#1	樓夢全集\紅樓	1		
成緒	00.00			and the				移除所有文件	匯入	匯出新增.	. 匯出完整

系統 設定							
標的語料庫	詞頻列表 關鍵語	同檢索 搭配詞 關鍵詞分	布 共現	頻率 共現分析 參照語料庫 詞語顯著性 字	串頻率 詞語:	詞典停用詞	同類詞 文檔檢視 詮釋資料欄位 更
团人	編輯詮釋	重新斷詞			移除選擇		
* 標題			詞	段落 檔案路徑			
1 001	版士隱夢幻識通靈	夏雨村風塵懷闇秀.docx	6072	1 F:\research\CORPRO\corpus\紅樓書	全集\紅樓 、	005	開生面夢演紅
	置夫人仙逝揚州城	冷子興演說榮國府.docx	4502	1 F:\research\CORPRO\corpus\紅樓書		000	
	金陵城起復賈雨村	榮國府收養林黛玉.docx		1 F:\research\CORPRO\corpus\紅樓書		+	
4 004 1	專命女偏逢薄命郎	葫蘆僧亂判葫蘆案.docx		1 F:\research\CORPRO\corpus\紅樓事	「全集\紅樓	樓夢	立新場情傳幻
5 005	制生自夢演紅樓臺	立新場情傳幻境情.doc		1 F:\research\CORPRO\corpus\紅樓書	全集\紅楼		
	實寶玉初試雲雨情 ※宮花周瑞運英蓮	劉姥姥一進榮國府.doc)		1 F:\research\CORPRO\corpus\紅樓書		上空ル主	daav
	古名化向场课央理	談肄業業鐘結寶玉.docx 賈寶玉大醉絳芸軒.docx		1 F:\research\CORPRO\corpus\紅樓團 1 F:\research\CORPRO\corpus\紅樓團		境情.	aocx
	律員奴小志未曾阮 善周流情友人実勢	夏賀玉八軒解云町.doc) 記嫌疑頑童開學堂.doc)		1 F:\research\CORPRO\corpus\紅樓事			
10 010 1	全實總會利權受受	張大醫論病細窮酒,doc)		1 F:\research\CORPRO\corpus\紅樓書		ALL 1-1 1 10 100	葳蕤擁 繡衾 · 恍 随 仙子 到 紅塵;
	要書辰寧府排家宴	見服風曹瑞把浮心.doc		1 F:\research\CORPRO\corpus\紅樓書			寄 景 · 千古 風流 造 業人 •
	王熙鳳舞設相思局	賣天祥正照風月鑒.doc		1 F:\research\CORPRO\corpus\紅樓書		第四回中既	將薛家母子在榮府內寄居等事略
13 013	秦可卿死封龍禁尉	王熙鳳協理寧國府.docx	3678	1 F:\research\CORPRO\corpus\紅樓書	「全集\紅樓	已表明・此	回則暫不能寫矣。
14 014 1	林如海揭館揚州城	賈寶玉路謁北靜王.docx	4065	1 F:\research\CORPRO\corpus\紅樓書	全集\紅樓	切合 日銀 林台	電子自在 榮 府 以來 · 賈母 萬般 憐愛
15 015	王鳳姐弄櫂鐵檻寺	素鯨卿得趣饅頭庵.docx		1 F:\research\CORPRO\corpus\紅樓書		, 宿食 把层	· 一如寶玉 · 迎春 · 探春 · 惜春 =
	買元春才選風蕩宮	秦鲸卿夭逝黃泉路.docx		1 F:\research\CORPRO\corpus\紅樓書		THE PARTICULAR	
17 017 7	大觀園試才題對額	賈寶玉機敏動諸賓.docx		1 F:\research\CORPRO\corpus\紅樓書		個親孫女倒	
18 018	林県土誤對音義 殺	賣元春歸省慶元宵.docx		1 F:\research\CORPRO\corpus\紅樓書	全類\紅樓		 亦自較別個不同 日則同行同
19 019 f	青切切良宵花解語	意綿綿靜日玉生香.docx 林黛玉俏語誠嬌音.docx	6899	1 F:\research\CORPRO\corpus\紅樓書 1 F:\research\CORPRO\corpus\紅樓書		坐・夜則同	息同止,真是言和意顺,略無參問
	照周止百年3月8日	修圖工词語語為自.docx 倘平兒軟語救費鏈.docx		1 F:\research\CORPRO\corpus\紅樓畫			
	國家八炳明成員工	制增排曹政悲讀語.doc		1 F:\research\CORPRO\corpus\紅樓書		不想 如今 忽然	* 來了一個 薛寶叙 ・ 年歳 雖 大 不多
	西廂記妙詞通戲語	社丹亭靜曲警芳心.doc		1 F:\research\CORPRO\corpus\紅樓書			方,容貌豐美,人多調黛玉所不
	粹金剛輕財尚義俠	痴女兒遺帕惹相思.doc		1 F:\research\CORPRO\corpus\紅樓書		· 流 血信 确, 及 ·	7 · 台帖皇天 · 八岁间黑玉所不
25 025	屬魔法姊弟逢五鬼	紅樓夢通靈遇雙真.doc)		1 F:\research\CORPRO\corpus\紅樓書			
	峰腰橋設言傳心事	潇湘館春图發器情.docx		1 F:\research\CORPRO\corpus\紅樓書		而且寶釵行為	A needed to be a set of the set o
	商翠亭楊妃戲彩蝶	埋香家飛燕泣殘紅.docx		1 F:\research\CORPRO\corpus\紅樓書		自許・目無	「塵・故比黛玉大得下人之心。
	傳玉菡情贈茜香羅	薛寶釵羞龍紅麝串.docx		1 F:\research\CORPRO\corpus\紅樓書		便是那些小	(頭子們,亦多喜與寶釵去頑。
29 029	享福人福深還禱福	痴情女情重愈斟情.docx	6218	1 F:\research\CORPRO\corpus\紅樓書	全無\紅樓		▶ 便 有些 悒鬱不忿 之 意 · 寶釵 卻 渾

						C	ORPRO Alp	ha D:\	詞步	夏與文	⊄件頻ጃ	∝ Word F	Frequency and DF
系統	設定							AutoCorrection					. ,
INTO	ER WILLIAM	104百万()王 8		「おお田」	開設部分本社	田橋東	土田分析	·林田双主新来注意	1 1015 25	装柱 空泉	46 SE 1015	12 曲 (体田 22	同類詞 文檔檢視 詮釋資料欄位 更
1980H 2			NO SHE IN'S LLL .	N IN BURS	1010 DE 010 73 10 21		II CONTRACTOR	A CONTRACTOR	10000000000	State Coloradore	and the second	1 196346 13.113 840	
詞		• 精密比對				搜尋	回到	完整列表	□歐用	□ 啟用	□忽略		授辱 清除 < > -/
# 1	同	35	频	文件頻率								005	
1			41289	80							1	005	開生面夢演紅
2			19755	80								000	
3	7		12630	80								Jack Hall	
4 1	10		8423	80								樓夢	立新場情傳幻
5			8164	80								135 2	エ州の月寺ち
6	1		7977	80								1	
7	r		7763	80								境情.	docy
8	我		5346	80								5元 月・	UUCA
9	道		5312	80								題曰:春团	葳蕤 擁 繡衾 · 恍 随 仙子 到 紅塵;
10	他		4413	80									
11 1	你		4173	80									
12	2		4022	80								第四回中既	將薛家母子在榮府內寄居等事略
13 3	來		3757	80								已表明・此	回則暫不能寫矣。
14	t.		3546	80									電子 自在 榮 府 以來 · 賈母 萬般 憐愛
15	V		3485	80									
16	去		2994	80									 一如寶玉 · 迎春 · 探春 · 惜春 三
17	宫		2887	80								個親孫女倒	且靠後・便是寶玉和黛玉二人之
18	英		2808	80								親宠方愛成	·亦自較別個不同 · 日則同行同
19	著		2739	80									息同止,真是言和意顺,略無參商
20	-		2696	80								王·拉则问	心内止, 具定百加应限, 始無参問
21]	寶玉		2591	77								1	
22 -	20.02		2376	80								不想 如今 忽然	* 來了一個 薛寶釵 · 年歲 雖 大 不多
23	?		2128	80								, 然品格 [[]]	方·容貌 豐美·人多調 黛玉所不
24 1	(E)		2080	79								及。	1 11 11 11 11 11 11 11 11 11 11 11 11 1
25	有		1951	80									
25	Å		1918	80								而且寶釵行為	₿ 豁達 ・ 随分 從 時 ・ 不比 黛玉 孤高
27 1	困		1848	80								自許 · 目無下	「應 · 故比 黛玉大 得下人 之 心 ·
28			1740	80								便是那些小丫	
29			1729	80									
匯出	-										1	2 因此 黛玉 心 9 然不覺 ・	▶便有些悒鬱不忿之意 · 寶釵卻渾

系統 設定														
標的語料」	■ 詞頻列表 關鍵詞檢算	索 搭配詞 !	關鍵詞分布 共現	現率	共現分析 参照語	料庫 詞	间語願著性	字串頻率	詞語;	詞典 停用	調 同類調 文	て檔檢視 詮釋資料	科欄位	更.
詞	• 精密比對 •			搜尋	回到完整列制	ŧ		▶ 啟用 □	忽略				新增	刪除
# 詞	詞頻	文件頻率								啟用	停用詞	來源		
1 .	41289	80							9	30	۰.	本次		
	19755								U.	8		本次		
	12630									8	7	本次	新增	
4 的	8423									8	的	本次	新増	
	8164	80 80								2		本次		
										8	:	本次	新增	
	7763	80								8	. Г.	本次	新增	
8 我	5346									8	我他你來也又去這著是	本次	新增	
9 通	5312									8	他	本次	新增	
10他 11你	4413									8	你	本次		
11 你	4173	80							100 C	8	來	本次		
12 誤	4022	80								8	也	本次	新增	
13 來	3757									8	又	本次	新增	
14 也	3546									8	去	本次		
	3485	80								8	這	本次		
16 去	2994									8	著	本次		
13 来 14 也 15 又 16 去 17 這										8	是	本次		
18 🕮	2808									8	-	本次		
19 著	2739									8	?	本次	新增	
19 著 20 是	2696	80								200				
21 寶玉	2591	77												
22 — 23 ?	2376	80 80												
23 ?	2128													
24 便	2080													
25 有	1951	80												
26 人	1918													
27 個	1848													
28 就	1740	80												
29 都	1729													
										-				
匯出										团人		匯出新增	匯出!	兒整

系統 設定				CORPRO	Alpha D:\CORP	RO Workspa	ce\red3\			· 拾	·��記(Collocati	on	
	1045 7018	關鍵的檢索 搭配調		旧稿家 计用公		1015 55 35 44	中田福安	ents.	(mm	(A FE E	(11) 新田 (文)	當檢視 詮釋資料	il INNA (h) सा
and the second second	금역 가면 가지 것이 같이	Contraction of the local division of the loc	and the second se	现现年 大現力		States -			E0194	行用的	回照的 又1			1
夏轉:		執行 2	文章		2 副用	停用詞	□忽略變7	F.					新增	刪除
#	Freq.	• N-gram							啟用		停用詞	來源		
	1	2105 笑道 521 兩個						4		2		本次		
	2	521 兩個						-		8		本次		
	3	445 鳳姐 兒						-		85	7	本次	新増	
	4	296 二人								8	的	本次	新增	
	5 6	285 湘 雲								80	3	本次	新墙	
	6	278 寶玉 道								8	:	本次		
	7	231 寶玉 笑								8	r	本次		
	8 9	152 一面 說								8	我他你來也又去這著是	本次	新增	
	9	151 在 這里								8	他	本次	新增	
1	0	149 都 笑 143 在 那裡								8	你	本次	新增	
1	1	143 在那裡								8	來	本次	新增	
1	2	141 寶玉 聰								8	也	本次		
1	.3	123 因問								100	又	本次	新增	
1	4	120 只說								8	去	本次	新増	
	4	119 作什麼									擅	本次		
	.6	117 有個								18	蕃	本次	新坡	
	.7	114 賈母笑						_		8	是	本次		
	8	111 有什麼								R.	-	本次		
		108 一切 切								8	2	本次		
	0	106 雨村												
	1	106 雨村 104 句話												
	9 0 1 2 3	103 只聽												
	3	101 便說												
	4	100 笑說												
	5	99 看時						_						
	6	98 眾人都												
-	7	98 冷笑 道												
	7	97 呢!												
	9	95 18 1												
		95 說 畢						R	-					
阿出									F	7		薩出新增	匯出	宗教.

Preq. N-gram 飲用 使用調 來源 1 2105 笑道 第 本次新聞 第 本次新聞 3 445 罵娘兒 第 本次新聞 第 ● ・本次新聞 4 296 二人 5 285 浩雲 ● ● ・本次新聞 5 285 浩雲 ● ● ・本次新聞 ● ● ・本次新聞 6 278 豐王達 ● ● ・本次新聞 ● ● ・本次新聞 9 •52 - 回日 ● ● ● ● ● ・本次新聞 10 151 在 道里 ● ● ● ● ● ● ● ● ● → → ● ● → → ● <th></th> <th></th> <th></th> <th></th> <th>0</th> <th>OPPPO Alaba DàCOP</th> <th>PPO Works</th> <th>取消</th> <th>停</th> <th>用詞 D</th> <th>isable S</th> <th>top Word</th> <th>s</th>					0	OPPPO Alaba DàCOP	PPO Works	取消	停	用詞 D	isable S	top Word	s
檀的語科庫 調練列表 解鍵詞合素 指趾調 解鍵詞合素 指趾調 解膜 調子 樂明 副規 帶用調 回酬 文檔佛 经 第 1 2015 笑道 文庫 《 松田停用調 ② 総健形 新油 第 1 2105 笑道 文庫 《 松田停用調 ② 総健形 第 6 - 本次新増 2 521 雨信 3 445 風伯兒 - 本次新増 8 - 本次新増 4 296 二人 28 7 - 328 万 8 - 本次新増 5 285 酒雪 1 195 世名	★ 体 助 中				ç	ORFRO Alpha D.(COR	PRO WORKS		1.7	/ 13 8 3 6			
(現象: 較子: 数行 2 文章 ◆ Wr-ray 2 2105 笑道 2 2105 笑道 3 445 馬娘兒 3 106 世者 3 107 119 作什麼 3 101 151 在選里 10 151 在選里 11 149 都笑 12 103 云的奶 13 114 寶田笑 20 111 有什麼 21 108 云的奶 21 108 云的奶 21 108 云的奶 23 105 世就 23 105 世就 23 105 世就 23 95 男孩 23 95 男人範 4 10 10 10 10 10 10 10 10 10 10 10 10 10	MOST REAL AND			-					_				
Preq. N-gram 政用 使用調 來源 1 2105 笑 通 第 本次新聞 第 · 本次新聞 2 521 兩個 第 · 本次新聞 第 · 本次新聞 4 296 二人 · <th>標的語料庫</th> <th>1 詞頻列表</th> <th>關鍵詞檢索</th> <th>搭配詞</th> <th>關鍵詞分布 共現頻率</th> <th>共現分析 鬱照語料圖</th> <th>1 詞語願著</th> <th>性 字串頻率 詞語</th> <th>8</th> <th>詞典 停用詞</th> <th>同類詞 文權</th> <th>拾視 詮釋資料</th> <th>開位 更.</th>	標的語料庫	1 詞頻列表	關鍵詞檢索	搭配詞	關鍵詞分布 共現頻率	共現分析 鬱照語料圖	1 詞語願著	性 字串頻率 詞語	8	詞典 停用詞	同類詞 文權	拾視 詮釋資料	開位 更.
1 2105 笑道 2 521 開催 3 44 296 二人 3 44 296 二人 3 238 異常 4 236 二人 3 238 異常 3 238 第二 3 9 52 20 3 238 第二 3 9 52 30 3 238 第二 3 100 本次新増 8 10 本次新増 8 11 149 都笑 12 136 也有 13 141 實工體 14 第二 4 26 二 14 14 15 123 四 15 123 四 15 13 141 實工體 10 151 在 15 123 四 15 123 四 15 123 四 15 13 141 實工體 10 本次新増 13 141 實工體 16 17 7. 101 月 行作 度 17 101 斤行 度 10 本次新増 13 141 實可笑 14 17 101 斤行 度 10 本次新増 14 104 可訪 15 130 四 114 責可受 14 14 14 14 15 123 四 15 130 奶 奶 14 14 14 14 14 15 133 141 責可受 14 14 14 14 14 14 14 14 104 可訪 25 103 只聽 14 14 14 14 14 14 15 133 141 責可受 14 14 14 14 14 14 14 15 103 只聽 14 14 14 <td< th=""><th>搜尋:</th><th></th><th>執行</th><th>2</th><th>文章</th><th>• 2 啟月</th><th>月停用詞</th><th>□忽略變形</th><th></th><th></th><th></th><th>新</th><th>增删除</th></td<>	搜尋:		執行	2	文章	• 2 啟月	月停用詞	□忽略變形				新	增删除
4 296 二人 5 228 算玉道 6 278 算玉道 7 231 寶玉笑 10 151 在道里 10 151 在道里 11 149 都笑 12 143 在那裡 13 141 寶玉陽 14 126 也不 15 123 因因問 16 120 只說 17 191 作作愛 18 117 有個 19 14 寶母笑 19 14 寶母笑 10 13 6 出資 12 13 6 出資 13 141 寶子 14 126 也不 15 123 因問問 16 120 只說 17 191 作作像 19 114 寶母樂 19 14 冒母樂 11 14 明母樂 12 108 口所 13 11 寶母樂 14 127 有個 15 123 四節 尚 16 120 只能 17 191 作 19 14 雪母樂 113 117 有個	#	Freq.	• N-g	ram						啟用	停用詞	來源	
27 100 天武 28 99 着時 29 98 眾人都		4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22	296 二 湘 章 296 二 湘 章 2985 第 9 6 二 湘 章 2985 第 9 章 278 章 278 \overline{278 } 2	《聖····································	>				.)		」: 我他你來也又去這著屋一	本 宋文 宋 永 宋 永 宋 永 宋 永 宋 永 宋 永 宋 永 宋 永 宋 永	鵋辧 巍巍巍巍巍巍巍巍巍巍巍巍巍
匯出 匯入 匯出新增 匯出完計	匯出	27 28 29	101 使訪 100 笑訪 99 看訳 98 眾人						-	FE 1		匯出新增 团	[出完整

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

玩庫博:人文社會學者的中文文本探勘工具

4. LA 20.00					ORPRO Alpha D:\COR	ine nonspe				op Words
系統 設定								1		
標的語料圖	1 詞頻列表	關鍵詞檢索	搭配詞	關鍵詞分布 共現頻率	共現分析 參照語料庫	1 詞語顯著性	字串頻率 詞語;	詞典 停用	詞 同類詞 文檔核	說視 詮釋資料欄位 更
叟辱:		執行	2	文章	• 2 胶月	目停用詞	□忽略變形			来斤均的一册
#	Freq.	• N-g						啟用	停用詞	來源
	1	2105 笑 連 521 兩 個	1				-	8		本次新增
	2	521 兩個	8					8		本次新增
	3	445 鳳姐	兒					8	7	本次新增
	4	296 二 人 285 湘 雪 278 寶玉 231 寶玉	2					S	的	本次新增
	5	285 湘 劉	ř.					8	4	本次新增 本次新增
	5 6 7	278 寶玉	道					8	:	本次新增
	7	231 寶玉	笑					8		本次新增
	8	196 世 祖						8	我	本次新增
	8 9 10 11 12	152 一面 151 在道 149 都 第	說					8	他你來也又去	本次新增
	10	151 在 道	里					8	你	本次新增
	11	149 都 第	E					8	來	本次新增
	12	1/12 77 10	(38)					0	也	本次新增
	13 14 15 16	141 寶玉 126 也不 123 因問 120 只訪	聽					8	又	本次新增
	14	126 也不	5					8	去	本次新增
	15	123 因問	1					8	這	本次新增
	16	120 只該	2						1	4 大部 40
	17	119 fF f1	- 155						易	本次新增
	18	117 有個 114 賈母					C C			大次新神
	19	114 曹母	笙					9	2	144 X 81 All
	20	111 右 付	- 101							
	21	108 _ 10	奶							
	18 19 20 21 22	106 雨木	1							
	23	108 二奶 106 雨村 105 包載 104 句慧 103 只能	t							
	24	104 句 記	Ē							
	25	103 🖓 🛤								
	26	101 便訪	1							
	23 24 25 26 27	101 便該 100 笑該	ž							
	28	99 看 眼	ŧ.							
	29	98 眾人								

標的語》	4庫 詞頻列表	關鍵詞檢索 搭配	調 關鍵詞分布 共現頻率 共	共現分析 參照語料庫 詞語顯著	性 字串頻率 詞語;	調典 停用詞 同	同類詞 文檔檢視 詮釋資料欄位 更
叟辱:		執行 2	文章	•3 啟用停用詞	□忽略變形		新加
ŧ.	Freq.	• N-gram				詞	來源
	1	2105 笑道			4	鳳姐兒	本次新增
	2	521 兩個			2	湘雲	本次新增
		445 風姐 見				二奶奶	本次新增
	4	296 二人				雨村	本次新增
	5	285 湘雲				Second Second	
	6	278 寶玉 道					
	7	231 寶玉 笑					
	8 9	196 也有					
	9	152 一面 說					
	10	151 在 這里					
	11	149 都 笑					
	10 11 12 13 14 15 16	151 在 這里 149 都 笑 143 在 那裡					
	13	141 寶玉 聽					
	14	126 也不					
	15	123 因問					
	16	120 只說					
	17 18	119 作 什麼					
	18	117 有個					
	19	114 賈母 笑					
	20	111 有什麼					
	21 22	108 二奶 奶					
	22	106 雨村					
	23 24	105 也就 104 句話					
	24	104 句話					
	25	103 只聽					
	25 26 27	101 便說					
	27	100 笑 說					
	28 29	99 看時 98 眾人都					

價的語料庫 詞頻列表 關鍵	調検索 探察部 開鍵調分	5 井頂橋室	共現分析 参照語料庫 詞語顯著性 字	电積寒 詞請:	調典 停用詞 同類	記 文標檢過 於護	皆料欄位 軍
應入 編輯詮釋. 標題	重新斷詞	詞段落		移除選擇	a []	來源	新增一副
1001	具内+1/2を使用したの 一次のの1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2を使用したの 一次の1/2 での したの1/2	6072 4502 6272 4348 55639 5296 5378 4945 4120 3667 4120 3148 3678 3428 3678 3693 3148 3678 3693 3693 3693 3941 6378 6899 3976 6378 4535 5204 4535 5202 4535 5202 6288 2528 5812	1 F:\research\CORPRO\corpus\紅標 1 E:\research\CORPRO\corpus\新闻 2 進行斷詞	会報用指示式建築。	■ 相兒 湘雲 二 奶奶 雨村	本次新增 本次新增 本次新增 文列 (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)	

International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班: 數位人文的發展與應用

	_	יני	把戈訶集	「水建山	同類詞集合	Drag	o Gro	oup v	ords/	wit	h Simila	ir Subj	ects
系統 設定									_				
標的語料庫	詞頻列表 關鍵詞	檢索	搭配詞 關鍵詞分	分布 共現頻率	共現分析 鬱照語料圖	詞語願著性	字串頻率	詞語;	詞典例	用詞	同類調文檔	檢視 詮釋算	戰料欄位 更
詞	・精密比對・			搜尋	回到完整列表	□ 啟用	□ 啟用 □	忽略	釈行	増	編章	R	刪除
# 詞	詞頻		文件頻率						啟用	#	Category	Words	Source
43 等		1120	80					-	- V	1	主角	[寶玉, 賈	本次新聞
44 上 45 質母		1119 1096	80 66					_	10000			President Permit	本次新增
46 呢		1063	80						\checkmark	1 (1)	寶玉		
47 04		1057	78 78						\checkmark	1 (2)	賣母		本次新增
48 什麼		1019	78						1	1 (3)	鳳姐		本次新增
49 一個		972 962	80 80					_	1000				
50 還 51 問		952	80						V	1 (4)	製人		本次新增
52 吃		906	71						\checkmark	1 (5)	重重		本次新增
53 看		840	80 80						1	1 (6)	寶釵		本次新增
54 不		820	80						100				
55 們		817	78 60					_	V	1 (7)	老太太		本次新增
56 風烟 57 興		778	80						V	1 (8)	鳳姐兒		本次新增
58 4		769	76						V	1 (9)	湘雲		本次新增
59 才		767	78 78 76 75						N.	1 (5)	/10 998		44-90/17/1 148
60 _		762	78										
61 F		750	76										
62 事		748	75										
63 和		737	77 57 79					_					
64 根人 65 我們		733	5/										
66 倒		726	80					_					
67 到		717	00										
68 你們		676	80 75										
69 一面		674	78										
70 館玉		672	60										
71 如今		671	80						0				
爾出			=-						Ē		(EE)		

			COF	RPRO AI 同类	傾詞合併統計詞]頻 Fre	equer	ncy of C	Catego	ries
系統 設定										
標的語料庫	詞頻列表 關鍵詞	檢索 搭配詞	關鍵詞分布 共現頻率	共現分析 參照語料庫	詞語顯著性 字母碩率 詞語	§ 詞典	停用詞	同類詞 文檔	檢視 詮釋算	和欄位 更
詞	・精密比對・		搜尋	回到完整列表	□啟用 ◎啟用 □忽闘	ł	新壇	編章	R	刪除
# 詞	詞頻		同類詞			啟用	#	Category	Words	Source
7 1	77 77			000 001 8 (01770 001)		5+1	1	主角	[寶玉, 賈	本次新增
8·主角 9 我	53	46 80	質工[2591,//],頁母[1	096,66],周32[778,60],	農人[733,57],黛玉[672,60],					
9 我 10 递 11 他 12 你 記 14 來	53	12 80				• <	2	丫環	[靖雯, 麝	本-汉新增
11 他	44					► V	3	主子	[迎春, 探	本次新增
11 他 12 你 13 說	41)			1000			• (The section of the	
13 説	40	22 80)							
14 來	37									
15 也 16 又	35	46 80)							
15 也 16 又 17 去 18 這 19 笑 20 著 21 是	34	85 80								
17 去	29)							
18 這	28									
18 這 19 笑 20 著 21 是 22 · 主子 23 賈璉 24 探春 25 賈政	28	08 80)							
20 著	27									
21 是	26	96 80)							
22 * 主子	24	33 76	育理[422,41],探春[32	5,40],賈政[323,32],賈3	[321,35],尤氏[309,32],李納					
23 贾 璉 24 探春	4	22 41	Providence of the state of the second s			22				
24 探春		25 40)							
25 賈政	3	23 32 21 35	2							
26 賈珍		21 35								
27 尤氏 28 李紈		09 32	2							
28 李紈		51 40)							
29 曹装		00 25 25 38 79 29 78 30								
30 迎春 31 主子 32 惜春		25 38								
31 主子		79 29	1							
32 惜春		78 30)							
33 -	23	76 80								
34 ? 35 便	21									
35 便	20)			7 4				
國出		~ ~~					匯入	(E)	4	

系統 設定		trace of Law					u latanaan	Loon						
And the second s		同機案	能詞開鍵詞分	布 开現頻率 手	現分析 鬱照語料』	■ 詞語願者	性 字串頻率	詞語;	, 詞典 停用詞	同類詞 又	當機視 目	E種資料	欄位	更.,
詞	子字串比對	玉		搜尋	回到完整列表	回啟用	□ 啟用 □	忽略		授辱	清除	<	>	-/-
# 詞	精密比對	文	件頻率											
1 2 課紙 五 工業 玉 工業 工工業 工工業 工工業 工工業 工工業 工工業 工工業	(括)略 玉	2591 672 268 172 48 15 11 8 5 5 5 4 4 4 4 4 3 3 3 2 2 2 2 2 2 2 2 2 2 1 1 1	77 60 39 51 7 6 9 2 3 3 4 4 3 3 4 4 3 3 3 1 2 2 2 2 2 1 1 1 1											

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

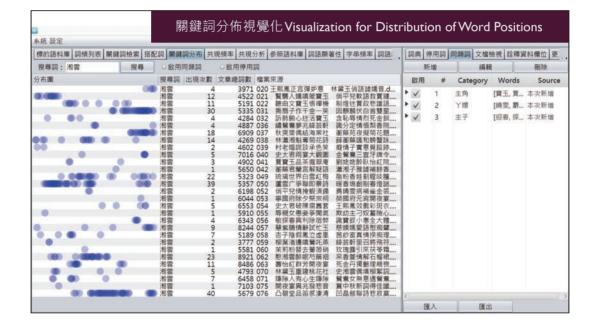
玩庫博:人文社會學者的中文文本探勘工具

COLOR.	設定												L					
標的	語料庫 詞頻列	康	關鍵詞檢索	搭配詞	關鍵詞分	布 共現	見頻率	共現分析 鬱	照語料庫	詞語顯著性	字串頻率	詞語;	詞典	停用	詞	可願詞 文檔相	会視 詮釋資	(料欄位) 更
搜寻	詞:寶玉		執行 左數:	2	2 右數:	5	文章			• □ 同類詞	●停用詞 □	忽略	-	新增		編朝	1	刪除
	左方出現詞	嗣	右方出現語	1		來源	文章						啟	用	#	Category	Words	Sourc
133	聴見・	ŋ.	在車上見	這般醉		007		花周瑞嘆英臻	談肄業!	柴鐘結寶玉.c		ń	F V	1		主角	[寶玉, 賈	太灾新博
134	號的	m3	七央告道	: 1		007	送宫石	花周瑞嘆英蓮	該肄業!	秦鐘結寶玉.c								
135	鳳姐和			過眾人		008		設小恙梨香院		大醉絳芸軒.c			P 1	2	2	丫環	[靖雯, 麝	本次新增
136	思人・	- M -	先便回明			008	薛寶			大醉絳芸軒.c			P V	3		主子	[迎春, 探	太次新博
	王夫人 林黛玉		等過去看	戲·至		008	醇寶			大醉絳芸軒.c			1.16	- 37		-T 1	form may reason	AP 27 411 28
138	· 卻說	1 1 1 1		1 米回		008	醉寶			大醉絳芸軒.c								
139	見了	<u>ĝ</u>]		省		008		設小恙梨香院		大醉絳芸軒.c		_						
140	說的	10	也笑了。	於是	-	008 008	醉寶	取小志采香院		大醉絳芸軒.c		_						
141	多日未見		 ・ 超米 都 ・ 忙 上來 	111 日	+	008	評質			大醉絳芸軒.c 大醉絳芸軒.c		_						
143	3日 不元 請安 ·	341.3	杜 令笑 擁	他起來		008	萨爾			大醉維芸軒.0		_						
144	14.	1	2 位 日 人 周	在		008		設小恙梨香院		大醉絳芸軒.0								
145		20-1	等道:「	芣值		008	前寶			大醉綠芸軒。								
146	· 且說	1	來至梨香	院中,		008	薛寶	的小羔梨香院		大醉綠芸軒。								
147	46 ・	11	忙請了安			008	范留:	線小羔梨香院		大醉絳芸軒.0								
148	来。	1	因問:「	哥哥不	在家	008	前寶	四小羔梨香院		大醉絳芬軒.0								
149	* 4	111	猫:「娃	姐可		008	薩寶	設小恙梨香院	晋寶玉:	大醉維芸軒.0	locx							
150	1.1	寶日	、聽說・忙	下了		008	薛寶	設小恙梨香院	賈寶玉:	大醉絳芸軒.c	locx							
151	朣 •	寶日	新廠一邊	步 進去		008	薛寶	設小恙梨香院	質寶玉:	大醉絳荟軒.c	locx							
152	守拙・	寶日	5 一面看 ·	一面問		008	薛寶	設小恙梨香院		大醉絳芸軒.c								
153	抬頭 只見	寶日	· 進來 · 連	忙起身	含笑	008		設小恙梨香院		大醉絳芸軒.c								
154	一面看	ER	頭上戴著	景		008	薛寶			大醉絳芸軒.c								
155	下來的	寶3				008		釵小恙梨香院	費寶玉:	大醉絳芸軒.c								
156	來。	M	「亦凑了上	去・		008	醉寶	取小志梨香院	買貸玉	大醉絳芸軒.c								
157	• 1	<u> </u>	「聴了・竹	笑		008		設小恙梨香院		大醉絳芸軒.c		_						
158	* J	10 H		好		008	薛寶			大醉絳芸軒.c								
159		寶.				800		設小恙梨香院		大醉絳芸軒.c		_						
160 161	永淵		看了,也			008		設小恙梨香院 設小恙梨香院		大醉絳芸軒.c 大醉絳芸軒.c								

系統 設定							Lc T				
標的語料庫	詞頻列表 關鍵	詞檢索搭	記詞 關鍵詞分布 共現頻率 共現分析	析 參照語料庫 詞	語顯著性 字串頻3	目詞語;	詞典(亨用詞	同類詞 文檔	会視 詮釋算	和欄位 更
搜尋詞: =	E角 執行	左數:	2 右數: 5 文章	• 30	同類詞 □停用詞	□忽略	Ŵ	i増	編朝		刪除
# 左方と	出現詞	關鍵詞	右方出現詞	來源	文章		啟用	#	Category	Words	Source
2583	1.1	寶玉[主角		080	懦弱迎春腸回九		FV.	1	主角	(寶玉,賈	本次新聞
2584	2.3	寶玉[主角		080	儒弱迎春陽回九			•			
2585 2586	1.1	寶玉[主角		080	情報迎春陽回九I		► ✓	2	丫環	[請愛, 麝	本次新增
2586	• 1	寶玉(主)		080	儒弱迎春腸回九 像弱迎春腸回九		F V	3	主子	[迎春, 探	太次新博
2588	著	寶玉住角		080	儒羽迎春陽回九 像弼迎春腸回九		1.00	- -		Free may reason	New Yorkin Mill
2589		寶玉[主]		080	儒弱迎春腸回九						
2590	又吩咐	寶玉江主角		080	儒弱迎春腸回九						
2591	~ ~ ~ ~	寶玉江主角		080	像弱迎春腸回九						
2592	• 當下	費用[主角	All the set of a set	003	金陵城記復寶面;						
2593	過・	習母(主角	1 又說:「講	003	金陵城記復曹雨	村 榮					
2594	· 不免	晋母门主角		003	余陵城起復曹雨	村 榮					
2595	* 1	晋母[主]		003	余陵城記復寶雨	村 榮					
2596	接見 ·	費母(主角		003	金陵城把復曹雨	村 榮					
2597	送至	費母[主角	1 身邊坐下 · 因笑	003	金陵城起復寶雨	村 榮					
2598	拭淚。	晋母[主]	前 笑道:「我	003	金陵城起復寶雨	村 榮					
2599	揻 .	費母[主角	1 命 兩 個 老娘娘 帶	003	金陵城起復寶南	村 榮					
2600	* _	費母[主角		003	金陵城起復賈雨	村 榮					
2601	・比	費母[主角		003	金陵城起復賈南	村 榮					
2602	· 便是	費母[主角		003	金陵城起復寶雨	村 榮					
2603	() ()	費母[主角		003	金陵城起復費雨	村 榮					
2604	推讀。	費母[主角		003	金陵城起復賈南	村 榮					
2605	7 .	贾母[主角		003	金陵城起復費雨	村 榮					
2606	茶・	置母[主角		003	金陵城起復實雨	村 榮					
2607	<u>7</u> ·	置母[主角		003	金陵城起復費南	村 榮					
2608	THE REAL	費母[主角		003	金陵城起復費雨	村 榮					
2609	寶玉[主角] 向	費母[主角		003	金陵城起復費南	村 榮					
2610 2611	安 · 形狀 !	費母[主角		003	金陵城起復寶南 金陵城起復寶雨	村 榮					

與統 設							1.00				
標的語	料庫 詞頻列表	關鍵詞檢索 搭配詞 關鍵詞	分布 共現	見頻率 共現分析 參則	服語料庫 詞語顯著性 字串頻率	詞語;	詞典例	亨用詞	同類詞 文檔相	会視 詮釋資	科欄位 更
搜尋訪	月:寶釵黛玉	執行 左數: 2 右數:	5	文章	•□同類詞 □停用詞 □	忽略	新	增	編輯		刪除
1 2	上方出現詞 關.	. 右方出現詞	來源	文章			啟用	#	Category	Words	Source
638	請了寶	观 來 分 晰 前日 的	078	老學士閒征姽嫿詞	痴公子杜撰芙蓉誄.docx	ŕ	P V	1	主角	[寶玉, 賈	* * **
639	居住 ・ 寶	双 陪笑 道 : 「 我	078	老學士聞征姽嫿詞	痴公子杜撰芙蓉誄.docx			1	土円	[與五, 頁	平火机墙
640	* 」 資料		078	老學士閒征她嫿詞	痴公子杜撰英蓉誄.docx		P V	2	丫環	[請愛, 麝	本次新增
641	去了寶	以 等 一 處 · 迎 春 雖	078	老學士閒征姽嫿詞	痴公子杜撰芙蓉誄.docx		N 1002		+		
642	薛寶釵・寶	図 久 察 具 不軌之心 ·	079	薛文龍悔娶河東獅	賈迎春誤嫁中山狼.docx		► ✓	3	主子	[迎春, 探	本灾新增
643	字,寶	观亦不在意·只因	080	儒弱迎春腸回九曲	姣怯香菱病人膏肓.docx						
644	香菱・寶	X 笑道:「咱們	080	懦弱迎春腸回九曲	姣怯香菱病人膏肓.docx						
645	•」寶	20 笑道:「他	080	懦弱迎春腸回九曲	姣怯香菱病人膏肓.docx						
646	果跟随寶	図 去了 · 把 前面	080	懦弱迎春腸回九曲	姣怯香菱病入膏肓.docx						
647	· 乳名 黛	Ⅰ · 年方五歳	002	賈夫人仙逝揚州城	冷子興演說榮國府.docx						
648	那女學生 黛	E · 身體方愈 ·	003	金陵城起復費雨村	榮國府收養林黛玉.docx						
649	? J 🕅	E 聽了 ・ 方 瀨浜	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
650	· 依附 黛	⊾ 而 行 ・ 有日 到	003	金陵城起復賈南村	榮國府收養林黛玉.docx						
651	· 且說 黛	E 自 那 日 棄 舟	003	金陵城起復費雨村	榮國府收養林黛玉.docx						
652	大字 • 黛	E想道:這必	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
653	・扶黛	E 下 轎・ 林黛玉 扶	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
654	*」篇	E 方 進入 房 時 ·	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
655	上來 · 黛	E便知是他外祖母	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
656	涕泣 · 黛	E 也哭個不住。	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
657	了,黛	五方拜見了外祖母・	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
658	指與黛	E:「這是你	003	金陵城起復賈南村	榮國府收養林黛玉.docx						
659	•」篇	拜見過・費母	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
660	妝飾・ 黛	E 忙 起身 迎 上來 見禮	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
661	說些 黛	E 之母如何得病,	003	金陵城起復寶雨村	榮國府收養林黛玉.docx						
662	摟了 黛	E.在懷,又鳴咽	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
663	眾人見 黛	E 年貌 雖小 · 其	003	金陵城起復費雨村	榮國府收養林黛玉.docx						
664	?」 🏛	E 道:「 我 自來	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
665	!」黛	E 納罕 道 : 「 這些	003	金陵城起復賈雨村	榮國府收養林黛玉.docx						
666	開・黛	E 連忙 起身 接見 • 費母	003	金陵城起復賈雨村	榮國府收養林黛玉.docx		2				

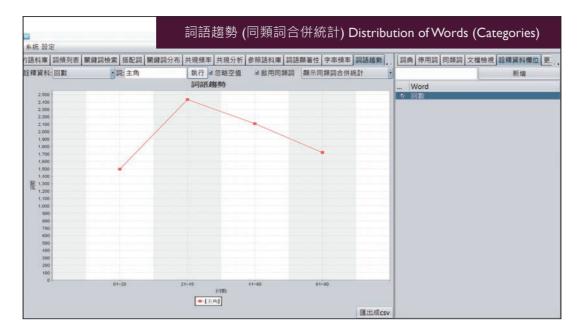
■ 系統設定		關	鍵詞檢索((連續查詢	同詞) K	eyword	d in C	Conte	ext (Adjacer	nt Wor	ds)
標的語料庫 詞頻列表	開鍵調検索 搭載	2週 開線調分	布 井頂頻率 井頂	分析 泰昭語影應	調講師著作	宇忠稲忠	調講:	調曲	序用詞	同類調文檔	白田 於爾留	彩欄位軍
授尋詞: "寶釵 黛玉"		2 右數:	5 文章			□停用詞 □		Lange de la constante de la co	F增	編員		刪除
# 左方 關鍵詞	右方出現詞	來源3	A REPORT OF THE					啟用	#	Category	Words	Source
3 • 那邊 寶釵 黛		氏 043 不以為 049	壬熙軍 正言彈炉愛 開取裝提界白馬 新規 整 制 置 前 販 着 一 二 一 二 一 二 一 算 炉 使 愛 馬 慶 一 一 一 二 二 算 炉 愛 慶 慶 一 一 一 一 一 二 二 一 一 一 二 二 一 一 二 二 一 一 一 一 一 一 二 二 二 二 二 一 二 二 二 二 二 一 二 二 二 二 二 一 二 二 二 二 二 一 二	林 編云(時語總書社 深了(時語總書社 照) 服 宗 香 菱 博 解 石 州	香.docx 膻.docx				1 2 3	主角 丫嬛 主子	【寶玉, 買 【時受, 覇 [迎春, 探	本次新增
匯出 搜尋到4筆結果						上一页	百一不	<	重入	(E)		

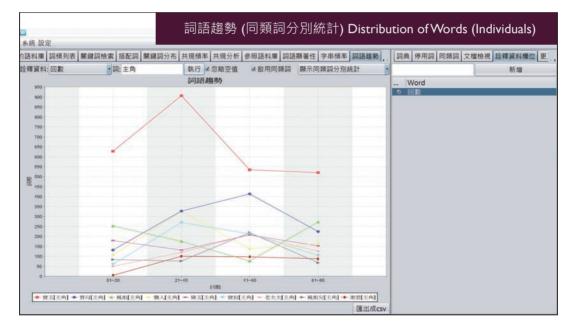


	tη λ	、文件的詮釋資料 Ad	ld Matadata ta F	
■ 系統 設定	COR /JII/	人又什的註梓貝州 AC	d Metadata to L	ocuments
標的語料庫詞積列去 開線調檢索 搭配詞 關鍵詞分布	共現頻率 共現分析 参照語	料庫 詞語顧著性 字串頻率 詞語志	詞典 停用詞 同類詞 文檔	会視 詮釋資料欄位 更
匯人 编辑詮釋 重新斷詞	Contract Provident Provident	移除選擇	Transfer the state of the state	新增
# 標題 部	同 段落 回數 档	*案路徑	Word	
1 001 転士陽學幻識通靈 頁而村是屬懷閣秀.docx (2 002 置夫人組織攝例/// 金子蘭演習結晶的 docx 4 3 003 金陵城老 4 004 獨命女優 5 005 開生品夢 6 006 買賣五彩 7 007 送宮花周 5 005 配置銀小 7 007 送宮花周 5 005 配置銀小 5 005 配置電子 5 005 配置銀小 5 005 配置 1 011 金壽振舞 606 菜和盼替去蘭葱粥 606 菜和盼替去蘭一個小 606 菜和盼替去蘭葱粥 606 菜和盼替去蘭一個小 606 愛爾爾子 605 毫易和和子 606 夏元都 605 毫易和和子 606 夏元都 605 毫易和和子 606 夏元都 605 毫易和和子 606 夏元都 605 毫易和和子 606 夏金和 605 章和助替去。 606 章 605 毫易和 605 章 605 章 605 章 605 章 606 章 605 章 7	6050 1 01~20 F 4463 1 01-20 F 度料借位 回數 21~40 i 040 史太君兩宴大觀圖 金 039 村老嫗誤該承色笑 與 038 林瀟湘魁寒菏花詩 薛 037 秋爽驚偶結海樂社 著 036 繡鬚鴦夢兆経芸軒 識 035 白玉銅親嘗蓮葉羹 黃	**ratu: //research/CORPRO/corpus/近 */research/CORPRO/corpus/近 */research/CORPRO/corpus/近 */rCORPRO/corpus/近 */research/CORPRO/corpus/近 */research/CORPRO	w Word	
19 019 (備切可良窓花秘語) 整線級第日王生著。docx 20 020 王陽麗正宮復空會 林憲工価格請請進資金docx 21 021 賢嬰人講職業資玉 位乎兒歌語教育通過docx 22 022 翻四記契詞通動語 社只学書細層芳心。docx 23 023 西周記契詞通動語 社只学書細層芳心。docx 24 024 節金剛觀別向蓋供、現文完選和影相思。docx 25 025 隔震法防導違元。 紅種等通過夏風。docx 26 026 錄聲機能言傳心事 満指態者回發影問。docx 26 026 錄聲機能言傳心事 進指態者回發影問。docx 28 028 施王強情謝影響著 證實契約	6899 1 01-20 F. 3971 1 01-20 F. 5191 1 F. 5191 1 F. 525 1 F. 6286 1 F. 5809 1 F. 5417 1 F. 5470 1 F.	Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al Aresearch/CORPRO/corpus/al for analyzing data fr	om different per	spectives)

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

玩庫博:人文社會學者的中文文本探勘工具





的語料庫	調頻列	表 關鍵	同檢索 搭配	詞關鍵	同分布 共	現頻率 共現分析	参照語料庫 詞語	願著性 字串頻率 詞語趨勢 .	詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位 更
史辱詞:	Contractor	執行		And Designation of the Party of	標點	•距離限制:	門檻值:	☑ 啟用停用詞□ 啟用同類詞	
	Fr	eq.	Co-Words		文章				
	1	2173	笑道		段藩			4	1
	2	642	寶玉道		句子			1	1
	3	545	個兩						
	4	526	F 4		標點				
	5	304	二人		-				
	6	287	寶玉笑						
	7	266	忙遁						
	8 9 10	256	夏母 迎						
	9	233	上在						
	10	225	便說寶釵道						
	11 12 13 14	224	寶釵道						
	12		也有						
	13	217	呢還						
	14	204	襲人 道 道 黛玉						
	15 16	203	迎黑玉						
	10	202	道鳳姐						
	17 18	194	個有						
	18	192	因 道 寶玉 見						
	19 20	188							
	21	172	寶玉聽						
	21 22	171	莫五郎 笑都						
	23	169	一面說						
	23 24	165	(雨 2話						
	25	162	常設						
	26	157	在這里						
	25 26 27	156	只說						
	28	153	と笑在記 記 記 記						
	29	151	笑費母						

a 44 80.	44 V					on no nip	ha D:\COR	rito m		見分析					
系統 設	E														
語料庫	調頻列表	關鍵調檢索	搭配詞 關鎖	建詞分布	共現頻率 共	現分析参	照語料庫	詞語顯著性	字串頻率	詞語趨勢。,	詞典 停用詞 [可類詞 文	當檢視 詮釋	資料欄位	更
廣玉		1	東點				執行	Jel .	歐用停用詞	□ 啟用問類詞		授辱	清除	< >	-/
#	Word.	z_score	t_score	MI	MI^3	chi-sq	simple-ll	dice co	odds-ra	log-like					
	1 通	17.924					-222.996								
	2 忙	6.701				364455									
	3 笑	11.098	18.899	1.536	1.741	133468	-110.660	0.106	29.580						
	4 聽	9.658	18.735		3.358	47513.5	-54.914								
	5 就	-8.413	-5.033		-1.564	18688.6	59.290								
	6 !	-12.512	-5.503			15735.1	62.231	0.005							
	7 與	4.292	6.395		0.732	10651.3	-24.790	0.036							
	8 聽說	6.887	18.036			8897.739									
	9 看	3.282				7631.831	-20.198								
	10 可	-9.163	-3.351	-2.903		5471.140									
	11 和	4.117	6.104	1.137	0.706	4660.928	-23.145								
	11 和 12 因	4.492				4607.900				14.234					
	13 我們	-10.907				4574.739									
	14 問	4.185	5.964		0.513	3878.022									
	15 姑娘	-9.803	-3.799			3793.080	29.019								
	16 在 17 同	3.335	4.129			3565.468									
	17 同	1.922	2.667	0.94		3271.343	-6.275								
	18 一個	-3.267				3016.438	19.251			3.271					
	19 房中	2.448				2780.118									
	20 呢	-8.017	-4.337	-1.773		2667.453	41.440								
	21 鳳姐	-4.968	-3.143			2493.186				5.769					
	22 人	-4.134				2456.802	35.205								
	23 怎麼	-4.833	-2.718			2265.792	16.616								
	24 鳳姐兒	-4.833	-2.718			2265.792	16.616	0.003							
	25 襲人	2.610	3.364		0.189	1654.393	-15.274	0.026							
	26 喝	1.389		1.074	0.596	1592.703	-2.824	0.005	27.313	1.403					
	27 罷	-9.740				1459.767	32.813								
	28 吃	-2.042				1385.241	12.251								
	29 林黛玉	3.866		1.710	2.401	1298.896				14.097 -					

	標的語料	·庫 vs 參照語料庫	I / Target Co	rpus vs. Re	ferenc	e Co	rpus	
系統 設定								
標的語料庫 詞頻列表 關鍵詞檢索 搭配	詞 關鍵詞分布 共現頻率	共現分析 參照語料庫 詞語顯著	性 字串頻率 詞語;	詞典 停用詞 同算	頃詞文檔檢	視詮釋	資料欄位	更
匯人 編輯詮釋 重新斷計	1		移除選擇		授辱	青除	: >	-
# 標題	詞 段落	回數 檔案路徑	N. N. M. N. M.					
	香入家塾.docx 4954 1		RPRO\corpus\#I					
2 082 老學究講義警頑心 病瀟湘痴动	懇職惡夢.docx 6099 1		RPRO\corpus\#I					
	電銀吞聲.docx 5804 1		RPRO\corpus\紅					
4 084 試文字寶玉始提親 探觸風寶環	責重結怨.docx 5522 1		RPRO\corpus\#I					
	意放流刑.docx 5843 1		RPRO\corpus\紅					
6 086 受私賄老官翻案牘 寄閒情淑女	て解琴書.docx 4917 1		RPRO\corpus\紅					
	k入邪魔.docx 4855 1		RPRO\corpus\紅					
8 088 博庭歡寶玉贊孤兒 正家法賈珍	◎鞭悍仆.docx 4826 1		RPRO\corpus\紅					
	臺剛絕粒.docx 4425 1		RPRO\corpus\紅					
	8號叵測.docx 4439 1		RPRO\corpus\紅					
	E宴談禪.docx 3837 1		RPRO\corpus\紅					
	文參聚散.docx 5097 1	1 F:\research\COF	RPRO\corpus\紅					
	III展月案.docx 4671 1		RPRO\corpus\#I					
	重知奇禍.docx 5966 1	1 F:\research\COF	RPRO\corpus\#I					
15 095 因訛成實元妃薨逝 以假混真劉	王座順.docx 4688 1	1 F:\research\COF	RPRO\corpus\#I					
16 096 瞞消息鳳姐設奇謀 泄機關鍵兒	記述本性.docx 4754 1	1 F:\research\COF	RPRO\corpus\#I					
17 097 林黛玉焚稿斷痴情 薛寶釵出聞	間成大禮.docx 6984 1	1 F:\research\COF	RPRO\corpus\#I					
18 098 苦絳珠魂歸離恨天 病神瑛淚漸	#相思地.docx 4159 1	1 F:\research\COF	RPRO\corpus\#I					
19 099 守官廠惡奴同破例 簡邸報老章	自撤職.docx 4352 1	1 F:\research\COF	RPRO\corpus\#I					
20 100 破好事香菱結深恨 悲遠嫁寶日	E感離情.docx 4021 1		RPRO\corpus\#I					
21 101 大觀圖月夜感幽魂 散花寺神發	5 ¥ ¥ 第 第 8 8 8 8 8 7 <p< td=""><td>1 F:\research\COF</td><td>RPRO\corpus\#I</td><td></td><td></td><td></td><td></td><td></td></p<>	1 F:\research\COF	RPRO\corpus\#I					
	(驅妖孽.docx 3593)		RPRO\corpus\#I					
	· 空遇器.docx 4975		RPRO\corpus\#I					
	編前情.docx 4574		RPRO\corpus\#I					
	h平安州.docx 3883		RPRO\corpus\#I					
26 106 王熙鳳致禍抱羞慚 寶大君禱?	F 消禍患.docx 4021		RPRO\corpus\#I					
27 107 散余資費母明大義 復世職政制	*沐天恩.docx 4456		RPRO\corpus\#[
	個關患哭.docx 5068		RPRO\corpus\#I					
	z返真元.docx 6603 1		RPRO\corpus\#I					
AA AAA	1.44 I. S. I. APPEN		LARAL LAT					

						CORPRO	Alpha D:\CO	RPRO Works	pace\WS_r	red3	詞語顯著性 Word Keyness
k統 部	定										-
語料圖	調頻列表	創鍵調檢索	搭配詞	關鍵詞分布	5 共現頻率	共現分析	參照語料庫	詞語顯著性	字串頻率	詞語趨勢	詞典 停用詞 同類詞 文檔檢視 詮釋資料欄位 要
keyn	ess: log-likel.	正負肉	同皆顯示	• •	啟用停用詞	1	□啟用同類詞	a) (□忽略變形		授募 清除 < >
	詞 log-likeli	10頃(標	%(標的)	詞頻(參	%(參照)	eyness		p-Value	e		
1	" Chi-squa	re 0	0	326	0.163	32	8.21645693	4045	0.000	00000000	
1	#	0	0	316	0.158		8.14846745			000000000	
3	笑	2808	0.643	413	0.207		2.66278795			000000000	
4	實政	323	0.074	610	0.306		7.03279558			000000000	
5		1220	0.279	105	0.053	18	6.01133211	9457	0.000	000000000	
6	10	67	0.015	245	0.123		7.61107326		0.000	000000000	
7	紫腸	131	0.03	304	0.152	11	7.73707862	6902	0.000	000000000	
8	Ż	1466	0.336	267	0.134	10	1.45755290	4512	0.000	000000000	
9		1057	0.242	958	0.48	9	9.16830311	1803	0.000	000000000	
10	沒有	282	0.065	400	0.2	9	3.25446256	8015	0.000	000000000	
11	老爺	182	0.042	313	0.157	9	1.86888061	9550	0.000	000000000	
12	巧姐	3	0.001	102	0.051		1.84205434		0.000	000000000	
13	到	717	0.164	696	0.349	8	4.65215758	4122	0.000	000000000	
14	吃	906	0.207	147	0.074	7	4.58076572	8426	0.000	000000000	
15	鳳姐兒	447	0.102	34	0.017		3.71472871		0.000	000000000	
	聽見	303	0.069	377	0.189	7	2.76303672	8996	0.000	000000000	
	皆	316	0.072	12	0.006	7	0.71893415	0223	0.000	000000000	
18		513	0.117	52	0.026	6	9.36054173	1390	0.000	000000000	
19	雪雁	29	0.007	122	0.061		8.15378464			000000000	
20		3	0.001	75	0.038	6	5.44623201	0129	0.000	000000000	
	林黛玉	268	0.061	8	0.004	6	4.25183910	2017	0.000	000000000	
	薛蝌	14	0.003	89	0.045		8.62235318			000000000	
	勇	0	0	56	0.028	5	6.38074106	8425	0.000	000000000	
	豈知	Ő	Ő	56	0.028		6.38074106		0.000	000000000	
25	至	523	0.12	69	0.035	5	5.40165224	8267	0.000	000000000	
26	那里	0		55	0.028		5.37394212		0.000	000000000	1
27	丫鬟	224	0.051	6	0.003		5.15425962				
	的話	180	0.041	243	0.122		2.94005942			七表正向]顯著 Black for Positive Keyness
	心裡	135	0.031	203	0.102		1.01784222]顯著 Red for Negative Keyness

Playing with CORPRO: a Chinese Text-mining Tool for Humanities Scholars

玩庫博:人文社會學者的中文文本探勘工具

更正斷詞	岐異處 CORRECT CROSS AMBIGUITY	
A● 開始列表 緊張的地震 緊張的 B		
81년 2 (文1 Fr* N-gram	■ ● 私用学用同 男法:年政府 更正:年政府 + + + 和 = = = = = = = = = = = = = = = =	
48 : 唐 44 恒信 43 数约要 42 : 西	AR RE	
41	【 標的結果 - 國旗功能 展現時始度 展記時 展記時 共活成年 共活分析 4, , , , , , , , , , , , , , , , , ,	
37 · 11 37 · 11 36 人民 31	2211 文章 ● 2212 文章 ● 2112 第一 1122 第一 第二 1122 第一 1122 1122	
37 · 權 36 人民的 36 重素的 36 素 页 35 编唱的 34		
34		
	27 5 5 人間的 27 55 第2 5 年代 約次 27 55 第2 5 年代 28 5 年代 29 5 7 (19) 29 5 1 (19) 20 5 年代 20	
	24 35 161 117 (1972) 117 (1972) 118 1187 117 (1973) 1187 (1973) 117 (1973) 1197 (1973) 1198 (1983) 119	
	14 44 mm 15 40 mm 15 40 mm	
	18 36 単 和月	
	19:37 · 增 20:37 · 增 21:49 人民的	
	22 56 人間的 22 56 原数的 23 56 原数 24 55 年間的	
	KH CHEN, AND HC CHUEH) ICDHS	





Taiwan Biographical Ontology (TBIO) and the Study of Taiwanese Elites 臺灣人物誌本體資料庫以及臺灣菁英份子研究

Táňa Dluhošová, 路丹妮

Research Fellow, Oriental Institute of the Czech academy of Sciences, Czech 捷克科學院亞非研究所研究員

Abstract 摘要

Who were Taiwanese elites? How do we examine the sources of their power and prestige? Did it change in time and did it signal any deeper changes in the structure of Taiwanese society in time? Can the new technological advances help us to address these issues? These will be the main questions, the presentation is going to respond to.

The study of elites, is by no means a novel topic. From at least the nineteenth century onwards, Western scholars have been paying attention to Western society's upper crust and its power to decide the fate of nations. According to Mills, elites in any society represent a unique social group that concentrates wealth, power, and prestige. They also enjoy a distinct way of life.¹ Another characteristic feature is that the positions of elites are supported by hierarchically structured institutions— bureaucratic domains for example, presided over by elite members. In the manner of a feedback loop, these institutions generate prestige that elites accumulate and, again, use to secure their positions.

The study of social structures and institutions is one way to understand how elites shore up their positions. But whereas some institutions help bolster elite status, others may work against elite interests. Elites do not form a single, homogeneous group, as Charles Wright Mills implied, and therefore we need to carefully define the social

¹ Charles Wright Mills, *The Power Elite* (New York: Oxford University Press, 2000 [1956]), 9.

groups we study rather than assume their existence from the outset.²

In addition to the state of the field in research on Taiwanese elites, this talk will tackle the question of whence the prestige of Taiwanese elites originated. It will introduce a new tool and an innovative method to approach the topic. Firstly, we will reconstruct complex intermarriage networks of Taiwanese families based on family trees as they appear in a monograph by Chen Ruoji.³ Social Network Analysis will help determine various centrality measures and identify important actors. Using Gephi,⁴ we will be able to visualize and understand the function of larger structural units within this network. Secondly, using data from the Taiwan Biographical Ontology (TBIO), a database developed with partners at the Oriental Institute (Czech Academy of Sciences), we will analyze a two-mode network consisting of members of prominent Taiwanese families and their connections to various organizations. By classifying these organizations, we will be able to detect how the organizations acted as sources of prestige for individuals, families, and clusters of families. The study of organizations will shed a light on mainly economic elites from a diachronic perspective (Japanese colonial period and the early post-war period), but will also enlarge the scope of inquiry, including a large variety of subjects and entities (6,518 people and organizations, 14,367 relationships, 86 families and family lineages), which gives this investigation into the nature of Taiwanese elites an unprecedented scope.

² Shamus Khan, "The Many Futures of Elites Research," *Sociologica*, 2 (2016), 1–11.

³ Chen Ruoji, *Zongtong de qinqi: Tuokai Taiwan quan guizu de qidai yu qundai guanxi* (Taipei: Shibao wenhua, 2011).

⁴ M. Bastian, S. Heymann, M. Jacomy, "Gephi: an open source software for exploring and manipulating networks," (International AAAI Conference on Weblogs and Social Media, 2009).

誰是臺灣精英?我們如何檢視其權力和威望的來源?它是否與時改變,是 否顯示臺灣社會結構產生任何更深層次的變化?隨著科技進步之各項新技術的 發展,能否幫助我們解決這些問題?這將是本次講題內容的重點。

有關精英的研究並不是一個新穎的話題,至少自十九世紀開始,西方學者 一直關注西方社會的上層地位及其決定國家命運的力量。Mills (2000)認為, 任何社會中的精英都代表著一個集財富,權力和聲望於一身的獨特社會群體。 他們也享受獨特的生活方式。另一個特徵是,精英的地位由層級結構的機構 -官僚領域支持,例如由精英成員主持。以反饋循環的方式,這些機構產生了精 英積累的聲望,並再次用於確保其立場。

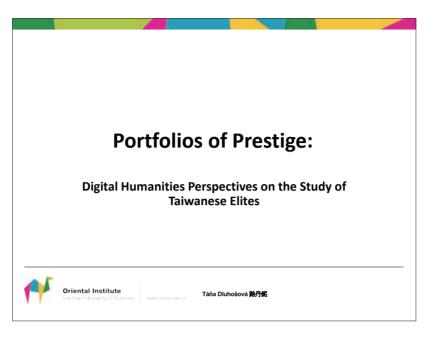
對社會結構和制度的研究是了解精英如何支撐其立場的一種方式。但是, 有些機構幫助提升精英地位,而其他機構可能會反對精英利益。因此我們需要 仔細定義我們研究的社會群體,而不是從一開始就假設他們的存在。

除了臺灣精英研究領域的現狀之外,本次演講亦將討論臺灣精英聲望的起 源問題,並將介紹一種新工具和一種創新方法來探討這一主題。首先,我們將 重建陳柔縉所著之《總統的親戚關係》一書中基於家譜的臺灣家庭複雜的通婚 網絡。社交網絡分析將有助於確定各種中心性措施並確定重要的參與者。使用 Gephi,我們將能夠以視覺化方式檢示該網絡中較大結構單元的功能。其次,利 用與東方研究所(捷克科學院)的合作夥伴共同開發的資料庫:臺灣人物誌本 體資料庫(TBIO)的數據,我們將分析一個由臺灣著名家庭成員及其與各種組 織的聯繫組成的雙模式網絡。 通過對這些組織進行分類,我們將能夠發現組織 如何充當個人,家庭和家庭群體的聲望來源。對組織的研究將從歷時的角度 (日本殖民時期和戰後初期)主要闡述經濟精英,但也將擴大調查範圍,包括 各種各樣的主題和實體(6,518人和組織,14,367個關係,86個家庭和家庭血 統),這項臺灣精英主題研究調查,其調查範圍與規模是前所未有的。

189

Taiwan Biographical Ontology (TBIO) and the Study of Taiwanese Elites

臺灣人物誌本體資料庫以及臺灣菁英份子研究



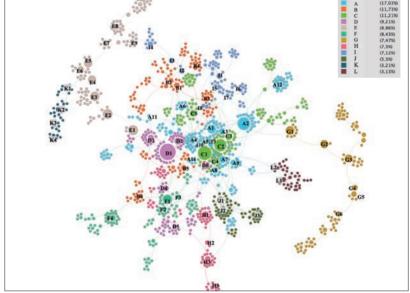
General Theoretical & Methodological Rramework

- Millsian understanding of the importance of biography and history
- Bordieusian approach oriented towards cultural and institutional factors.
- methodological framework of Social Network Analysis, which helps to identify the key families in the intermarriage network.
- Positional analysis based on biographical information about different types of organizations the family members were involved with—stored in the Taiwan Biographical Ontology (TBIO)—adds a second layer to this network which helps us to describe the various sources of prestige these families enjoyed.

Methodology and Sources

- Dataset identified by Chen Roujin consisting of members of the old families
 - 4,065 people (1,665 females; 2,390 males)
- Terminology
 - Five Old Prominent Families
 - Five New Prominent Families
 - locally important families nouveaux riches who made their fortune under the Japanese—and became suitable economic and marital partners both for the old and new families

Family Networks of Taiwanese Elites

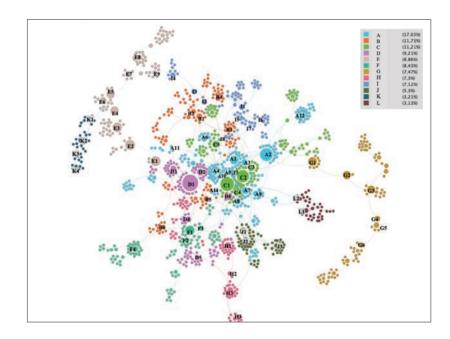


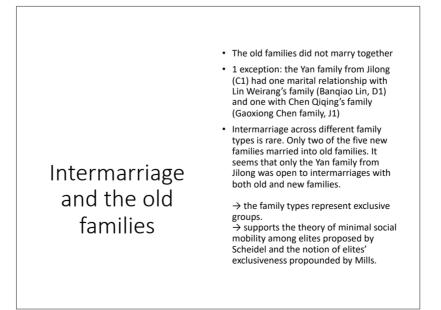


臺灣人物誌本體資料庫以及臺灣菁英份子研究

Intermarriage: Political Mainland Elites and the New Families

 political Mainland elites (clusters E and K) and the new families (cluster F) tend to cluster together and did not marry Taiwanese elites



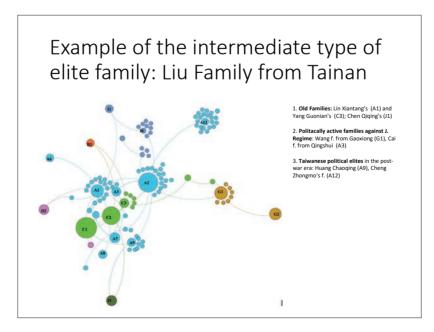


International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用

	Cai f.	Cai	He	He	He	Wu	Yu	Wu	Wu	SUM
	A4	Wancai f. A10	Yi's f. F1	Chuan's f. F2	Yong's f. F3	Huoshi's f. F4	Yuli's f. B6	Sanlian's f. I5	Xiuji's f. H1	
Yan Dexiu's f. C1	2	3	1	2	2	2	2	3	2	19
Yan Yunnian's f. C2	3	3	2	3	3	3	3	2	1	23
Lin Xiantang's f. A1	2	2	3	3	2	4	4	3	3	26
Chen Qiqing's f. J1	4	3	2	3	3	3	3	3	2	26
Lin Weirang's f. D2	3	4	2	3	3	3	3	4	3	28
Yan Guonian's f. C3	4	4	3	4	4	4	4	2	2	31
Gu Xianrong's f. B1	3	4	3	4	4	4	4	2	3	31
Lin Wenming f. A5	1	2	4	4	3	5	5	4	4	32
Chen Qichuan's f. J2	5	4	3	4	4	4	4	4	1	33
Lin Wencha's f. A7	2	3	4	4	3	5	5	5	5	36
Lin Weiyuan's f. D2	4	4	3	4	4	4	4	5	4	36
Gu Zhengfu's f. B2	4	5	4	5	5	5	5	3	4	40
Lin Weide's f. 林維德家系	5	5	4	5	5	5	5	6	5	45
SUM	42	46	38	48	45	51	51	46	39	

Family	Count	Family	Count
Yan Dexiu's f. Cl	40	Gu Xianrong's f. B1	4
He Yi's f. Fl	38	Cai f. from Qingshui A3	4
Yan Yunnian's f. C2	21	Yan Fu's F. D3	4
Yang Zhaojia's f. A8	20	Lin Wenming f. A5	3
Lin Weirang's f. D2	16	He Yong's f. F3	3
Lin Xiantang's f. Al	15	Cai Wancai f. A10	2
Chen Shoushan's f. C5	10	Wu Xiuji's f. H1	2
Lin Weiyuan's f. D2	9	Xu Ziqi's f. 許子秋家族	2
Ding f. B3	9	Weng jiazu 翁家族	2
Chen Xingheng's f. 陳星恒家族	9	Lin Yuegui's f. 林月規家族	2
Zheng f. From Xinzhu A13	9	Liu f. from Tainan A2	2
Lin Zhengwei's f. A15	8	Xu Boyun f. C4	2
Lin Rongsan's f. 林榮三家系	7	Luo Wanche's f. D6	2
Wu Jingqiu's f. 吳鏡秋家族	6	Yang Yanran's f. 楊曼然家族	1
Chen Qiqing's f. J1	6	Lin Zixi's f. 林白西家系	1
Chen Wangzeng's f. 陳宝曾家族	6	Cai Lun's f. 蔡瑜家族	1
Cai f. A4	6	Yan Guonian's f. C3	1
Li Zhaowei's f. 李兆偉家族	5	Yang Lanzhou's f. 相能測定系	1
Gao f. from Tainan II	5	Huang f. From Tainan B4	1
Cao Yonghe's f. 曹永和家系	4		



臺灣人物誌本體資料庫以及臺灣菁英份子研究

Portfolios of Prestige

Different sources of prestige among Taiwanese elites

Methodology and Sources

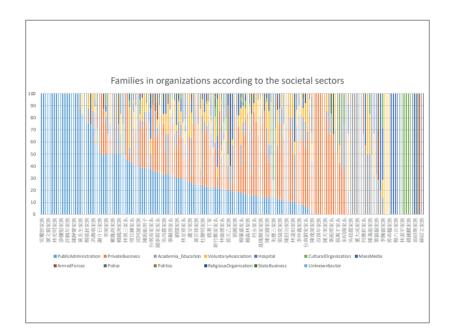
- Cross-referenced the Chen's dataset with TBIO database for further biographical information (overlap of 353 members) to conduct a positional analysis
- identify and visualize the structure of Taiwanese elite families and their portfolios of prestige
 - traced through family members' affiliations to administrative organizations, private businesses, cultural organizations, and the like. This approach resembles that of studies analyzing the behavior of elite groups in various institutional settings by tracking their careers

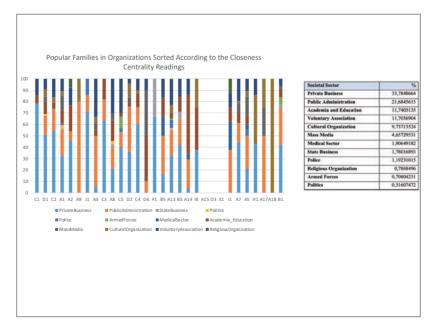
Societal Sector	Qing Period (Taiwan)	Qing and Japanese Period (Taiwan)	Post- War	Japanese Colonial Period (Taiwan)	Pre- Retrocession Republican Period (China)	Unknown Period	Total for Societal Sector
Private Business	3	67	10	610		68	758
Public Administration		85	89	379	2	29	584
State Business		4		18			22
Politics		4	5	14	1		24
Police		3		30		1	34
Mass Media		1	12	79		8	100
Medical Sector		19		55		16	90
Academia Education		100	17	177	7	13	314
Armed Forces		4	4	15	6	2	25
Religious Organization	1	2		8			11
Cultural Organization			-	2		87	89
Unknown Sector		-				1	1
Voluntary Association		8	40	202		51	301
Total No. of Organizations for Periods	4	297	177	1,589	10	276	2,353

International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用

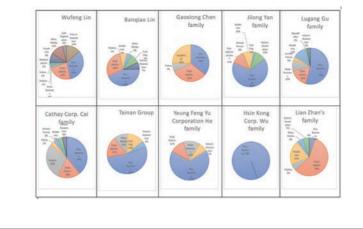
Societal Sector	Capital
Private Business	Economic capital
Public Administration	Power and social prestige
State Business	Economic capital and partly political power
Politics	Political capital
Police	Social prestige (social standing)
Mass Media	Cultural capital and indirect political power
Medical sector	Social prestige (social standing)
Academia Education	Cultural capital
Armed Forces	Military prestige and partly political power
Religious Organization	Social prestige (social standing), spiritual capital?
Cultural Organization	Cultural capital
Voluntary Association	Social prestige





臺灣人物誌本體資料庫以及臺灣菁英份子研究

Portfolios of Prestige: The Old and the New Families



Conclusion

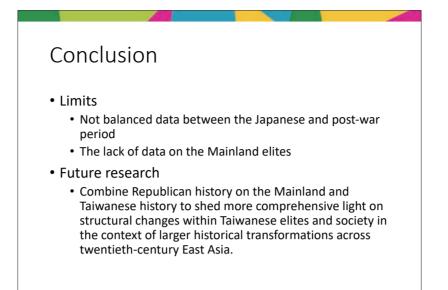
Network

- large scale networks of 1,207 elite families and investigated patterns of intermarriage between widely recognized groups of elite families: the five old families, the five new families, and four Mainland families.
- the core elite consists of the old families and certain other families who were primarily attached to the old families, but also connected them with other types.
- the new families are arranged around this core
- peripheral positions are occupied by the Mainland families

Conclusion

• Prestige

- prestige portfolios helps us understand the disposition of each family as well as network-wide trends and tendencies specific to the two main elite types
- social prestige derived from public administration seems to be the dominant element, and is often accompanied by economic success
- some elite members were active in the academic and educational sector, evoking the older concept of the highly educated and politically active scholar-official.
- Involvement in culture and arts is another salient feature of the five old families
- the new families have little academic and cultural capital



Introduction of the Database: Taiwan Biographical Ontology (TBIO), www.tbio.orient.cas.cz

Táňa Dluhošová and Wen-hsing Cheng, Oriental Institute (Czech Academy of Sciences)

Proposed poster would like to introduce TBIO database as a tool for the prosopographical study of Taiwanese elites. In addition to the introduction itself, we will show a few case studies how to use the database.

TBIO Introduction

Taiwan Biographical Ontology (TBIO) is a database developed at the Oriental Institute (Czech Academy of Sciences) storing biographical information about Taiwanese elites and those related to them. It can be used for prosopographical studies of particular groups of people and for analysis of their life and career-path trajectories. Currently the database stores information about 19,372 people; 49,967 organizations; and about 1,379 various positions within the organizations.

The database is designed as an open database where registered users can visualize but also download needed data. We also welcome colleagues to contribute with their datasets so as those who are happy to improve the current content. Their contribution is being credited on the TBIO webpage.

1. Sources

The selection of concrete people for the ontology is closely connected to the question of sources. Each of our sources has its own policy for the selection and by accepting the source, we are accepting the unavoidable bias as well. Our strategy to overcome the bias is to include a variety of different types of materials to get more plastic picture of Taiwanese society. Some of the main sources for the biographical information are:

- Database of Taiwanese Gentry under the Japanese Rule 楊建成日治時期台灣人士 *神*資料庫(1915-42). It includes 5,880 people. The sources for the database are 16 (out 28) biographical dictionaries of Taiwanese gentry (士紳名鑑) processed by Prof. Yang Chien-cheng.
- Biographical dictionary of the Post-War period《臺灣時人誌》(台灣省行政長官公署, 1947); it includes 999 people, but the contemporary editors included only those connected to the regime. The selection divides people according to the place of origin (i.e., Taiwanese (507 people) and Mainlanders (492 people).
- Employment in the state organizations 《台灣省各機關職員錄》(台灣省行政長官公署, 1947): includes 3,811 names and is used to identify state bureaucrats.
- 4. Taiwanese Writers and their works (2007作家作品目錄); this dataset is derived from a biographical dictionary compiled by the National Museum of Taiwanese Literature and currently includes only 105 writers of the post-war period.

- 5. **1940-1949 literary journals and supplements catalogue** includes 8,550 people (PR1940_49) who published in literary journals (1940-49) and selected literary supplements (1945-49);
- 6. **Wufeng Lin Family tree** as an example of prominent Taiwanese families from the secondary scholarship, 538 people.
- 7. Family trees based on Chen Roujin's 陳柔縉, 《總統的親戚:揭開台灣權貴家族的臍帶與 裙帶關係》 (時報出版, 2011).

Just to mention a few.

2. Information stored

The structure of TBIO is mainly influenced by the types of information an biographical analysis requires. Being a prosopographical tool, TBIO is designed to analyze the following aspects:

For People

- Kinship relationships
- Employment
- Education
- Association
- Regional affiliation
- Cultural activities (catalogue information and membership to literary groups)
- Political relationships (memberships in political parties and affiliation with political cliques) will be included later.
- Religion will be included later.

For Organization (to be implemented soon):

- Employed people
- Organizational structure included organizational parts
- Predecessor and following organization
- Level of operation
- Affiliation to a Societal Sector
- Regional affiliation

Users of TBIO can download datasets for Social Network Analysis, Principal Component Analysis, regression calculations, and GIS. Users will be able to choose from individual datasets.

3. Structure

Relational Database Management System (RDBMS) is designed for tabular data and the search is for well-defined questions. Out of the predefined questions, the SQL queries will include a large number of joins which lead to an explosion of complexity and computing resource consumption. RDBMS is also difficult to implement certain kinds of basic queries using SQL and relational databases, such as the shortest relation between two person nodes, unfortunately, which is what the researchers desired to know.

As a result, TBIO is using ontology to solve the semantic query problems. TBIO is organized as a **graph database**, meaning that it uses graph structures for semantic queries with nodes, edges and properties (also known as triple) to represent and store data. A key concept of the system is the *graph* (or *edge* or *relationship*), which directly relates data items in the store. The relationships allow data in the store to be linked together directly, and in many cases retrieved with one operation. Graph databases, by design, allow simple and fast retrieval of complex hierarchical structures that are difficult to model in relational systems by a query language, in our case SPARQL.

One of the reasons why we opted for the ontology structure was that we did not want to design all properties in advance, but rather followed the realities reflected in our sources. We also needed a more flexible system that would allow us to encode all available information and from them to retrieve the hierarchy. In other words, we used the bottom-up modeling of our ontology.

4. Users interface and editing mode

The web interface enables:

(i) Editors to add and edit data (Person page) without prior knowledge of the SPARQL language(ii) Registered users can browse and visualize biographical information as

Tables with Information overview

earch Person	Search	Information Timeline SNA FamilyTree	
Dan_Bird ludy ludyHillKornhoff Ricahrd_BPalmer		丁瑞彬 ▼Basic	
一背和枝 丁一枝 丁世線 丁仲態 丁伯總 丁伽藏 丁唯載 丁離戰		Basic Information	Edit Add
		Name 丁瑞彬@zh	Source Dataset
智盟 壽泉 守中 幼如 廣伊		Gender Male	Source Dataset
廣備 廣欽 憲社 根社 桐		Nth Child SecondSon	Source Dataset
榮松 淳兒 添		Birth (DD/MM/YYYY)	Source Dataset
「珍兒 「琴英		1/2/1898	yjc

Timeline

Search Person	Search	Information Timeline SNA FamilyTree						
Dan_Bird Judy JudyHillKomholf Ricahrd_B_Palmer 一青和枝 丁一天		丁瑞彬						
了一天 丁世傑 丁世豪 丁仲翹 丁伯雄		😬 出生						
丁叔豪 丁维農 丁普玖 丁普理		Θ						
丁善鞭 丁爵泉 丁守中 丁幼如 丁廣仔		1904						
了廣借 丁廣徽 丁廣鉱 丁念莊		唐港街協議會						
丁根杜 丁桐 丁榮松 丁厚兒		143議會員 790 1924						
丁添 丁玲兒 丁琴英 丁瑞園								
丁瑞彬 丁瑞魚 丁瑪池 丁祥烈		高砂鑄造株式會社理事會 常務董事						
丁秀卿 丁細寧 丁若夢 丁逸民		1924						
丁逸郎 丁逸龍								

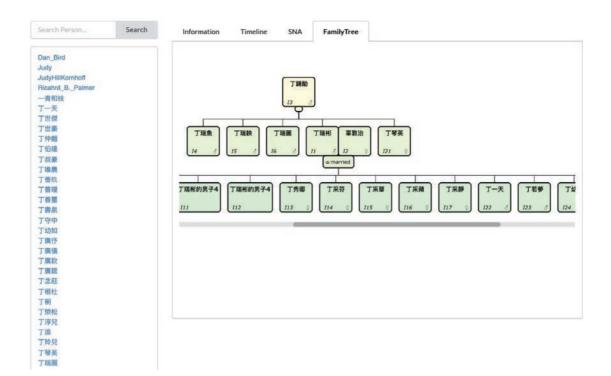
Social Network Analysis

-



- Family Tree

臺灣人物誌本體資料庫以及臺灣菁英份子研究



(iii) Advanced users can explore all data with SPARQL which can be downloaded

	TBIO Browse Query Sources About		A • •
• ?•			
SELECT * WHERE (?s ?p ?o . FILTER(sameterm(?o, tbio:Person))			Search stored queries
Save		Send	
			- Public queries
	P	٥	+ sddtaaaaaa
tp://tbio.orient.cas.cz#	http://www.w3.org/1998/02/22-rdf-syntax-nsil/type	http://tbio.orient.cas.cz#Person	• 3000388888 • a2en
ttp://tbio.orient.cas.cz#	http://www.w3.org/1999/02/22-rdf-syntax-nsi/type	http://tbio.orient.cas.cz#Person	a2
tp://tbio.orient.cas.cz#歐深的回男	http://www.w3.org/1999/02/22-rdf-syntax-nsiltype	http://tbio.orient.cas.cz#Person	
tp://ibio.orient.cas.cz#歐深的長男	http://www.w3.org/1999/02/22-rdf-syntax-nsiltype	http://tbio.orient.cas.cz#Person	Transition a2 zh Author
tp://tbio.orient.cas.cz#歐洲的次男	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://tbio.orient.cas.cz#Person	Nekdo jiny nekdo-jiny@gmail.com
tp://tbio.orient.cas.cz#歐深的三男	http://www.w3.org/1999/02/22-rdf-syntax-nsiittype	http://tbio.orient.cas.cz#Person	* test2
tp://tbio.orient.cas.cz/lisNthChild	http://www.w3.org/2000/01/rdf-schema#domain	http://bio.orient.cas.cz#Person	delect · where (
tp://tbio.orient.cas.cz#hasBirth	http://www.w3.org/2000/01/rdf-schema#domain	http://tbio.orient.cas.cz#Person	7m 7p 7o . FILTER(sameterm(7o, thiorPerson))
tp://tbio.orient.cas.cz#hasBirth	http://www.w3.org/2000/01/rdf-schema#domain	http://tbio.orient.cas.cz/Person	1
tp://tbio.orient.cas.cz#hasBirth	http://www.w3.org/2000/01/rdf-schema#domain	http://tbio.orient.cas.cz#Person	Translation
ttp://tbio.orient.cas.cz#hasBirth	http://www.w3.org/2000/01/tdf-schema#domain	http://tbio.orient.cas.cz#Person	Author Vincent whichang3448 gmail.com
tp://tbio.orient.cas.cz/fhasBirth	http://www.w3.org/2000/01/rdf-schema#domain	http:///bio.orient.cas.cz//Person	
ttp://tbio.orient.cas.cz#hasBirth	http://www.w3.org/2000/01/rdf-schema#domain	http://lbio.orient.cas.cz#Person	* select all

5. Acknowledgements:

Financial support:

- Bilateral project: "Concepts in Contexts: A Corpus-based Approach to Literary Field in Early Post-war Taiwan and its Application to the Sociology of Literature" (Alvin C. H. Chen, NTNU Taiwan; 2016–2019); collecting and cleaning data.
- "Memory in the Digital Age" of the Czech Academy of Sciences then supported the creation of the user's interface and various functionalities

Council:

- Chu Ping-tzu (National Tsing-hwa University)
- Chang Lung-chih (Institute of Taiwan History):

Shared Data:

- Mr. Yang Chien-cheng who made his HTML database of Japanese elites available for TBIO
- The National Museum of Taiwanese Literature which shared information on Taiwanese writers

Team:

- Táňa Dluhošová
- Petr Šimon, database designer
- Wen-hsing Cheng, database designer and IT support
- Pan Po-yang, IT support
- Meng-lun Teng, assistant
- Wei-jen Ting, assistant

Development of Digital Humanities in the Field of Buddhist Studies in Taiwan

臺灣佛學研究領域的數位人文發展

Jen-jou Hung 洪振洲

Associate Professor, the Department of Buddhist Studies and Director of the Library and Information Center at Dharma Drum Institute of Liberal Arts (DILA) 法鼓文理學院佛教學系副教授暨圖書資訊館館長

Abstract 摘要

In recent years, the mission to preserve old texts and manuscripts has taken on a new level with digital tools. Within the field of Buddhist studies, the digitization of Buddhist texts began in the early 1990s. More than two decades later, the initiative has developed into a comprehensive network of Buddhist digital research resources. Also, due to the vigorous development of digital humanities technology, many digital research tools and platforms in the field of Buddhist studies are created, enabling users to conduct advanced analysis and exploration of the huge amounts of Buddhist digital

resources in a simple way.

In this talk, the speaker will focus on aspects of the development of digital resources and digital tools for Buddhist studies in Taiwan over the past 20 years. Library professionals, archivists, scholars and those interested in developing digital humanities research resources will appreciate this opportunity to understand the journey of unifying the diverse array of Buddhist texts into an all-encompassing digital resource and learn more about the future direction and potential of DH in Buddhist studies.

數位人文與數位典藏,雖是近幾年才在人文研究領域新興起的話題。但在 佛學研究的領域之中,早在1990年代初期就已經開始進行佛教典籍的數位化工 作。歷經超過長年來的積累,佛學數位研究資源已發展成為十分全面的佛教數 位研究資源網絡。加上近年來數位人文技術的蓬勃發展,也出現了許多佛學數

位人文研究工具與平台,讓使用者能以簡易方式進行佛學數位資源的整合分析與探索。

本次講座將重點式的介紹臺灣近20年來所發展佛學研究數位資源與該領域 中數位人文工具的相關發展。除了讓與會這同仁能明瞭佛教數位研究資源的概 況外,同時間也讓有興趣致力於發展數位人文研究資源的專家學者,瞭解未來 可能的發展方向與潛力。

臺灣佛學研究領域的數位人文發展



Field of Buddhist Studies in Taiwan

Dharma Drum Institute of Liberal Arts

2019 International Workshop for Professional Librarians Tentative

2019/10/29

Program (2019.10.28-11.01)

The Origins of Buddhism



Buddhism originated in the area around the Indus River in Kapilavasthu (now in Nepal and the region of India bordering Nepal), whose royal prince, Siddhartha Gautama, became Sakyamuni Buddha, and whose teachings in the 6th century before our era became Buddhism. Later development led to Buddhism becoming one of the three major religions of the modern world.





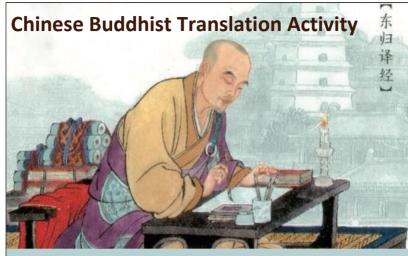


Buddhist Scriptures in India



When Sakyamuni Buddha was alive, Buddhist teachings were passed down to disciples through oral means. After the Buddha died, the disciples collected the contents of the Buddha's teachings during a large council.

臺灣佛學研究領域的數位人文發展



After Buddhism was introduced to China, large-scale translation began. This translation activity evolved into a national translation project with a scale of thousands of people and continued until the Song Dynasty.



International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用

Center for 檔案①	BuddhistStudies - Microsoft 編輯(E) 檢視(V) 移至	and the second	[明(田)			
← - -	€ 🔂 🛇 • •) 🖻 🧭 🎖 🗄	1 🖏 🛦 🎒 🗟	》 連結 《HotMail的	免費電子郵件 🥑 Internet	起始畫面
			Institute the August of the A	V. A. T. A. L. A. & BI		
				全會及 <u>諸大德</u> 贊助		
	.41	中心为或有BBS%-	"回入"那学时经中	心狮子吼站都迎彼	(H):	
	中心資訊	資料檢索	佛經原典	教學系統	其他資源	
	中心借介		建立佛典	AT 	Prot 併學網路	
	中心公告		<u>美能文佛典</u>		<u>读 佛學工具</u>	
	2 327.19.0		夏夏之供共		全 佛學院所	
	MARINE		日和佛典	2 3480	10 0007	
	聯絡電記	版傳真號碼:TEL	: (02)2363-0231 Ex	st. 3414, FAX: (02)2	1363-9029	
)	檢視頻道			一旦我的電腦	8	



Buddhist Studies Information Network Center in CHIBS (1998)

In response to the 21st century electronic information age, the rapid development of the Internet, the rapid growth of digital text, electronic publishing, information sharing and other trends in the new century, in March 1998, the Buddhist Studies Information Network Center was established to meet the needs of the times.

Objectives

- 1. Digitize Chinese Buddhist Tripitaka
- 2. Establish databases for Buddhist Studies
- 3. Maintain official website of CHIBS
- 4. Help CHIBS becoming computerized.

Work Plan

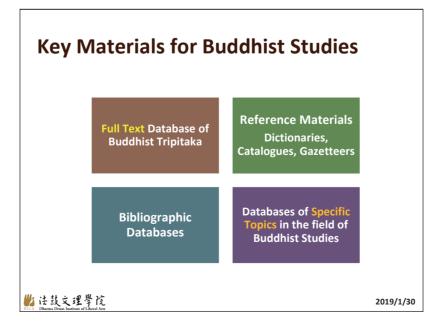
- Create Digital Databases of Buddhist Reference Materials
- Setup Online Buddhism Courses



臺灣佛學研究領域的數位人文發展



Buddhist Digital Archive Projects in DILA (法華經) 詞 2 東 1 舌 安建两级行 4 聖嚴 法師 位典誠 賢度法師華嚴著作集 【法社全集】 📕 法鼓文理學院



Full Text Database of Buddhist Tripitaka

📕 法鼓文理學院

📕 法鼓文理學院

2019/1/30

The History of the Digitization of Chinese Buddhist Texts - The Origins of the Idea

- Prof. Lewis R. Lancaster from UC Berkeley is the first person who noticed the potential of a digital Chinese Buddhist Canon.
- In 1993, he established the Electronic Buddhist Text Initiative.





2019/9/2



臺灣佛學研究領域的數位人文發展

The History of Digitization of Chinese Buddhist Texts -- Data Entry Period

- In 1990, the Yinshun Foundation start to type up 《印順法師佛學著作集》 (Collection of Ven. Yinshun's work).
- In 1994, the Electronic Buddhadharma Society (EBS) was founded by Mr. 沈家禎 (C.T. Shen) in Chuang Yen Monastery, New York. They begin to type up the 〈大般若波羅蜜多經〉 (Mahā-prajñāpāramitā Sūtra), 〈丁福保佛學大辭典〉(Ding Fubao: Dictionary of Buddhist Studies) the same year.
- From 1995, Mr.蕭鎮國(Xiao, zhen-guo) personally sponsored the typing work of Taishō Tripiţaka. By 1997, vols.1~25 had been digitalized.

📕 法鼓文理學院

2019/9/2

The History of Digitization of Chinese Buddhist Texts -- CD-ROMs

- From 1995-1997, 佛光山 achieved the digitization of the 《佛光大辭 典》(Fo Guang Dictionary of Buddhism)。It consists of around 22,650 dictionary entries, and in total more than 7 million characters. The digitization is published as a multi-media CD-ROM.
- In 1998, Yinshun Foundation released the first digitized version of 《印順法師佛學著作集》 (Collection of Ven. Yinshun's work), also on multi-media CD-ROM. It contains 42 volumes and more than 7 million characters.
- On the CD-ROM is not only the text data, but also an easy-to-use reader software.

📕 法鼓文理學院

2019/9/2

Chinese Buddhist Electronic Text Association

On February 15, 1998, Ven. Heng Qing (恆清) gathered those interested in the digitalization of Buddhist scriptures to establish the CBETA in Taipei.



2019/9/27

	Collection	Characters
CBETA Chinese	Taishō Tripiṭaka	78,770,000
Electronic Tripitaka	Shinsan Zokuzōkyō	71,220,000
	Passages concerning Buddhist activities from the Official Histories	333,000
Collection 200	Buddhist texts not contained in the Tripițaka	1,663,000
The world's largest	Selection of stone rubbings from the Northern Dynasties Period	74,000
collection of digital	Supplements from other editions of the Tripițaka	24,193,000
Chinese Buddhist Texts	Chinese translations of the Pali Canon	9,156,000
 Since 1998, CBETA has continued to expand data 	Selections from the Taiwan National Central Library Buddhist Rare Book Collection	6,801,000
and improve content	Supplement to the Dazangjing	20,478,000
quality.	Chinese Buddhist Temple Gazetteers	1,650,000
	Collection of Ven. Yinshun's work	7,540,000
	~2018.07.20	c.a.223,822,000

CBETA Online Reader http://CBETAOnline.dila.edu.tw ■ T0251 般若波羅蜜多心經 ④ / 巻/篇章 一 Ø English ♠ About III 六明太祖高皇帝御製序 般若波羅蜜多心經序 卷/篇章:[1] ■開閉目次 ○ 追定 希 經文傳送 Q.很非 般若波羅蜜多心經 唐三藏法師玄奘譯 觀自在菩薩行深般若波羅蜜多時,照見五蘊皆空,度一切苦厄。 「舍利子!色不異空,空不異色,色即是空,空即是色;受、想、行、識,亦復如是。 「舍利子!是諸法空相,不生不滅,不垢不淨,不增不滅。是故,空中無色,無受、想、行、 識;無眼、耳、鼻、舌、身、意;無色、聲、香、味、觸、法;無眼界,乃至無意識界;無無明 亦無無明盡,乃至無老死亦無老死盡;無苦、集、減、道;無智,亦無得。 「以無所得故,菩提薩埵依般若波羅蜜多故,心無罣礙;無罣礙故,無有恐怖,遠離顛倒夢想, 究竟涅槃。三世諸佛依般若波羅蜜多故,得阿耨多羅三藐三菩提。 「故知般若波羅蜜多,是大神咒,是大明咒,是無上咒,是無等等咒,能除一切苦,真實不虛, 故說般若波羅蜜多咒。」 目錄 卷/篇章 A Web-based Reader for CBETA Corpus, consisting of over 200 million Chinese Characters.

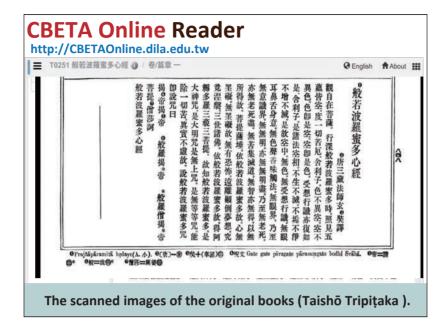
CBETA Online Reader

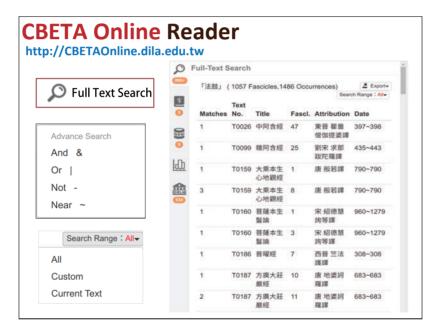
http://CBETAOnline.dila.edu.tw

T08n0251_p0848c04	般若波羅蜜多心經	
T08n0251_p0848c05		
T08n0251_p0848c06	唐三藏法師玄奘譯	
T08n0251_p0848c07	觀自在菩薩行深般若波羅蜜多時,照見五	
T08n0251_p0848c08	蕴皆空,度一切苦厄。 「舍利子!色不異空,空不	
T08n0251_p0848c09	異色,色即是空,空即是色;受、想、行、識,亦復如	
T08n0251_p0848c10	是。 「舍利子!是諸法空相,不生不滅,不垢不淨,	
T08n0251_p0848c11 :	不增不減。是故,空中無色,無受、想、行、識;無眼、	
T08n0251_p0848c12	耳、鼻、舌、身、意;無色、聲、香、味、觸、法;無眼界,乃至	
T08n0251_p0848c13	無意識界;無無明亦無無明盡,乃至無老死	
T08n0251_p0848c14	亦無老死盡;無苦、集、滅、道;無智,亦無得。 「以無	
T08n0251_p0848c15	所得故,菩提薩埵依般若波羅蜜多故,心無	
T08n0251_p0848c16	罣礙;無罣礙故,無有恐怖,遠離顛倒夢想,究	
T08n0251_p0848c17	竟涅槃。三世諸佛依般若波羅蜜多故,得阿	
T08n0251 p0848c18	耨多羅三藐三菩提。 「故知般若波羅蜜多,是	

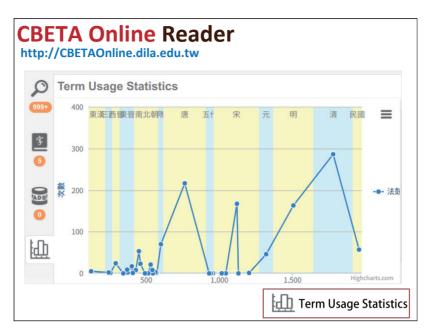
Development of Digital Humanities in the Field of Buddhist Studies in Taiwan

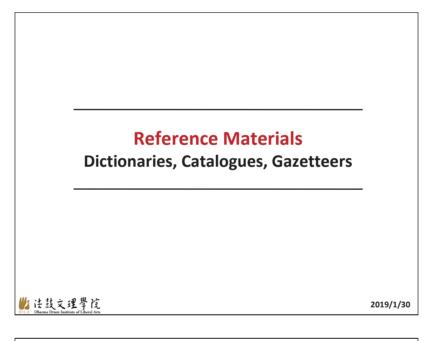
臺灣佛學研究領域的數位人文發展











Glossaries Home Digital Archives	Home Consenter Software & Fonts Search FAQ E	ゆ 中文				
	Inter-Y-concertratington Determiny-Y-1 Buddhist Studies Glossary	entries				
	Ding Fubao: Dictionary of Buddhist Studies, 1922					
Chinese-Chinese	「丁編保 (佛學大部典), 1922 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2					
Chinese-Chinese	Nanshan Vinaya Dictionary. Tapei: SeeLand Publishing, 1966-2000. 釋觀論 《闲山/律學辭典》 台北 图题序稿 1966-2000.					
Chinese-English-Sanskrit	Soothill-Hodous: A Dictionary of Chinese Buddhist Terms					
Chinese-English-Sanskrit	Seishi Karashima: A Glossory of Dharmaraksa's Translation of the Lotus Sutra. Tokyo: Soka University, 1998. 辛福時志 (正法等短期件)					
Chinese-English-Sanskrit	Seishi Karashima: A Glossary of Kumdrajiva's translation of the Lotus Sutra. Tokyo: Soka University, 2001. 中编群士 (坎)法理单控提供)					
Sanskrit-Tibetan-Chinese	Mahāvyutpatti - 《翻譯名義大集》	9,37				
Sanskrit-Manchu-Mongolian- Chinese	Pentagiot Dictionary of Buddhist Terms - (五單合理論度)					
	A Chinese Translation of A.P. Buddhadatta's "Concise Pall-English Dictionary" : Digital Edition Taipe: DDBC, 2010.					

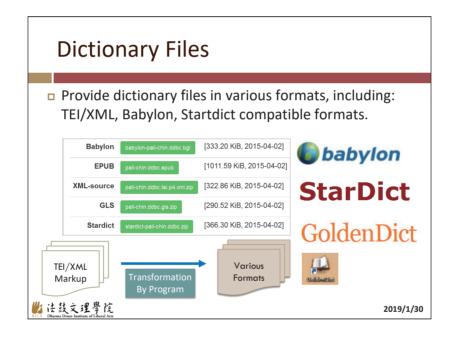
臺灣佛學研究領域的數位人文發展

DILA Glossaries for Buddhist Studies http://glossaries.dila.edu.tw

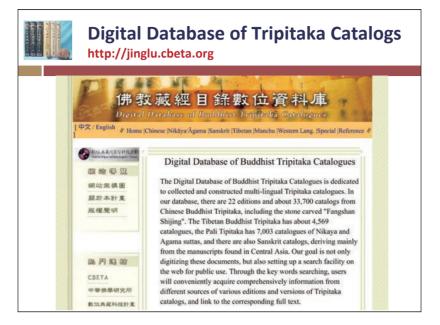
(10+ Dictionaries / 100,000+ Entries)

Provides digitization of the following dictionaries

- 1. Soothill-Hodous: A Dictionary of Chinese Buddhist Terms (Ch, En)
- 2. Mahāvyutpatti (Ch, Tibetan, Sanskrit)
- 3. Ding Fubao: Dictionary of Buddhist Studies (Ch)
- 4. Shi, Zhi-yu: Nanshan Vinaya Dictionary (Ch)
- 5. Pentaglot Dictionary of Buddhist Terms (Manchu, Ch, Sanskrit)
- 6. Seishi Karashima: A Glossary of Dharmarakṣa's Translation of the Lotus Sūtra
- 7. Seishi Karashima: A Glossary of Kumārajīva's translation of the Lotus Sutra
- 8. A Chinese Translation of A.P. Buddhadatta's *Concise Pali-English Dictionary* (Ch, En, Pali)
- 9. Dan Martin: *Tibskrit Philology* (Tibetan)
- 10. Jeffrey Hopkins: Tibetan-Sanskrit-English Dictionary

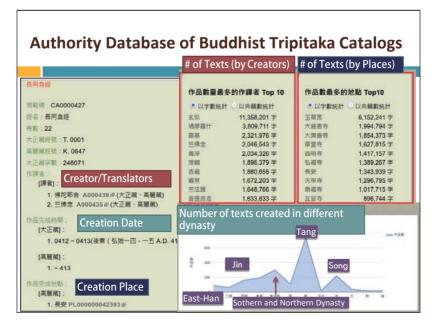


Online Services Online Search and WEB	API S	ervice	2
法鼓 Search 股每外部间典	AF	ין	
Dictionary of Buddhist Studies 注版	URL		
【雙喻】加脱減兵總罪以醫党之說法為減害養者。这等輕序品曰:「心大法處, 離大法数」」無量素輕上曰:「扣法数,吹進屬。」大集輕五十六日:「法續當摧 折,法鼓號亦應。」大煙懸違為曰:「撒鼓試兵,說数試人。」同篇祥疏曰:「扣 踏城车,合報送江魚常。梁讓於為。」		aries.dila.edu	.tw/search.json
或病来,"日時就法決束从",就是200日。」 福林之器。法堂般二鼓,其東北角之鼓,謂之法鼓,西北角之鼓,謂之茶鼓。見象	参数		
器箋十八。	Argument		Description
A Dictionary of Chinese Buddhist Terms by William	term	<string></string>	要查詢的詞條
Edward Soothill and Lewis Hodous	type	"match" "all"	 match: 詞條完全符合 all: 詞條完全或部份符合
法股 The drum of the Law, stirring all to advance in virtue.	dicts	"dila" "all"	預設值為 'all' • dila: 只搜尋 DILA 提供的詞典 • all: 除了 DILA 提供的詞典之外,也搜



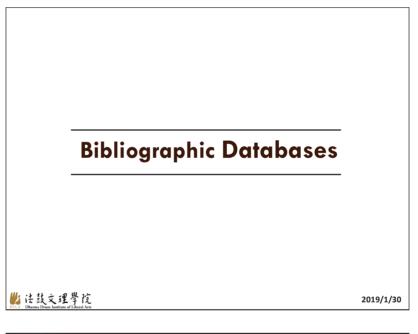
Digital Database of Tripitaka Catalogs List of Catalogs





臺灣佛學研究領域的數位人文發展







International Workshop for Professional Librarians : Development and Application of Digital Humanities



A AL			佛學裏	2位圖書館9	暨博物館	DIGITAL L	LIBRARY &	MUSEUM o	f BUDDH	iist studie
det								2. STATUS	AND AC	HIEVEMEN
-y-	The	world	's lar	gest B	uddh	ist Bi	bliog	raphic	data	base
	bik - Fu Jap - Bib	oliograp Il text c panese,	hic da lata in Englis hic rec	ta. 45 diff h, Gerr cords: 4	erent man, F 102,8 3	Langu French 3; full	ages, , Kore -text:	e forma includir an, Tibe 55,227 . мв.	ng: Chi etan.	inese,
	Chinese	Japanese	English	German	French	Korean	Tibetan	Multi- Language	Other	sum
Biblio- graphic	235,149 58.3%	102,548 25.5%	57,028 14.1%		1,680	382	181	2,551	549	402,833
full-text		18,084 32.7%	· · · · ·	349	70	30	1	1,006	22	55,227
http://bude	lhism.lib.nti	uedutw						4	流計至2018 ⁴	



臺灣佛學研究領域的數位人文發展



Databases of Specific Topics in Buddhist Studies

📕 法鼓文理學院

2019/1/30

A Digital Comparative Edition and Partial Translation of the 雜 譯 阿 Shorter Chinese Samyukta Āgama (T.100) 含 http://buddhistinformatics.ddbc.edu.tw/BZA □ The Bieyi za ahan jing 別譯雜阿含經 (BZA) consists of 364 sutras and belongs to the early Chinese Buddhist texts collectively called Ahan (Āgama) sūtras 阿含經. Generally, all of the 364 short sūtras contained the BZA have at least one parallel in Chinese and one Pāli parallel (with commentary). Often there are several parallels in Chinese and Pāli, at times even a fragment in Buddhist Sanskrit has survived. This project has created a digital comparative edition of the BZA, which connects these text-clusters. 🖐 法鼓文理學院 2019/9/2

別譯	阿含	BZA Project – Text Cluster Selector Please click on a BZA number in the left column to view a cluster.								
		Go to BZA No 別譯雜阿含 BZA (T.100)	0: bza001 雜阿含 ZA (T.99)	▼ 其他漢譯異譯本 Other Chinese Parallels	巴利語 Pāli	梵語/藏文/其他 Sanskrit/Tibetan/Others				
		bza001 Sujāta is praised 善生二種端嚴	za1062		SN,II,278 (Sujãto)					
	ome obr.s earch	bza002 The Ugly Monk 憔悴的阿羅漢	za1063		SN,II,279 (Bhaddi)					
	q	bza003 Devadatta and Ajātasattu 提婆達多貪利 養	za1064	增壹阿含(12.7) T.02.0125.0570b20 增壹阿含(23.7) T.02.0125.0614a18 雜寶藏經(37) T.04.0203.0465b20*	SN,II,241 (Pakkanta) SN,II,242 (Ratha) SN,I,154 (Devadatta)(gāthā) AN,II,073 (Devadatta) Vin,II,187 (Cv.VII.2.5)	Enomoto 1994, no.1064 (Sarighabh 2, pp.71-73; Uv 13.1f.)*				
		文理學院 um Institute of Liberal Arts				2019/9/2				

BZA Project – Reading Interface





臺灣佛學研究領域的數位人文發展

About the Yogācarābhūmi Database

- The Yogācārabhūmi database includes texts discussing different levels of practice and achievements of a yogi (the meditation practitioner or the "Zen Master").
- The Yogācārabhūmi also constitutes an encyclopedia of Mahāyāna Buddhism in its own right.
- The website of the "Yogācārabhūmi database" consists of the "Yogācārabhūmi-śastrā itself, its various translations and commentaries.



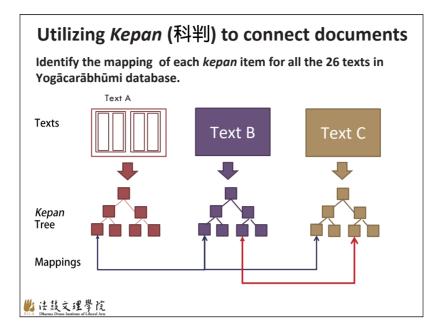
• It consists of 26 texts belonging to the Yogācarābhūmi corpus.

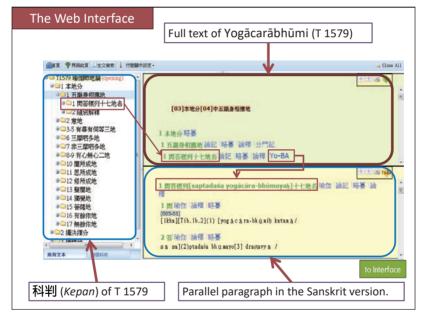
• Includes Chinese (6 translations, 7 commentary texts), Sanskrit (10 texts), and Tibetan literature (3 texts).

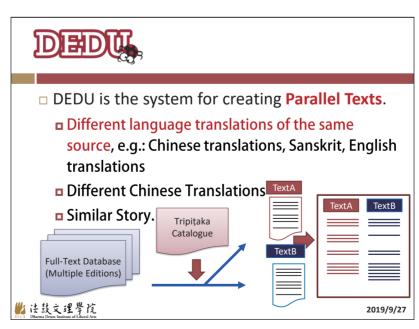
• Full-text data of all 26 texts are available on the web interface and are cross-referenced.

http://ybh.ddbc.edu.tw









Development of Digital Humanities in the Field of Buddhist Studies in Taiwan

臺灣佛學研究領域的數位人文發展



DATA Editing Tool – 「DEDU」 Text1 Text2 文章標題 III 文章標題 = (雷) 立宗:唯道無境 (雷) 立宗:唯道無境 (雷) 立宗:唯道無境 (雷) 道宗地域地域2四義沿難與称 (二) 二、小果四義迟難:處、時決 (二) 二、地主程[国](2) (二) 二、地主程[国](2) (二) 二、地主程[国](2) (二) 元、地注程[国](2) (二) 元、地注程[国](2) (二) 元、地注程[国](2) 查、正辨本宗,破計釋難 壹、正辨本宗,破計釋難 立實賬契 高仁作,承父獨分份下有后實池壹口。坐在 吳曆 边,即 圓頭圓地方,東至為為寬池,南、北至田;更有自己 手造后實池壹口,坐在 陳竹埔墘地方,東至仁理實池,西至 浦,南、北至 科技部教育部的 立貢斯提高仁作,承父闯分份下有后翼池壹口。坐在 吴厝 边,即 崩頭面地方,東至為為寬治,南、公至田;更有自己 手造后實池壹口,坐在 除竹埔墩地方,東至仁理異池,西至 浦,南、北至田 虎虎虎 虎虎虎 內改註為壹字,再照。中人: 高仁旋光緒整拾整年拾式月念 七日立貢新與高仁作在見: 胞弟高仁理、高仁根代筆:高 題制 Flexible 123, 456 Structure Tree 各毋反悔,今欲有遷,立貫断契重纸為照。 附據 吳遂龍與袭 断疑,共工能,再照。 內放註為重乎,再照。 中人: 高仁證 光緒整拾卷年拾式月念七日立實新與 高仁作在見:胞第 高仁 遵、 萬仁根代筆: 高應前 (壹) 立宗:唯識無境 (畫) 立宗:唯識無境 📕 法鼓文理學院 2019/9/27

n W	ith Built-in WYSIWYG HTML Editor	
		► \$ Q 0 I
Mattl - 《中語	[分別論》文本及注疏編撰 [123]	
盛寨 編輯 外部資	目來源 設定 說明	← BNRAB ⊙ BRAR
→ 《中還分別論》 → ◎ 壹.歸敬偈 → ◎ 貳.安立論 → ■ 貳2 → ■ 貳3	编輯節點「轉識論 真諦譚,序」 🔜 🛄 🎫	
→ 2 号、別申義 → 2 (壹) 排: → 2 一,1 → 2 一,1	調轉有二種:一個為原生,二個為法。一切所將不出此二,此二實施॥,但是調轉作二相說也123。 次用匙線有三種:	

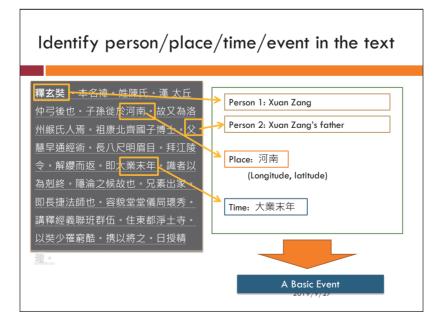
说阿彌陀經 	分享連結		× Dim
中文本(CBETA T0366)	https://dedu.dila.edu.tw/cr	eate/out/main.php?t=5cd269b3e4	43b5 網)
, 弗說阿彌陀經		複製到剪貼簿	
姚泰龜茲三藏鳩摩羅什譯	姚泰龜茲の三藏、鳩摩羅什譯す	Translated into Chinese during the Yao-Qin Dynasty by Tripi taka Master Kumārajīva of Kucha	姚秦三藏法師鳩摩羅什譯 Mapping
3	3	3	3
如是我間:	かくのごとくわれ聞けり。	Thus have I heard.	這部經是我阿難親自聽佛陀宣講 的。
* 一時,佛在舍衛國祇樹給孤獨 雨,與大壯斤傑千二百五十人	* 一時、佛、舍衛國の祇樹給孤獨 順に在まして、大社斤備、千二	At one time the Buddha was dwelling at Prince Jeta's	有一天,釋迦牟尼佛在舍衛國, 紙樹給紅獨圖裡說法,在場的有

Visualizing and Querying Chinese Buddhist Biographies

The collections of hagio-biographies of eminent Buddhist monks and nuns in China between c. 200 and 1700 CE.

Buxu Gaoseng Zhuan 補續高僧傳 by Minghe 明河, dated 1647
Chu sanzang jiji (Part 5: Biographies) 出三藏記集 (列傳部) by Sengyou 僧祐, completed 510-518
Biqiuni zhuan 比丘尼傳 by Baochang 寶唱, dated 516
Liang gaoseng zhuan 梁高僧傳 by Huijiao 慧皎, dated 519
Tang gaoseng zhuan 唐高僧傳 by Daoxuan 道宣, dated 665
Song gaoseng zhuan 宋高僧傳 by Zanning 贊寧, dated 988
Mingseng zhuan chao 名僧傳抄 (X.1523), by Sōshō 宗性, 13th cent.
Ming gaoseng zhuan 明高僧傳 by Ruxing 如惺, dated 1600

Inm Total About 2900 biographies.



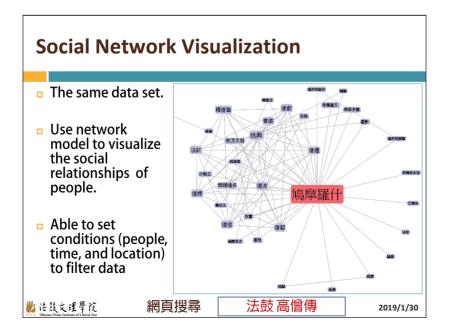
臺灣佛學研究領域的數位人文發展

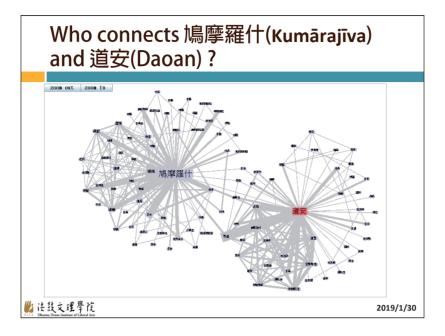
Geographic visualization of Buddhist Biographies

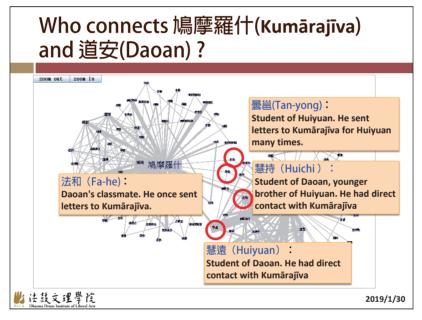


URL: http://buddhistinformatics.ddbc.edu.tw/biographies/gis/interface/ 2019/9/27

<section-header>

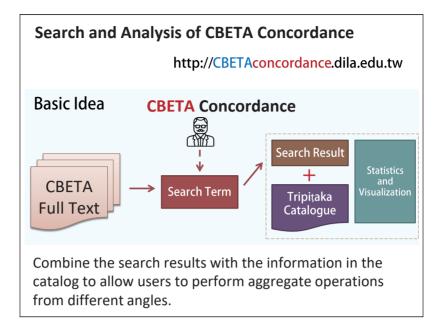




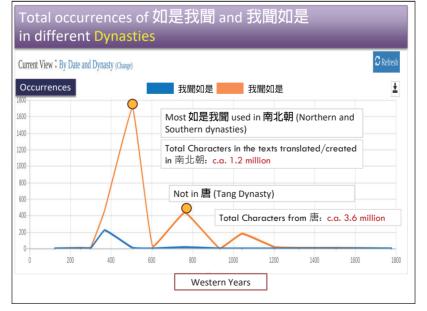




臺灣佛學研究領域的數位人文發展

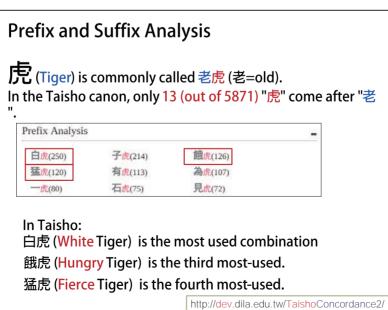












Development of Digital Humanities in the Field of Buddhist Studies in Taiwan

臺灣佛學研究領域的數位人文發展



Experiences and Perspectives – The Tools of Digital Humanities in Taiwan and their Future

經驗與展望—臺灣的數位人文建設與發展方向

Che-chia Chang 張哲嘉

Associate research fellow of Institute of Modern History & Deputy Director of Center for Digital Cultures, Academia Sinica

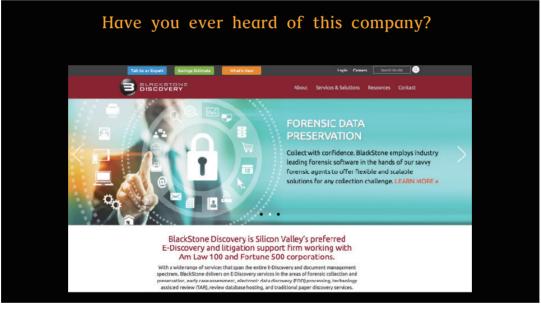
中央研究院近代史研究所副研究員暨數位文化中心副召集人

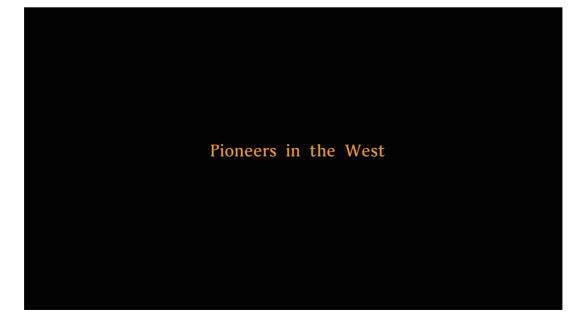
Abstract 摘要

Since the late decades if the last century, Taiwan has been promoting digital collections with rich resources supported by the government. At the first stage, a few academic institute also invested resources on constructing research databases. Such efforts keep on going and become more and opener, freer, and more user-friendly. Among the world, Taiwan's contributions to the digital humanities of sinology cannot be neglected. This lessen will begin with the emergence of digital tools in Chinese studies in the late Twentieth century, and will explain Taiwan's status in that trend originally driven by Western scholars. By introducing some representative databases built by major academic institutes such as Academia Sinica, National Taiwan University and National Cheng-chih University, I shall show the ideas, features, and powers of those resources. Also I would use the digital research platform newly developed by Digital Cultural Center of Academia Sinica to explain how the academic leaders on digital humanities in Taiwan consider and push the next step of the trend, and suggest the possible future influence over the scholars of the younger generation.

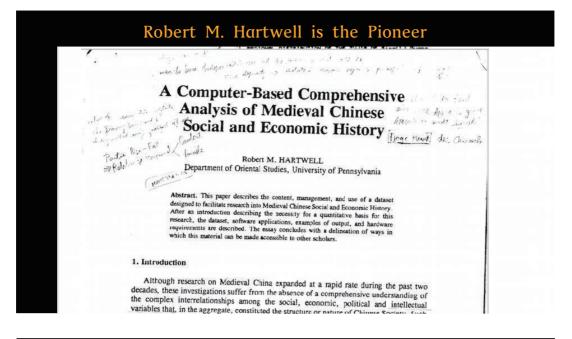
臺灣從上個世紀起即以國家力量推動數位典藏,並有學術機構挹注經費建 置數位資料庫,至今仍持續不斷且更趨開放,是華文文史學界不容忽視的寶貴 資源,即使與歐美與中國大陸相較也毫不遜色。本課程首先從歐美開發中國史 數位人文研究的背景談起,介紹臺灣在這個風潮中的地位,同時介紹中央研究 院、臺灣大學、政治大學等機構所建置、富有代表性的非營利性數位人文資料 庫的背景、理念及特色。此外,將中央研究院數位文化中心新開發的學術研究 平台為例,說明目前臺灣的數位人文領袖所關注的發展趨勢,以及未來可能對 學界所造成的影響。

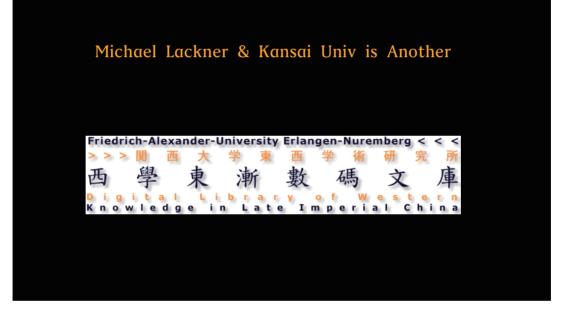


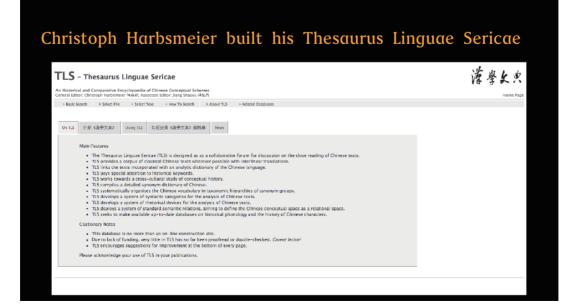


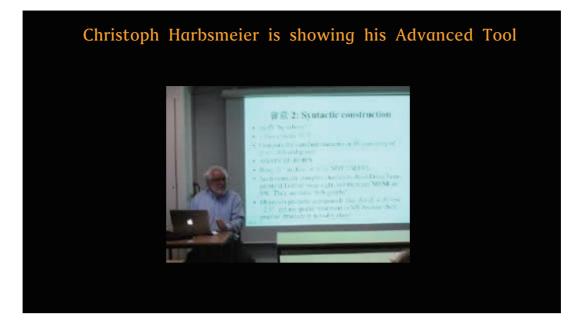


International Workshop for Professional Librarians : Development and Application of Digital Humanities











	▲ 業業主文務業系 実業端子 当首見 中			
	120002710022110		資料庫	
	前科演名稱	制作单位	······································	
	業薪全文資料庫	史語所漢籍全文資料庫計畫	申请使用	249779.271+
1 m	古漢語文劇語科庫	結合所文載活料庫研究室 資訊所詞庫小組及 中採所運講卒文資料庫計書	7 13,499,174 中請使用	
34	古漢語語科庫(與放)	部言所文載語料庫研究室 資訊所詞庫小組及 中語所運講卒文資料庫計書	1 672 384	
	基準文素權利	臺史所史籍自動化室	原臺灣方市、臺灣檔桌 47,054,288 及臺灣文戲(一)區 五) 均同今供存於資料庫	
adal	行代史杂文資料庫	- And	59.332.639(Whotestein) 新印音代人十 申請使用近代支全文資料	
The Col	清代經世文編 五現代中國史際日島	HAR T	18,292,747 検索 30章 清代即世文編 2,036.5 3 包含十個資料庫	Chi A
THE SAL	折氏紋 内開漢文顯本專題權案:	about the	119,798	
21 - 22	利利婦原類提度 文心理論 体好三婦	普訊所文獻表理研究室	1,700,01	
	装み位著作集 総合大学場	文朝所資料要并	951,560 1.542,048	
	正統道藏		6,092,586	
	詞話集成	文習所適料提其 ※期你该時小和其何非是為	95,347	
	<u>斯油史-本紀</u> 製店訪集	間支館清史組 即大国文系製	670.629 633.146	
		「計算する無所容易」		
	其常語俗曲唱本「歌仔冊」	王順連数授製	2,366.353	
	人文優科課師生造1.1	中央研究院 会容操自上刻音和库	24,907.254	11.5

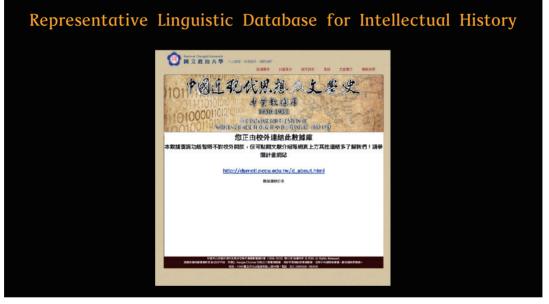
International Workshop for Professional Librarians : Development and Application of Digital Humanities

Mainl	y f	or Se	arching Key	words	
····································			学士 首页 書曰明麗 基本校	李 講院授家 吉英・筆記	山县縣 使日
1.	[命中書日	與意治】 (RH 328#8	1.19578回世际: 武平电HAB		
【基本給准】 (単行) 世界(タギ型合) 上北大 人口 働名 は カス ロ 計算	N . / 13			L	© \$\$:4000
6 武振7 (100:10-6) 同義词 (100:8018-80)		86	finit	版率	
CRUCHS (TARRES THE)	2	1. 111方版型 1. 111月前日 1. 111月前日	(細鮮)金融酸糊:創江當中醫師內約,总州中醫院性 (細)金融酸藥		4230 500
二 元素 義美 四百年人物 再成五代 法会会 元 第 第 第 再成			(第)等者回編 (第)時存留写編		4042 004
THAN 48 (228.03 2 HAN 1)		太平殿 書方	(本)工模種編創	國家業者並總自於損於本	1001 101
20220007500 000 000 (*:000007) 000 000			大師会然差	印刷州主	1921 171
			中國世界集	光潮刊市	1418 187
DML	7 7 8	5 III:44	(場)調中初,部以正,遂發神編解	上海重防備影響率	500 76
02-		a statement of the	(第)指导的重要; 他先不知多的		546 65
07.	* * *		(像)王辉健		784 BN
n a. .	. 78		(唐)所印刷 新一会市:為三朝新会		129 46
			時一世祖:周田羅斯作 (元)朱麗帝國:並且倫中醫局研究院大助研究会編校		300 48
			(近:市政設備)		150 40
		TART	(第1時前書)		104 28
			()(())()()()()()()()()()()()()()()()()		158 28
	* 7 8	1 105 I	升级康藤繁:题附以寄注用		118 21
	17 2 #	- 798.02			101 21-
		東垣動魚	(金)宇庆祖御		105 18
			(清)文慶,黄坂,黄富等敗萬	清内石沙平	55 13











International Workshop for Professional Librarians : Development and Application of Digital Humanities



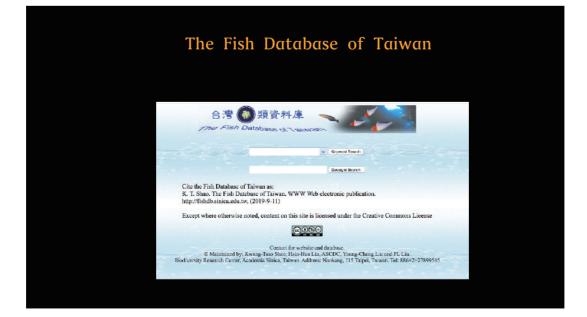




wan, o	r Uni	on Ca	talogue		
Taiwan - Culture &	Nature				
2005 7 Sptcial Petrane 7 Antodoci fra atacianisti	Faceta of Microbility Themas take active 1 at	iana contradi vialionalia asistena of Gance. Adi il Buditaliana Suationa il	-		
	10000				
Tarrier Management Management print Colo Manufactures and Mercury Tarrier Netto they are to face and mercury and the set of the set the set of the set of the set of the set the set of the set of the set of the set the set of the set of the set of the set the set of the set of the set of the set the set of the set of the set of the set the set of the set of the set of the set of the set the set of the set of the set of the set of the set the set of the set the set of the set the set of the set	Han of Trinson and the Person 7 cm police The Yaha was shift server if its "Markan channe of the Data 1 find more	Constraint Constraint Constraint Constraint Constraint Constraint Observations and page Constraint observations and sound and report, constraints			
and the second second second					
Two Reard Conferent Contribution We have not be carfulated We have not be carfulated	Transvers of Scampo to help proves consecut coord adverse, the distribution re-scale a late of four size stample on, head more	Bainbox Sheach Olice leads The Kanace Seach Cob Berle- Nat A man apparative with a berry length of it. P. A. BH., BL , Isaid Tree			
		Taiwan - Cuturo & Naturo Value value Value	TOTAL STATE TOTAL STATE	Value Part (P) Image: Pice Pice Pice Pice Pice Pice Pice Pice	

A Representative showcase: Taiwan Encyclopedia of Life





International Workshop for Professional Librarians : Development and Application of Digital Humanities







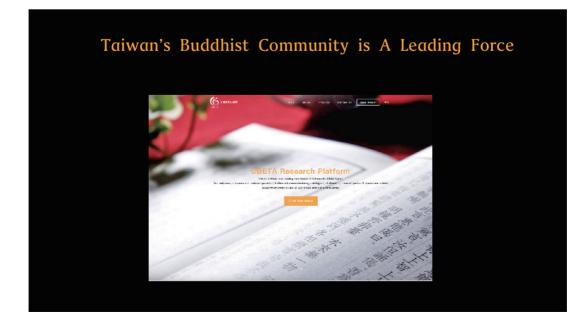
Experiences and Perspectives - The Tools of Digital Humanities in Taiwan and their Future

經驗與展望一臺灣的數位人文建設與發展方向

いて	
(영국) - 福祉(新) (1993) 1 (1993)	

Cheng-chih Univ's Database of ROC Official Posts











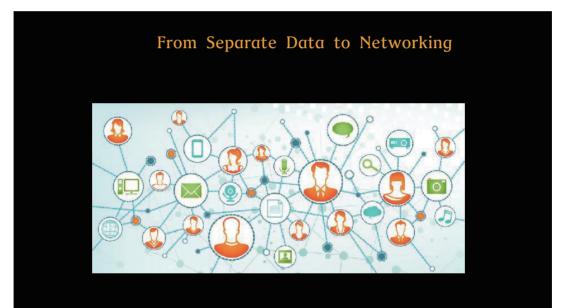
ce Price Database Provides Economic Data
清代僅價資料律用 ****
0 安期秋年 🚍 🔳
1814/1 1820/13次定(2)小安・孝子・赤孝 供付銀化 2xxx 1921(2)(2)21121 For 1215100
2000 - 124124 10 1242410
100 C
and the second
e
年代 1月 2月 3月 4月 5月 6月 7月 8月 * 月 10月 11月 12月
18 4 220,719 220,796 220,400 220,400 220,406 220,420 220,420 216,420 216,420 264,420 200,420 200,420
18 6 201,500 208/500 209/500 218/500 228/500 220/500 221/500 214/500 200/500 200/500 200/500
18 7 220(300 207/500 220/500 219/500 219/570 220/470 221/476 222/466 227/470 242/490 265/486 265/485 18 8 250/460 230/459 210/440 230/440
100 23/100 23/100 23/100 23/100

Database Increases the Chances for Cooperation











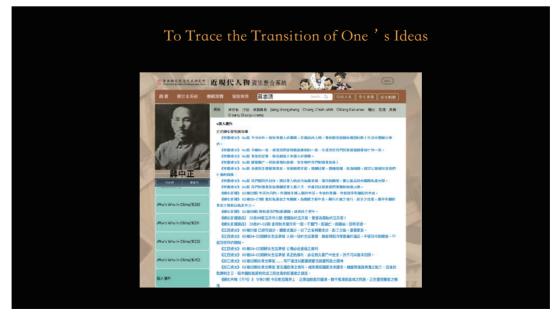
'IS Allow:	s to Further Linking to Biographies
查約 瀏览 ● 西暦 ○ 中暦 ○ 干支	▲ 近代春秋 TJS 系統 銀刻今折 鼓定 ■ E 至 10 日 11 日 ④ 時途性 55頃页 ■ 日本月 12 日日 ● 日本本 20 日日本 20 日日本 20 日日本 10 日日本 11
● 単 10: 4単 ○中華民意の単13: 4単 ○荒火江戸紙 4単 ○抗火雪茄、4単 ○王社会目話、4単 以16部 会155 山	





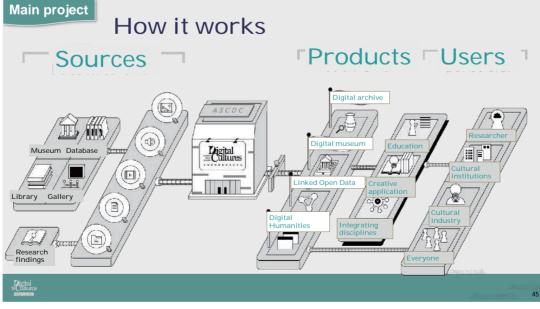
· 中央研究院近代支研究所	近代春秋 TIS 突然
至詞 瀏覽	绿铜兮折 银定
◎ 西暦 ○ 中暦 ○ 王文	
金減 2年 ○月 人名获祥 ○: 李鴻章	●日 単 ●月 ●日 ◎ 周敏進所提演
Anteo . 7/44	Mauc daas
89 AC	田田が方 ② 業券取引系統 ② 株本林 ② 規想取利用
 5近代中間史事目誌 1223年 6年華民國史學日誌 1年 5済公大事長編 6年 6栄水昌日記 2差 5千世杰日記 4差 4236年 全部列出 	近代中国史専 🚫 時間時 📄 今日 両元:1847年6月7日星期一 中暦:道光27年4月25日 丁水年乙巳月癸酉日
	6.7[四·二五] 谢银之黑·沈祥芬,华宗殿,李瑞章,定荣构,何谓《郭雷素,西新道等独士出身。(mmatx)
	相關人物:亨洛章 骤波表 汉藻伯 沈桂芬
	近代中国史第 ⑤ 時間目 ● 今日 西元·1850年6月1日星第六 中暦·道光30年4月21日 改成年年已月交末日
	6. 引四,二一,根李鸿章、沈桂芬,沈葆振与朝御,《《编》公)
	和蜀人物;李鴻章 沈东横 沈世芬
	3.4



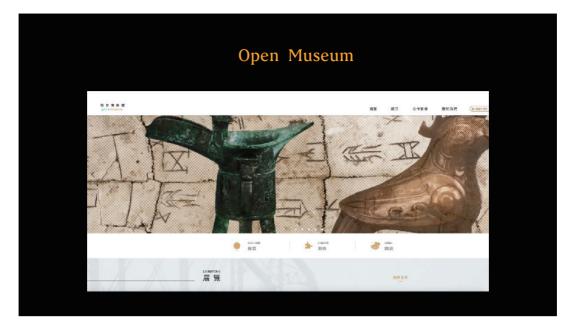






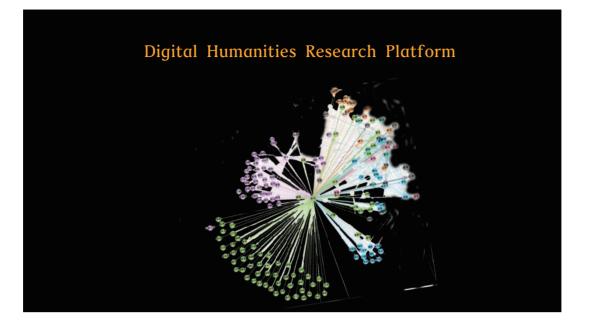


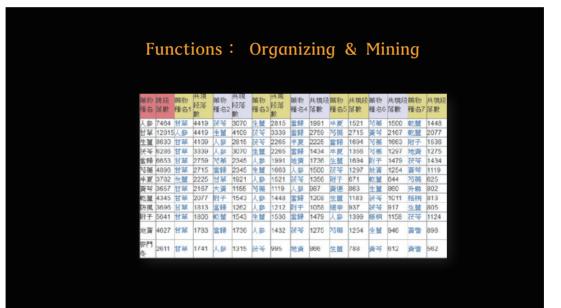
International Workshop for Professional Librarians : Development and Application of Digital Humanities

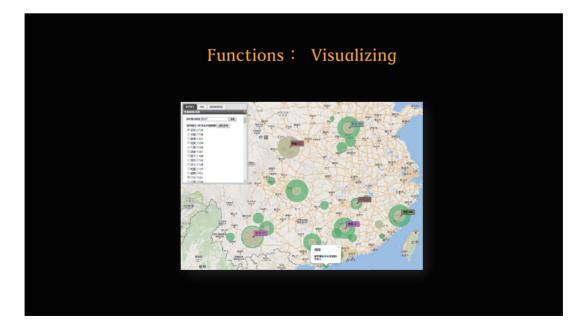




經驗與展望一臺灣的數位人文建設與發展方向







International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用



Marketing and Editorial Board



Achieved Appliances

Experiences and Perspectives - The Tools of Digital Humanities in Taiwan and their Future

經驗與展望一臺灣的數位人文建設與發展方向







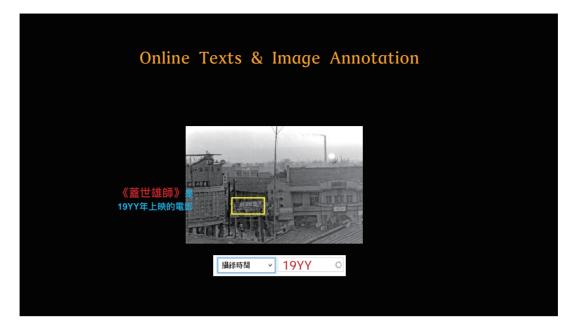


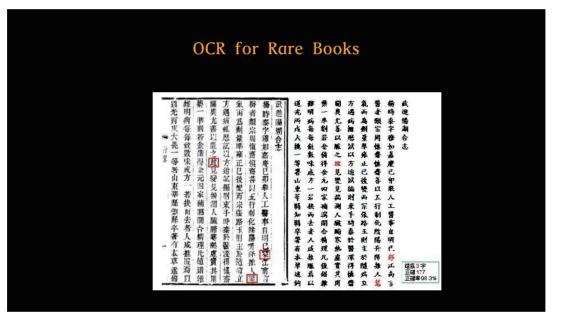




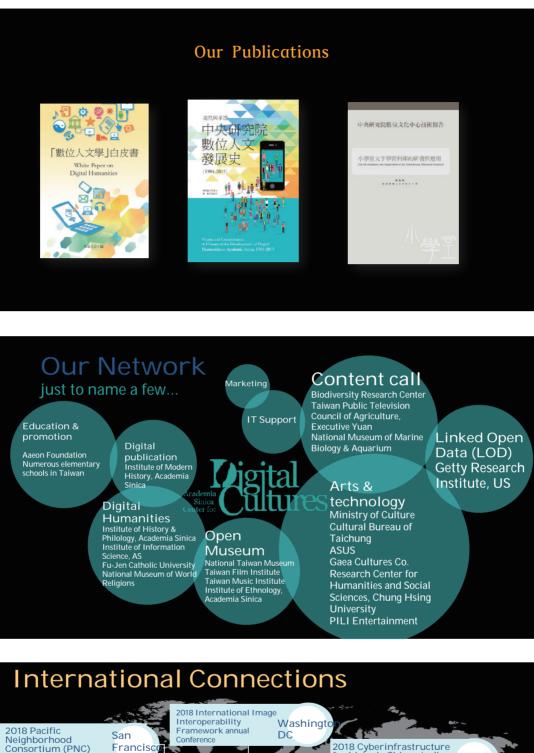
Experiences and Perspectives - The Tools of Digital Humanities in Taiwan and their Future

經驗與展望一臺灣的數位人文建設與發展方向









Experiences and Perspectives - The Tools of Digital Humanities in Taiwan and their Future

經驗與展望一臺灣的數位人文建設與發展方向



如何在學術圖書館發展數位人文研究平台支援數位人文研究 - 國立政治大學圖書館的經驗

Chih-Ming Chen 陳志銘

Professor & University Librarian Graduate Institute of Library, Information and Archival Studies National Chengchi University 國立政治大學圖書資訊與檔案學研究所教授暨圖書館館長

Abstract 摘要

Digital humanities (DH) is work at the intersection of digital technology and humanities disciplines. It includes the systematic use of digital contents and tools in the humanities disciplines and makes possible new kinds of teaching and research. In recent years, DH has received a particular attention in both the humanities and computer science fields. Particularly, academic libraries, such as Research Libraries UK and National Central Library of Taiwan, also got started to develop infrastructure to support DH research as well as the close collaboration between librarians and researchers is disclosing new opportunities for librarians to actively collaborate in DH research. Humanities research involves the analysis, comprehension, production, and sharing of a huge amount of digital sources. Moreover, humanists in DH field heavily relied on using digital tools, such as text mining, annotation, social networks analysis, geographic information system (GIS), and natural language processing (NLP) technologies to analyze data from digital sources in order to extract useful information, clues, and find new knowledge.

In recent years, Chinese ancient books is arising an unprecedented challenge and opportunity due to the rise of international Chinese study around the world. To increase the research efficiency and production, there have been a lot of international Chinese studies to explore the issues about Chinese culture, religion, economy, and politics through Chinese ancient books or literatures with the support of digital humanity tools. To provide more convenient and friendly research environments for humanists engaging in international Chinese study, increasing effort has been made not only to

make digital contents and tools available, but also to create a complete virtual research environments (VREs) that provide interpretative frameworks for making sense of cultural artefacts. This course thus focuses on how to develop and promote a Chinese ancient books digital humanities research platform (CABDHRP) in National Chengchi University Libraries which provides basic data analysis functions, such as full-text search with a single word or double words, term frequency analysis and post-query classification, as well as advanced data analysis functions, such as automatic text annotation, textual analysis, character social network analysis, and user behavior analysis, to support digital humanities research.

數位人文 (digital humanities, DH)為數位內容、工具與人文學科相互結 合的新興研究領域,主要聚焦於如何系統地使用數位內容和工具支援人文學科的 研究,並據此發展新的教學與研究模式。近年來,數位人文在人文科學和計算機 科學領域都受到極高的關注。特別的是,諸如英國研究圖書館和台灣的國家圖書 館等學術圖書館也開始發展支持數位人文研究的基礎設施,圖書館員和研究人員 之間的密切合作,正在為圖書館積極開展數位人文研究發展新的服務模式與新的 機會。人文科學的研究涉及大量數字資源的分析、理解、解譯和共享,數位人文 領域的學者依賴使用包括文本探勘、標註,社會網絡分析、地理資訊系統 (geographic information system, GIS)和自然語言處理(natural language processing, NLP)等數位工具,分析來自數字源的數據,以便提取有用的數據, 並透過信息分析來發現新知識。

近年來,由於全世界對於國際漢學研究的興起,透過中國古籍發展數位人文 研究正形成前所未有的挑戰和機遇。為了提高研究效率與生產量,在數位人文工 具的支援下,已有許多透過中國古籍或文獻探索中國文化,宗教,經濟和政治問 題的國際漢學研究被提出來。為了提供從事國際漢學研究的人文學者更方便和友 善的研究環境,必需不斷提升數位內容和數位人文工具的可用性,發展完整的虛 擬研究環境(virtual research environment, VRE),提供人文學者探索漢學研 究的框架。本課程重點在於介紹國立政治大學圖書館如何開發和推廣中國古籍數 位人文研究平台(Chinese ancient books digital humanities research platform, CABDHRP),該平台提供基本數據分析功能,如單詞或雙詞全文檢索, 詞頻分析和查詢後分類,以及高階數據分析功能,如自動文本註釋,文本分析, 社會網絡分析,以及使用者行為分析等,以支持數位人文研究。

如何在學術圖書館發展數位人文研究平台支援數位人文研究 - 國立政治大學圖書館的經驗



2

How to develop a digital humanities research platform in an academic library for supporting humanists' research: Experiences from NCCU Libraries



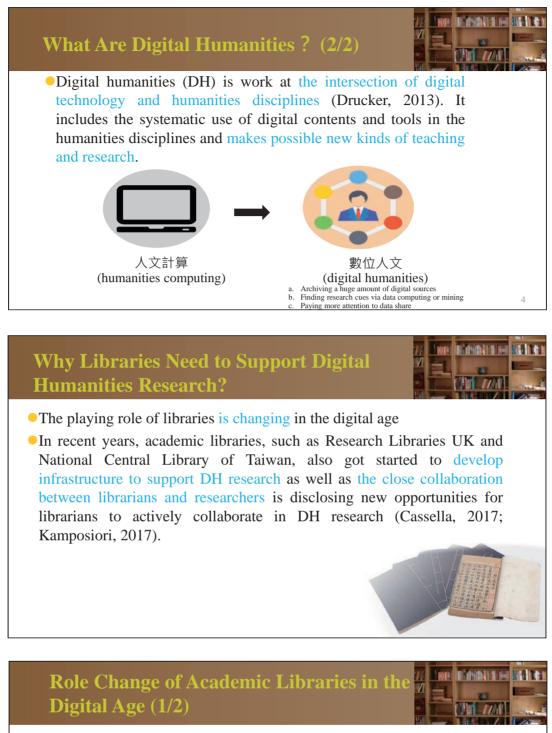
Chih-Ming Chen, Professor & University Librarian of National Chengchi University

Outline

- What digital humanities are?
- Why libraries need to support digital humanities research?
- Introduction to digital archives developed by NCCU Libraries
- Introduction to digital humanities research platform of Luo Jialun writings developed by NCCU libraries
- Experiences of developing digital humanities research platform from NCCU libraries
- Promotion outcomes of digital humanities research platform developed by NCCU libraries on campus & internationality
- Future development of digital humanities research platform developed by NCCU libraries
- Conclusions
- Q & A

What Are Digital Humanities ? (1/2) "the application of information technology as an aid to fulfill the humanities' basic tasks of preserving, reconstructing, transmitting,

and interpreting the human record." (Frischer, 2009)
"the practice of humanities research in and through information technology, and the exploration of how the humanities may evolve through their engagement with technology, media, and computational methods." (Digital Humanities Quarterly)



- The role of the library is to support learning, teaching and research as well as supporting research is the most important task of the academic library, but it is not active enough and effective in acting.
- In the face of digital era, the use of academic information has changed rapidly. However, library has not changed in time, thus leading to the gradual exclusion from the research process.
- Change is the best survival way to libraries in the digital era, and libraries should see the digital age as an excellent opportunity to develop new services for readers.
- Library is no longer just a "space" for collecting books. It should return to its more fundamental definition: a "media" of researchers and knowledge.
- The core business of library will be to connect people and knowledge. It is not in providing books. It is imperative to develop a research environment that promotes interaction between people and digital materials.

如何在學術圖書館發展數位人文研究平台支援數位人文研究 - 國立政治大學圖書館的經驗

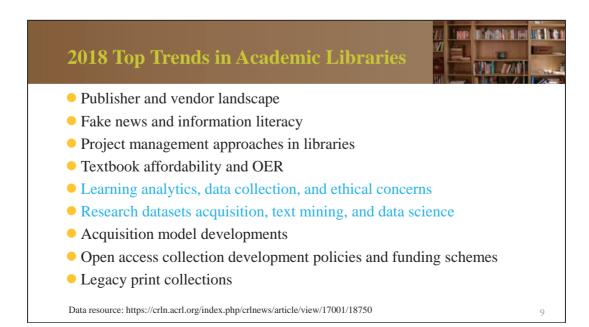


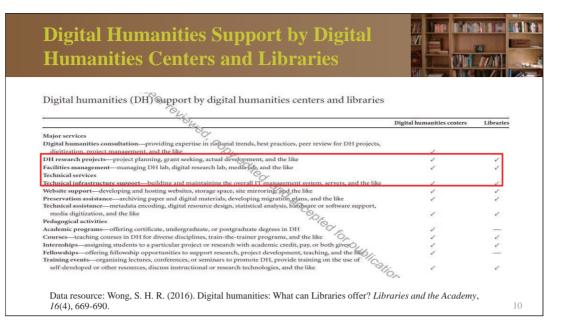
- The digital humanities are the comprehensive academic achievements of traditional humanities research with the support of digital technology. In 2014, the Association of College and Research Libraries (ACRL) published new trends in academic libraries, among which digital humanities research was listed.
- In 2014, IFLA published the Library Horizon Report at the annual meeting in Lyon, pointing
 out that the library faces six key trends, one of which is the rise of new forms of
 interdisciplinary research, with particular reference to digitalization. Technology will be
 widely and deeply applied to the research and teaching of traditional humanities and social
 sciences.
- At present, it is an important point in the life and death of the library. If library do not actively invest in the field of knowledge management in the new media era, especially in the field of digital humanities, the future will be downgraded or eliminated by the society. The reason why academic research in the humanities field requires a library not only depends on it to preserve the literature, but also requires it to provide the functions of management knowledge and knowledge sharing, particularly for digital resources.

2016 Top Trends in Academic Libraries

- Research data services (RDS)
- Data policies and data management plans
- Professional development for librarians providing RDS
- Digital scholarship
- Collection assessment trends
- ILS and content provider/fulfillment mergers
- Evidence of learning: Student success, learning analytics, credentialing
- Digital fluency in the Framework
- Critical information literacy in the Framework
- Altmetrics
- Emerging staff positions
- Open Educational Resources (OER)

Data resource: https://crln.acrl.org/index.php/crlnews/article/view/9505/10798





Examples of Libraries with Digital Humanities Services



- The professional team of the Center for Digital Scholarship (CDS) at Brown University consists of several academic librarians and information technology librarians to provide 11 studies services, including data curation, information visualization, spatial analysis, metadata creation, digital humanities, consulting, digitization, workshops, conferences, and research.
- Harvard University's Yanjing Library integrates digital resources and all Chinese rare books that have been digitally read and provided open sharing to support digital humanities research and the open sharing of resources. It is also a kind of propaganda for library resources, and to a certain extent, it also expands the influence of libraries.
- In 2016, the Shanghai Library pioneered the digital genealogy service "Family Spectrum Knowledge Service Platform" based on the openness of related materials in the public library community in mainland China, enabling the public to go through the full text of more than 6,000 genealogy on the Internet. Moreover, the Shanghai Library Open Data Platform has been developed to support digital humanities research services.

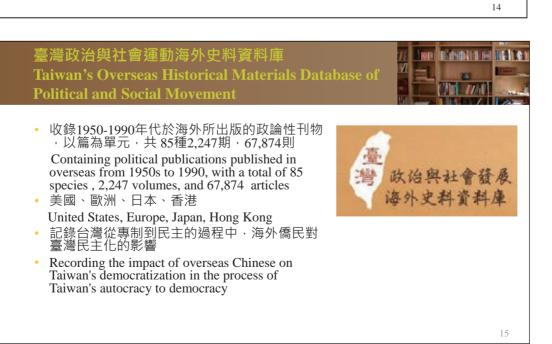
Digital Archives Developed by NCCU Libraries



Digital archives developed by NCCU libraries	Digital archives cooperatively developed by NCCU libraries and			
羅家倫文存	other institutions			
我樂多齋	中華救助總會數位資料庫			
筆匯	臺灣人權促進會人權檔案數位典藏			
知識分子與臺灣民主化:大學雜誌	臺灣總督府府(官)報資料庫			
陳芳明手稿				
臺灣政治與社會發展海外史料資料庫				
民俗臺灣				
民國38年前重要剪報資料庫				
中華民國政府官職資料庫				
顧正秋藝術網站				
茶言觀政:政大校園影像記憶網	12			



- 收球陳方明教授于個, 共1,669篇 Contained Professor Chen Fang-Ming's manuscript, a total of 1,669 articles
- 1983-2012年 From 1983 to 2012







 收錄中華民國體制以來總統府公報人 事任免命令

Collecting the appointment and dismissal order announced in the presidential office bulletin of the Republic of China

 已建置85萬筆以上資料以及15,708筆 公報影像檔,並持續新增資料

More than 850,000 documents and 15,708 bulletin image files have been built, and new data has been added



18

如何在學術圖書館發展數位人文研究平台支援數位人文研究 - 國立政治大學圖書館的經驗



臺灣總督府府(官)報資料庫 Taiwan Governor's Office Newspaper Databas

彙集臺灣總督府於日治時期發行的府報及官報
 ·共13,027期·135,970筆。

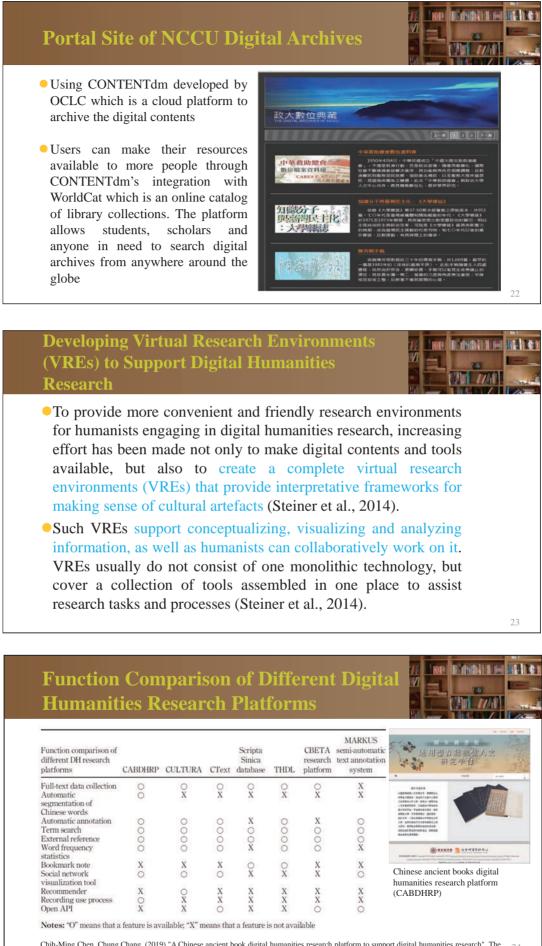
Bringing together the government newspaper and official newspaper issued by the Governor's Office of Taiwan during the Japanese occupation period, totaling 13,027 issues, 135,970 articles

 除法律、行政命令及重要行政措施外,亦包含 人事、通信、財政、司法、警察、統計、衛生 等面向之資料。

In addition to laws, administrative orders and important administrative measures, it also includes personnel, communications, finance, justice, police, statistics, health and other information



21



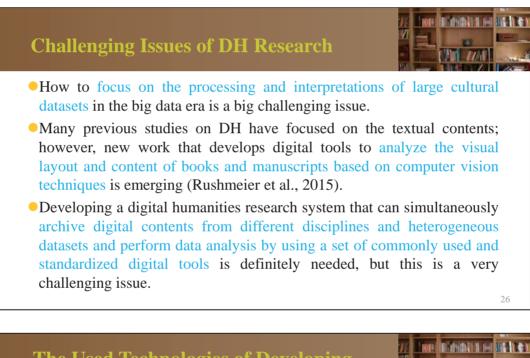
Chih-Ming Chen, Chung Chang, (2019) "A Chinese ancient book digital humanities research platform to support digital humanities research", The Electronic Library , Vol. 37 Issue: 2, pp.314-336

如何在學術圖書館發展數位人文研究平台支援數位人文研究 - 國立政治大學圖書館的經驗

Difficult Issues of Developing Digital Humanities Research Tools



- •Although a lot of digital tools have already been developed for supporting digital humanities research, many of these tools do not properly fit the needs of the humanities community due to technical difficulties or scientific objectives (Picca & Egloff, 2017).
- •Furthermore, different types of humanists have considerable differences in their knowledge of the collections, requiring varying levels of support, and every individual humanist has their own particular interests and priorities.

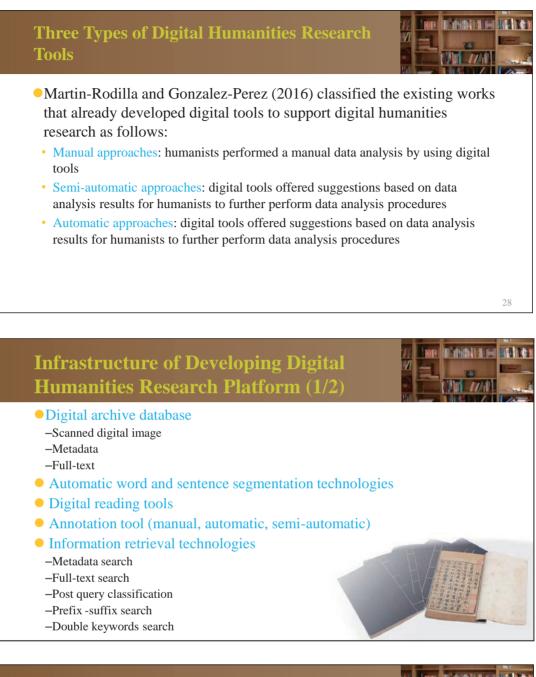


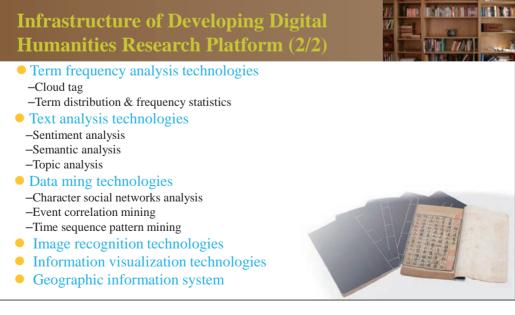
The Used Technologies of Developing Digital Humanities Research Too

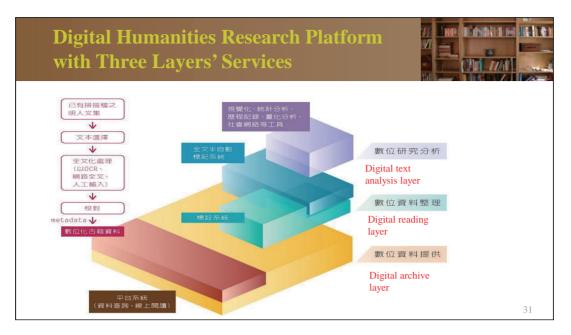
• In the future, DH needs to increase the development and demand in the following fields:

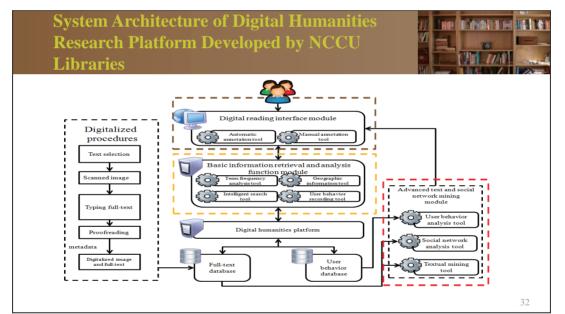
- Computational science
- · Social network theory and analysis
- Digital literacy
- Computational pedagogy
- Computational literacy
- Machine and deep learning
- Neuro and cognitive learning and analysis
- It is obvious that DH is an emerging and interdisciplinary research field aiming at enhancing and redefining traditional humanities scholarship through digital means.

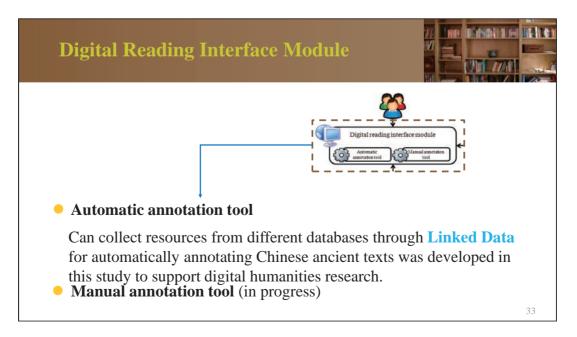
International Workshop for Professional Librarians : Development and Application of Digital Humanities

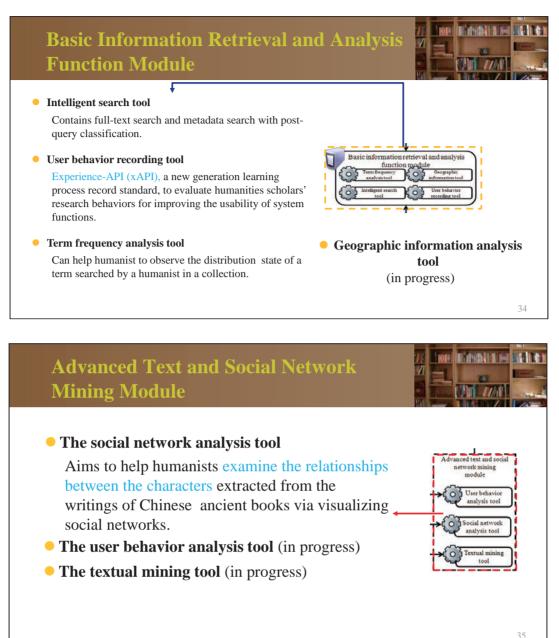












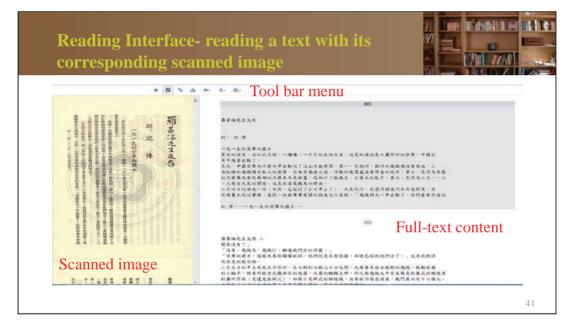


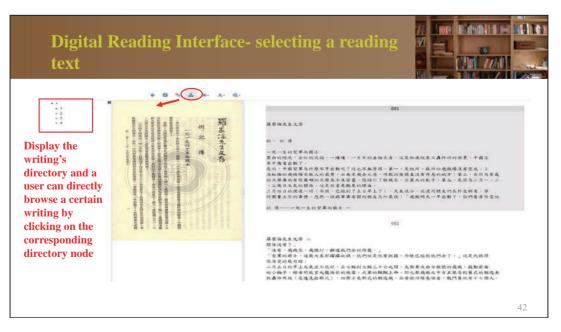


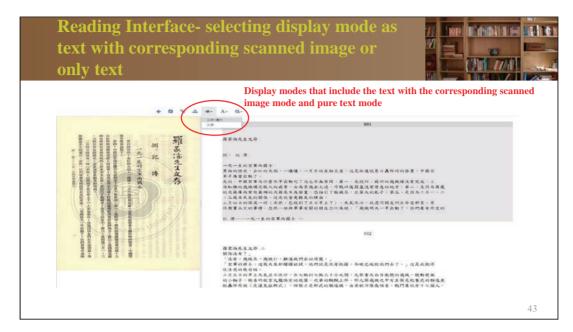


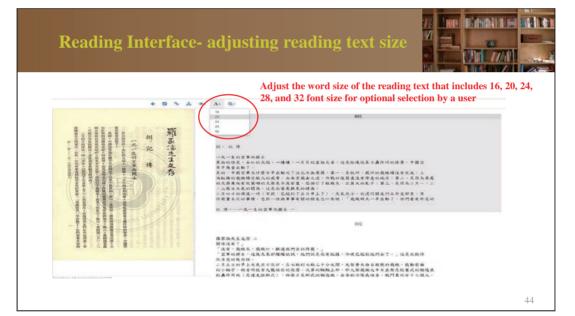


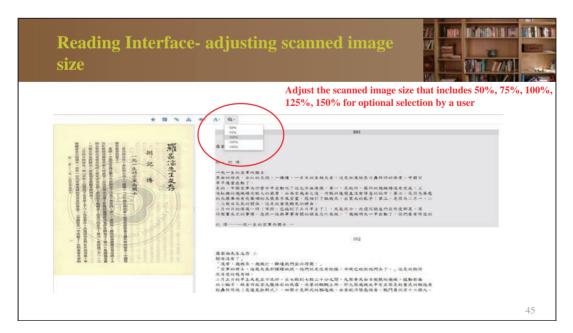


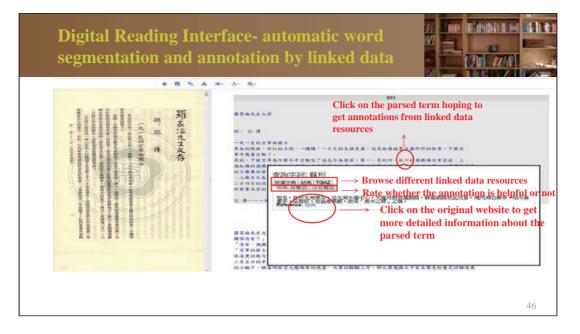


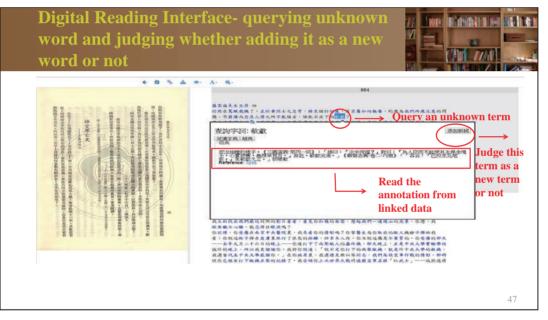












Observation Tool of Viewpoint Migration and Age Division on Digital Humanities Research Platform of Luo Jialun Writings

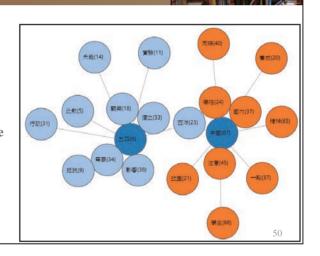
- Digital humanities research platform of Luo Jialun Writings develops an observation tool of viewpoint migration and age division (OTVMAD) for supporting research on digital humanities, allowing humanists dividing texts with time continuity into several time intervals to train word vectors.
- •Word vectors are regarded as the reference for determining highly relational vocabularies with input keywords that a humanist would like to observe.
- Clustering technology is also applied to assist humanists in analyzing whether two neighboring time intervals require split or merge to determine better perspective change observation vocabulary and age division.

如何在學術圖書館發展數位人文研究平台支援數位人文研究 - 國立政治大學圖書館的經驗

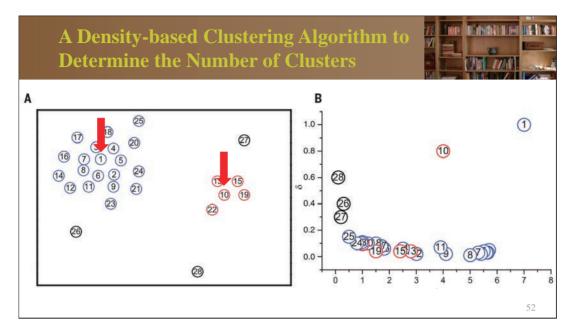


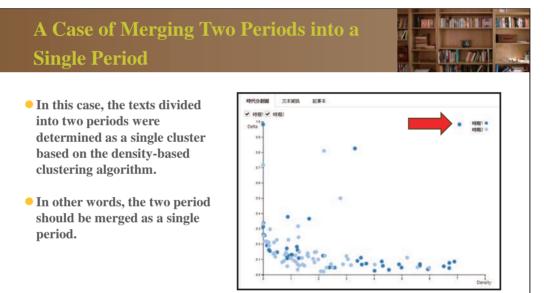
The Term Relationship Map for the Top Ten Most Relevant Terms with the Input Two Keywords

- The blue nodes represent the two input keywords
- The light blue and deep yellow nodes represent the top ten most relevant terms with the keyword "五四" and "中國", respectively.
- The length of the link between the input keyword and a certain relevant term represents the relevance degree. The length is more relevant with the input keyword if the length is more short



Find Out the Texts with the Observed Term		日本 1 (1993), 1, 1] 時期2:[1920, 1, 1]~[1933, 1, 1] ● 五四 ● 中國			
		詞彙	出現文章數	與關鍵詞向量	
• When a user clicks on a term	4	+ 五四	6	該點為主要節點	
on the term relationship map,		+ 影響	38	0.9996081590652466	
the system will show the position that appears the term		+ 観察	18	0.9995607733726501	
in this text for a user's		+ 獨立	33	0.9995565414428711	
reference.		+ 需要	34	0.9995362758636475	
• If a user clicks on the icon +,		+ 西洋	23	0.9998965859413147	
then the system will show the		+ 抵抗	9	0.9995177984237671	
ten characters lying in before and after the term.		+ 此款	5	0.9994977712631226	
		+ 實驗	11	0.9994933009147644	
			21	0.0004005550314331	



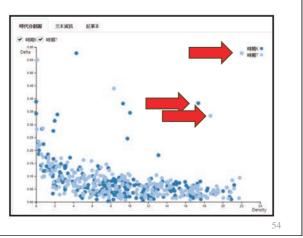


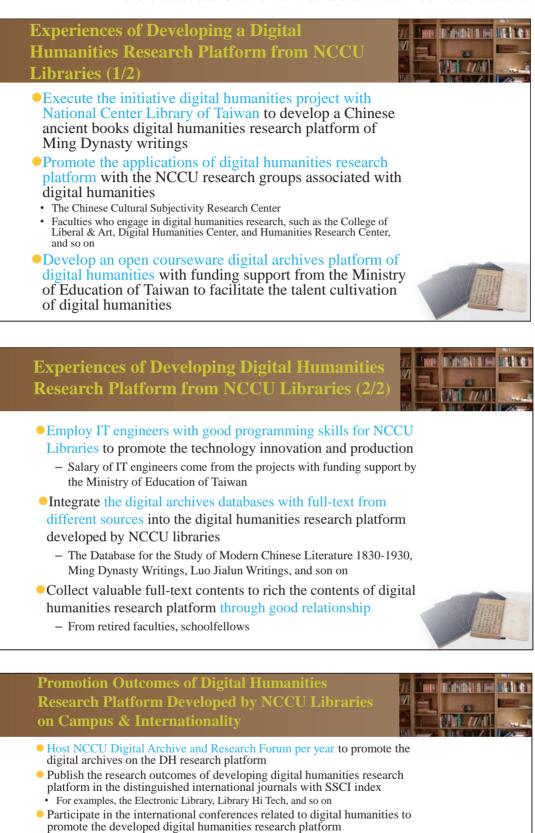
A Case of Dividing Two Periods into Three Periods



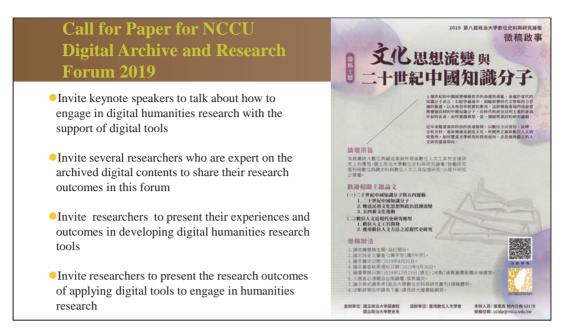
53

- In this case, the texts divided into two periods were determined as three clusters based on the density-based clustering algorithm.
- In other words, the period seven should be divided as two periods. In total, this texts in this case should be determined as three clusters.





- For examples, International Conference on Digital Archives and Digital Humanities, Japanese Association for Digital Humanities
- Regularly cooperate with the research fellows of Chinese Cultural Subjectivity Research Center of NCCU which is a top research center with funding support from the Ministry of Education, Taiwan to engage in the research issues related to the digital archives on the DH research platform since 2017
- Develop an International Sinology Research Center at Dah Hsian Library of NCCU



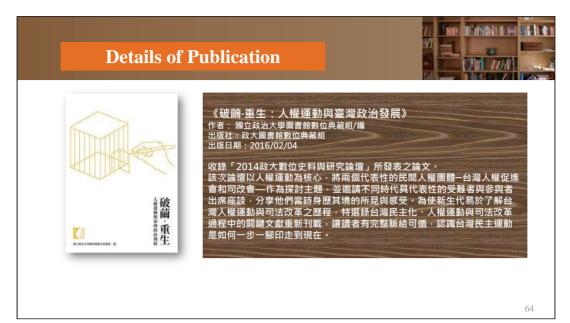
















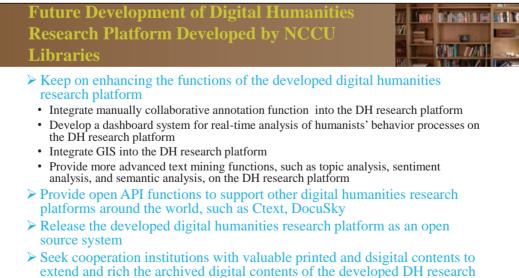






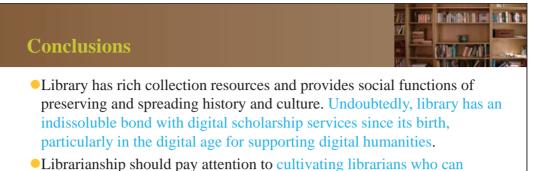






2

platform



- Librarianship should pay attention to cultivating librarians who can support the promotion of digital scholarship and humanities services and assist in the development of digital humanities education in the future.
- Taking the library as the center to drive digital humanities research based on developing and improving the infrastructure of supporting digital humanities research can more specifically highlight the importance of libraries in the era of knowledge economy. 73



Development and Trends of Digital Scholarship Services in Academic Libraries

學術圖書館數位化學術研究服務發展與趨勢

Hao-ren Ke 柯皓仁

University Librarian & Professor Graduate Institute of Library & Information Studies National Taiwan Normal University 國立臺灣師範大學圖書資訊學研究所教授暨圖書館館長

Abstract 摘要

Ushering in the digital era, people are now inhabiting in a digital world. Digital information on the Internet is massively increasing and having significant impacts on people in many areas, such as education, recreation, research and culture. The primary mission of academic libraries lies in supporting teaching, learning, and research via information resources and services. The subject service is one important kind of services offered by research libraries. In recent years, many academic libraries around the word have started the provision of the digital scholarship services. Digital scholarship can be defined as the utilization of digital evidences, investigation methods, research, publishing, and preservation to achieve goals of scholarship and research, and present the specialty of a discipline.

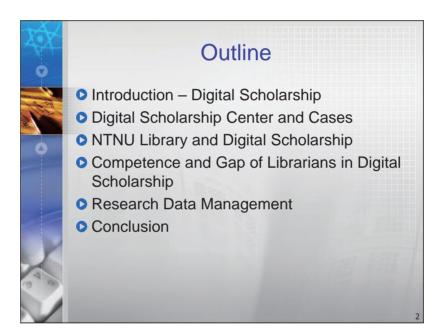
This lecture aims at introducing digital scholarship, presenting the case studies of digital scholarship services in academic libraries, elaborating the services that academic libraries can offer to support digital scholarship, and highlighting the competences and gaps of librarians in offering the digital scholarship services.

數位和網路科技的蓬勃發展已經影響人類生活的許多面向,並改變了資訊和 知識的產生、處理及傳播方式,在教育和學術研究也不例外。大專校院圖書館的 主要任務在於提供資訊資源和服務以支持教學與研究,研究型大學的圖書館多設 有學科服務。近年來,全球許多研究型大專校院的圖書館多開始提供「數位化學

術研究」,所謂「數位化學術研究」,可定義為運用數位化的證據、探究方法、研究、出版、保存來達成學術與研究的目標,並展現領域的專門性。

本次演講旨在介紹數位化學術研究的定義、學術圖書館建立數位化學術研究 中心的案例,學術圖書館對學者從事數位化學術研究所能提供的支持與服務,以 及學術圖書館學科館員在提供數位化學術研究服務上所需的知能與能力落差。









Digitalization of content
Networking of peers and content
Openness
Technology – open source software, API, standard
Content – open access, OER, open data
Mindset – sharing of thought, research materials/data, discussion

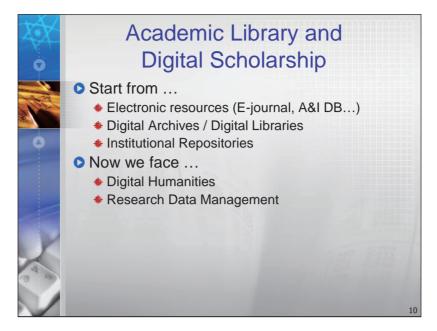




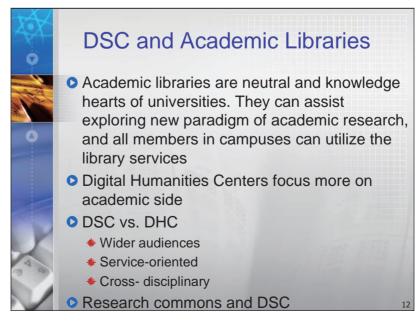




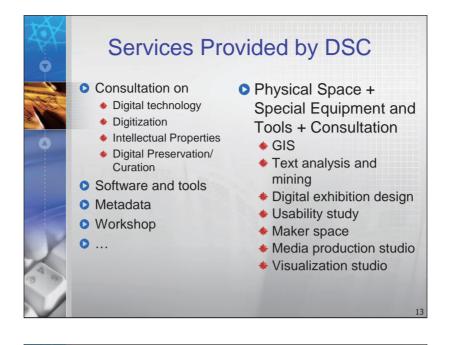
International Workshop for Professional Librarians : Development and Application of Digital Humanities

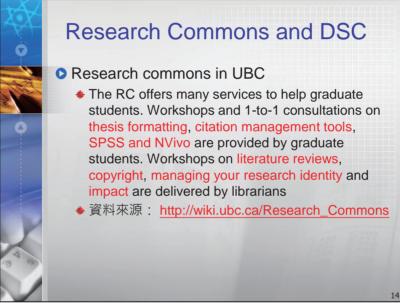






Development and Trends of Digital Scholarship Services in Academic Libraries







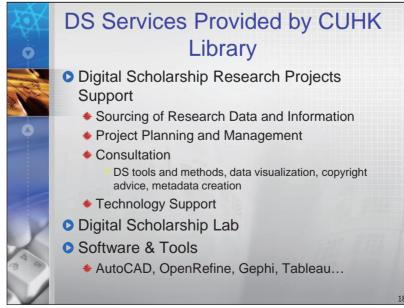


CUHK's Definition of DS

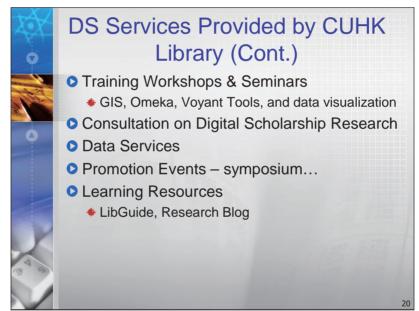
 At CUHK, the Library intends to support the use and reuse of digital evidence, digital tools, methods, and processes to help faculty, researchers and postgraduate students to accomplish their scholarly research

 In collaboration with faculty and researchers, we provide services, online platform, cuttingedge space, tools, training, consultation and advice to support and promote digital scholarship research at the University through the entire research life cycle

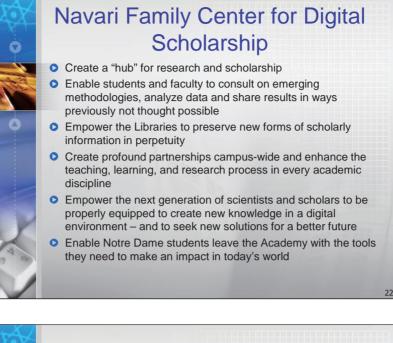


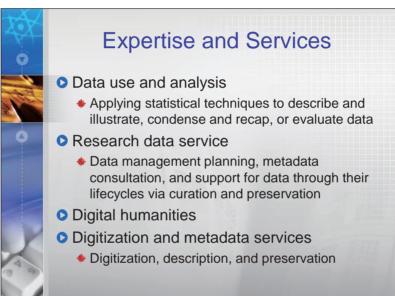








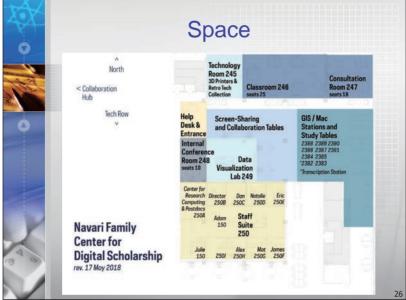






Development and Trends of Digital Scholarship Services in Academic Libraries

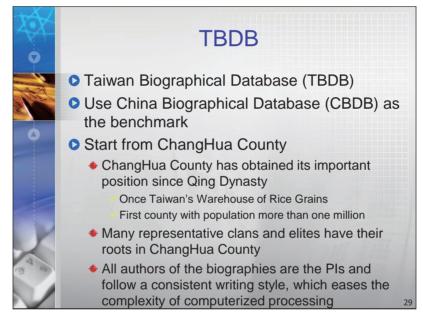


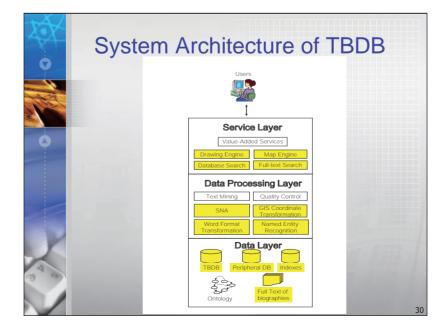




International Workshop for Professional Librarians : Development and Application of Digital Humanities

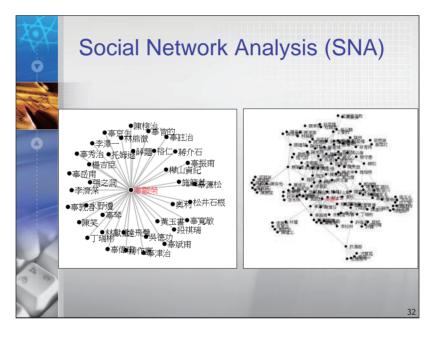


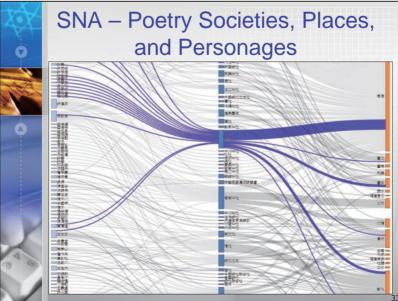




Development and Trends of Digital Scholarship Services in Academic Libraries

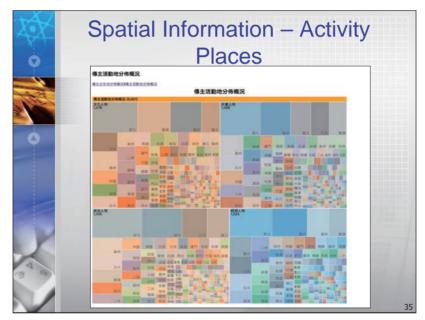
	50 SI				總筆	數:68筆			
71/14.	総計人間(19) 又化人物(18) 政治人物(18)	姓名	出生地	寧順發生地	MBI	生卒年	查看		
	社會人物(13)	黃天	演州	溪州	政治人物	1908-1950	檢視傳文	检视21群	GDF
	出生地	莱旭	花壇	花壇	經濟人物	1873-1952	检视博文	检视社科	GDF
0	間道(10) 1時均(10) 第5代(9)	萬服	線西	绿西	經濟人物	1908-1978	檢視轉文	檢視社群	GDF
	線西(7) 環語(6)	黃秋	鹿港	鹿港	經濟人物	1887-1959	檢視傳文	极视和群	GDF
	花場(5) 田堤(5) 大村(3)	黃龍	12 講	竹環	經濟人物	1865-1932	极很缘文	檢視社群	GDF
	竹環(2) 税費(1)	两周	和美	和英	文化人物	1899-1957	檢視傳文	檢視社群	GDF
	馬水 (1) 百昧 (1)	丙冊	田尾	田尾	文化人物	1902-1976	檢視傳文	檢視社群	GDF
	編用(1) 送州(1) 送肥(1)	黃況	線西	線西	社會人物	1895-1979	檢視佛文	檢視社群	GDF
2	定与 (1) 田中 (1)	黃石柱	花壇	花壇	政治人物	1922-1980	檢視傳文	檢視社群	GDF
r	急調(1) 福建鑽江(1)	黃合益	大村	大村	政治人物	1900-1974	檢視傳文	檢視社群	GDF
	調帶損害(1)								





International Workshop for Professional Librarians : Development and Application of Digital Humanities











Competence related to the subject/discipline

- 1-1 Deep knowledge of their discipline/subject
- 1-2 Excellent knowledge of content (in all relevant media) available to their discipline/subject
- 1-3 Excellent knowledge of bibliographic and other finding tools in the discipline/subject
- 1-4 Knowledge to advise on relevant archive and special collections locally and elsewhere



Competence related to the research Process

- 2-1 Awareness of current and changing local research interests
- 2-2 Understanding of a typical researcher's experience, including their workflow, and how researchers access and use information, within a discipline/subject and at different stages of the researcher's career
- 2-3 Knowledge of sources of research funding to assist researchers to identify potential funders
- 2-4 Ability to gain an appreciation of individual researcher/project needs, including effective listening skills

Competence related to partnerships

- 3-1 Skills to build strong relationships with researchers and other professionals
- 3-2 Skills to establish collaborative partnerships externally
- 3-3 Skills to manage client relationships
- 3-4 Awareness and ability to recognise the value of services and opportunities provided by national and international collaborative initiatives and agencies
- 3-5 Skills to participate effectively in research projects, including identifying a role for the library in the project, and assisting with bid and report writing
- 3-6 Ability to proactively advise and market appropriate library services to researchers

Competence related to Information Literacy

- 4-1 Outstanding skills in information discovery, literature searching etc.
- 4-2 Ability to synthesize, analyze and provide digests of, 'discovered information'
- 4-3 Knowledge to advise on citing and referencing, and the use of bibliographic management software
- 4-4 Ability to assist individual project to conduct literature analysis and provide SDI service
- 4-5 Excellent skills to design information literacy training (both face to face and online) to meet the identified needs of different types of researchers

Competence related to research data

 5-1 Good knowledge of data sources available in the discipline/subject

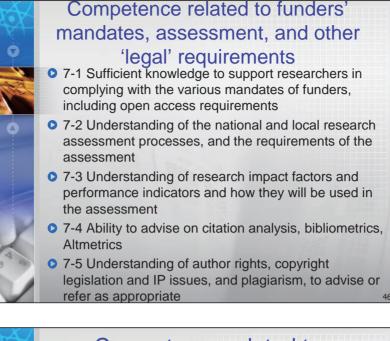
- 5-2 Knowledge to advise on data management and curation, including ingest, discovery, access, dissemination, preservation, and portability
- 5-3 Knowledge to advise on the management of researchers' information, such as reference management
- 5-4 Knowledge to advise on potential data manipulation and visualization tools used in the discipline/subject

Competence related to scholarly communication

- 6-1 Understanding of scholarly communication cycle and the resources and supports that the library can provide in each phase of the cycle
- 6-2 Ability to advise on current trends, best practice and available options in research publication and dissemination methods and models nationally and internationally, including scholarly communication and open access publishing
- 6-3 Ability to advise on current trends, best practice and available options in research publication and dissemination methods and models nationally and internationally, including scholarly communication and open access publishing with a variety of communication methods (face-to-face, FB, Youtube...)

Competence related to scholarly communication (Cont.)

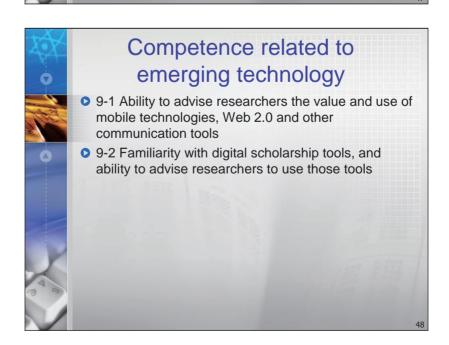
- 6-4 Ability to advise on the quality of journals and other academic publication (such as advise on predatory journals/conferences)
- 6-5 Ability to advise on how to increase research visibility and impact
- 6-6 Ability to assist researchers to evaluate their research visibility and impact
- 6-7 Ability to advise on preserving research outputs
- 6-8 Ability to advise on the preservation of project records e.g. correspondence



Competence related to metadata

 8-1 Knowledge to advocate, and advise on, the use of metadata

 8-2 Skills to develop metadata schema, and advise on discipline/subject standards and practices, for individual research projects



	Gap – Subject/Discipline										
	專業知能	Variable	М	STDEV	р	Gap					
1 miles	1-1具備學科/主題的深	Importance	4.08	.662	.000	.96					
•	入知識	Competence	3.12	.812	.000	.90					
	1-2精熟學科/主題相關 之各種媒體類型的資 訊資源	重要程度	4.22	.549	.000	1.00					
		具備程度	3.22	.776	.000	1.00					
		重要程度	4.24	.610	.000	.95					
2	的書目及查檢工具	具備程度	3.29	.828	.000	.90					
X	1-4能對機構內外、國 內/國際與學科/主題相	重要程度	4.04	.615	.000	1.11					
3 9	關的檔案和特藏提供 建議	具備程度	2.93	.866	.000	1.11					

	Gap – Research Process									
	專業知能	Variable	М	STDEV	р	Gap				
1 ton		Importance	4.03	.681	.000	1.10				
	前和未來的研究熱點	Competence	2.93	.836	.000	1.10				
	2-2 瞭解不同學科/主題、 處於不同職涯階段之研究 人員的經驗・包括他們的 工作流程・以及研究人員 如何取得和使用資源	重要程度	4.09	.675	.000					
		具備程度	2.91	.892		1.18				
	2-3 瞭解研究資金來源·以 協助研究人員識別潛在的	重要程度	3.60	.850	.000	1.19				
Y	協助 研究 入員 識	具備程度	2.41	1.006	.000	1.19				
A	2-4 體認個別研究人員/研	重要程度	3.91	.739	000	1.20				
X 520 1	究計畫的需求	具備程度	2.71	.962	.000	1.20				

	Gap –	Partne	ers	hip		
R. Mark	專業知能	Variable	М	STDEV	р	Gap
•	3-1 能和研究者及其他專業人	Importance	4.08	.619	.000	1.05
	員建立穩固關係	Competence	3.03	1.029		1.05
	3-2 能建立外部合作關係	重要程度	4.02	.648	.000	1.07
	3-2 肥建立介配口下開床	具備程度	2.95	.986		1.07
1		重要程度	3.99	.667	.000	.97
	5-5 共闻自注音广闸际的汉船	具備程度	3.02	.954	.000	.97
40						
						51

*	Gap – Pai	rtnersh	nip	(Cor	nt.)	
1/145-	專業知能	Variable	М	STDEV	р	Gap
R.	3-4 能意識並體認到館際合作	Importance	4.11	.623		
•	相關計畫或機構所提供服務和 機會的價值,如中華民國圖書 館學會、中華圖書資訊館際合 作協會、國際圖書館協會聯盟 等	Competence	3.39	.943	.000	.72
	3-5 具備有效參與研究計畫的	重要程度	3.89	.715		
	技能,包括確定圖書館在研究 計畫中的角色,以及協助爭取 計畫和撰寫報告	具備程度	2.84	1.022	.000	1.05
	3-6 能主動給予研究人員建議	重要程度	4.16	.623	000	00
	並推薦其適合的圖書館服務	具備程度	3.20	.933	.000	.96
A B						52

•	Gap – Information Literacy											
	專業知能	Variable	М	STDEV	р	Gap						
1100-		Importance	4.39	.546	000	70						
0	尋的相關技能	Competence	3.66	.798	.000	.73						
	4-2 能對查找到的資訊進行	重要程度	4.25	.600	.000	.91						
	綜整、分析並提供摘要報告	具備程度	3.34	.869	.000	.91						
	4-3 具備給予引用與參照建	重要程度	4.21	.607	.000	70						
1	議・以及使用書目管理軟體 的知識	具備程度	3.45	.842	.000	.76						
	<mark>4-4</mark> 能協助個別研究計畫進 行文獻分析與提供新知通報	重要程度	4.08	.639	000	02						
X	们又刷分析與提供新知通報 服務	具備程度	3.15	.901	.000	.93						
A	4-5 精熟設計資訊素養訓練(重要程度	4.16	.636	000	05						
	面對面或線上)的技能·以 滿足研究人員的需求	具備程度	3.21	.959	.000	.95						

XX	Gap – Research Data Managent										
	專業知能	Variable	м	STDEV	р	Gap					
81100-	5-1 精熟學科/主題相關的研究	Importance	4.20	.583	.000	1.09					
0	資料來源	Competence	3.11	.873	.000						
	5-2 能為研究資料管理和庋用 提供建議,包括攝取、發現、 取用、傳播、分享、保存和可 移植性	重要程度	4.07	.608	000	1.10					
		具備程度	2.98	.924	.000	1.10					
	5-3 能對研究人員之資訊管理(重要程度	4.07	.648	.000	.95					
	如參考文獻的管理)提供建議	具備程度	3.12	.930	.000	.95					
	5-4 能提供操作和呈現研究資	重要程度	3.99	.678	000	1.00					
	訊的相關建議	具備程度	2.97	.959	.000	1.02					
V						54					

	Gap – Schola	rly Cor	nm	unica	atio	n
1 and	專業知能	Variable	М	STDEV	р	Gap
	6-1 瞭解學術傳播生命週期及圖	Importance	3.97	.699		
	書館對該週期中各個階段所能 提供的資源與支援	Competence	2.90	.878	.000	1.07
9	6-2 能就學術研究出版和傳播方	重要程度	4.00	.660		
	法與模式的當前趨勢、最佳做 法和可能選項提供建議,包括 學術傳播和公開取用	具備程度	2.84	.914	.000	1.16
	6-3 具備運用多元管道(如面對	重要程度	4.05	.646		
	面、網頁、社群網路、部落格 影音平臺)對研究人員倡導學術 研究出版和傳播方法與模式的 當前趨勢、最佳做法和可能選 項,包括學術傳播和公開取用	具備程度	2.98	.900	.000	1.07
A B				112		55

***	Gap – Scholarly Communication (Cont.)									
	專業知能	Variable	М	STDEV	р	Gap				
1/100	6-4 能對協助研究人員評估期刊和	Importance	4.11	.662						
	其他學術出版品的學術品質(如協助 判斷掠奪型期刊、掠奪型研討會)	Competence	2.84	.994	.000	1.27				
	6-5 能對提高研究人員能見度與影	重要程度	3.99	.656	.000	1.25				
	響力的方法提出建議	具備程度	2.74	.986	.000	1.20				
	6-6 能協助研究人員進行研究生產		3.94	.653	.000	1.30				
	力與影響力評估(如資訊計量分析)	具備程度	2.64	.955	.000	1.50				
	6-7 能就研究成果的保存提出建議	重要程度	3.88 2.76	.707 1.001	.000	1.12				
12	6-8 能就保存研究計畫的相關紀錄(重要程度	3.79	.725	000	1.20				
	如實驗室筆記、通訊紀錄)提出建議		2.59	.986	.000	1.20				
1 (D				112	120	56				

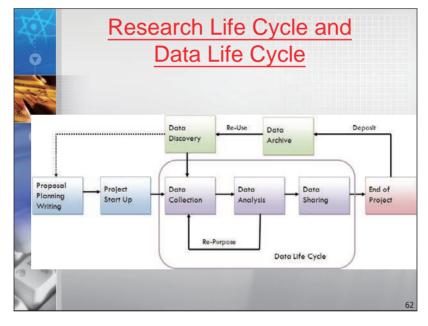
XX •	Gap – funders' mandates, assessment, and other 'legal' requirements									
	專業知能	Variable	М	STDEV	р	Gap				
1 Sar	7-1 具備足夠的知識,以支持研 究人員符合資助者的各種規定,	Importance	3.91	.690	.000	1.24				
2	包括公開取用的要求	Competence	2.67	.945	.000	1.24				
0	7-2 瞭解國家和機構內部進行研究人員評鑑的過程,以及研究人員評鑑的過程,以及研究人員評鑑的要求	重要程度	3.81	.731	.000	1.25				
		具備程度	2.56	.915	.000	1.20				
	7-3 瞭解研究影響指數和評鑑指 標 · 以及該等指數或指標如何用 在研究評鑑	重要程度	3.95	.684	.000	1.21				
		具備程度	2.74	.929	.000	1.21				
	7-4 能就引文分析、資訊計量、	重要程度	3.93	.667	.000	1.19				
	替代計量提供建議	具備程度	2.74	.890						
	7-5 瞭解作者權利、智慧財產權 法、知識產權、抄襲等議題,以	重要程度	4.08	.677	.000	1.02				
A G	便適當提出建議或轉介	具備程度	3.06	.840	.000	1.02				
						57				

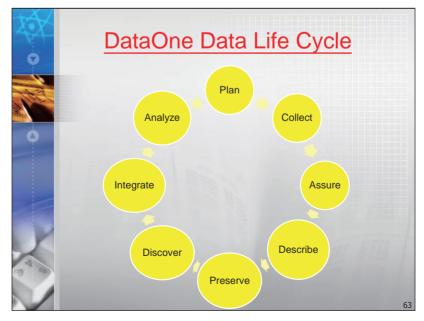
	Gap -	- Meta	dat	а		
1 1 and a	專業知能	Variable	М	STDEV	р	Gap
	8-1 能宣導詮釋資料的使用並 提出建議	Importance	3.81	.744	.000	1.04
		Competence	2.77	.911		1.04
	<mark>8-2</mark> 能為個別研究計畫開發詮 釋資料綱要・以及對學科/主	重要程度	3.74	737	.000	1.15
	題範圍內的詮釋資料標準和 實務提供建議	具備程度	2.59	.935	.000	1.15
A B						58

to a	Gap – Emerg	ging To	ecł	nol	ogy	,
1110	專業知能	Variable	М	STDEV	р	Gap
	9-1 能向研究人員提供關於行動 科技、Web 2.0和其他社群網路 工具(如ResearchGate)的價值 和使用建議	Importance	3.86	.693	.000	1.02
		Competence	2.84	.880	.000	
	9-2 精熟數位化研究工具 (如資 料處理與探勘、資料視覺化、社	重要程度	3.89	.720	.000	1.35
X	群網路分析、地理資訊系統) · 並能對研究人員提出建議	具備程度	2.54	.949	.000	1.00
						59



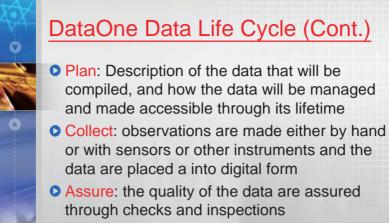






International Workshop for Professional Librarians : Development and Application of Digital Humanities

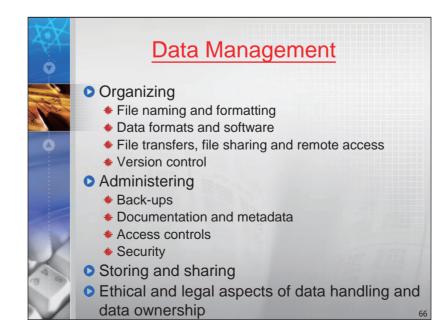
2019 國際圖書館專業館員研習班:數位人文的發展與應用



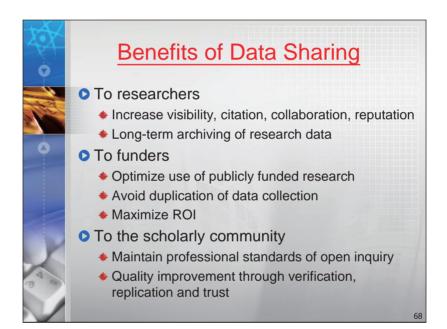
 Describe: data are accurately and thoroughly described using the appropriate metadata standards

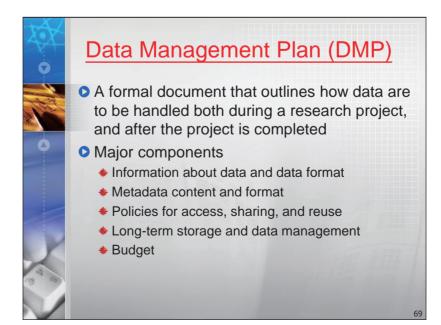
DataOne Data Life Cycle (Cont.)

- Preserve: data are submitted to an appropriate long-term archive (i.e. data center)
- Discover: potentially useful data are located and obtained, along with the relevant information about the data (metadata)
- Integrate: data from disparate sources are combined to form one homogeneous set of data that can be readily analyzed
- Analyze: data are analyzed



Benefits of Good Data Management Efficiency – makes your own research easier Safety – protects valuable data Quality – better research data leads to better research Reputation – enhances research visibility Compliance – with ethical codes, data protection laws, journal requirements, funder policies





International Workshop for Professional Librarians : Development and Application of Digital Humanities

2019 國際圖書館專業館員研習班:數位人文的發展與應用

Taiwanese Scholars' Perception of Research Data Management

- Many academic libraries have established institutional repositories (IRs)
- No mandates on open access, data management plan, and open data
- Difficult for academic libraries to acquire (post-print) articles for IRs, not to mention research data

No previous studies for understanding scholars' perception of managing and sharing of research data

Research Questions

- What are the sources of research data used in research? ?
- What are the types of research data used and produced in research?
- What are scholars' perception of and behavior in managing research data?
- What are scholars' perception of and behavior in sharing and reusing research data?





Two Major Sources of Questionnaire Items

 Chowdhury, G., Boustany, J., Kurbanoglu, S., Unal, Y., & Walton, G. (2017). Preparedness for research data sharing: A study of university researchers in three European countries. *Proceedings* of 19th International Conference on Asia-Pacific Digital Libraries, ICADL 2017, Bangkok, Thailand, November 13-15, 2017, 104-116.

 Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., Manoff, M., & Frame, M. (2011). Data sharing by scientists: Practices and Perceptions. *PLoS ONE*, 6(6), 1-21.

1000		Dem	ogra	phic An	aly		Percent
and		_	_			Responses	
1 Prais	Ages	Responses	Percent	Professor/Fel	low	454	42.5%
S.	26-35	47	4.4%	Associate		347	32.5%
	36-45	314	29.1%	Professor/Fel	low		
Q	46-55	421	39.0%	Assistant		253	23.7%
	56-65	278	25.8%	Professor/Fel	low		
	Over 66	19	1.8%	Postdoctoral Fellow/Lectur		15	1.4%
	Subject Disciplines Social Sciences Biology					2	ercent 2.3% 2.1%
		ring and Techr	nologies		234 209		9.7%
	Humanities				157		4.8%
15	Natural Sciences and Sustainable Development				134	1	2.7%
3 4	Management					5	.7%
		ic Education			29	2	.7%
							74

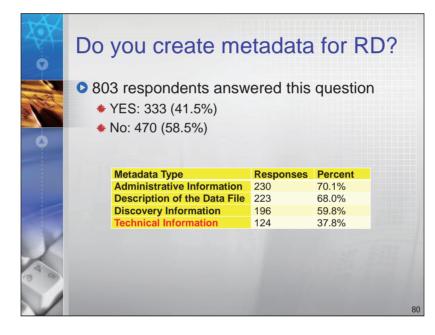
	Responses	Percent
Under 5 Years	54	5.0%
6-10 Years	152	14.1%
11-15 Years	236	21.9%
16-20 Years	192	17.8%
21-25 Years	191	17.7%
26-30 Years	145	13.4%
Over 30 Years	109	10.1%
Time Spent on Research	Responses	Percent
0%-20%	61	5.7%
21%-40%	324	30.0%
41%-60%	384	35.6%
<u>61%-80%</u>	216	20.0%
81%-100%	94	8.7%

		Data	Sour	ces		
Data Source	Never	Rarely	Occasionally	Frequently	Very Frequently	Average
Create New Data	34 (3.6%)	129 (13.5%)	171 (17.9%)	335 (35.1%)	285 (29.9%)	3.74
From Same Teams	34 (3.5%)	123 (12.6%)	195 (19.9%)	402 (41.1%)	225 (23.0%)	3.68
Through scholarly Communication Networks	82 (8.9%)	216 (23.5%)	264 (28.7%)	258 (28.0%)	101 (11.0%)	3.09
From Research Data Repositories	114 (11.6%)	93 (9.5%)	128 (13.0%)	354 (36.0%)	294 (29.9%)	3.63
4						76

	Types of Da Used and Proc					
Data Types (Multiple Choices)	Used Responses(Percent)/Rank	Produced Responses(Percent)/Rank				
Experimental	553 (52.1%)/1	538 (51.4%)/1				
Social Science Surveys	341 (32.1%)/2	287 (27.4%)/2				
Bibliographic/Citation Data	336 (31.6%)/3	193 (18.5%)/8				
Observational	312 (29.4%)/4	241 (23.0%)/5				
Interviews	305 (28.7%)/5	270 (25.8%)/3				
Text	301 (28.3%)/6	239 (22.8%)/6				
Computer Simulation	285 (26.8%)/7	263 (25.1%)/4				
Research-Related Computer Programs	276 (26.0%)/8	203 (19.4%)/7				
Data Models	164 (15.4%)/9	187 (17.9%)/9				
Biotic Surveys	72 (6.8%)/10	59 (5.6%)/10				
Log Data	72 (6.8%)/11	48 (4.6%)/11				
Remote-Sensed Abiotic	50 (4.7%)/12	22 (2.1%)/13				
Abiotic Surveys	42 (4.0%)/13	26 (2.5%)/12				
Remote-Sensed Biotic	7 (0.7%)/14	8 (0.8%)/14				
		77				

X	Sto	rage	and M	Manag	gemer	nt of F	RD
	I am satisfied with the process for	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Ave. SD.
No.	storing my data during the life of the project	4 (0.4%)	9 (1.3%)	85 (8.4%)	407 (38.8%)	517 (51.1%)	4.39 0.71
-	storing my data beyond the life of the project	8 (0.8%)	27 (2.6%)	133 (13.0%)	434 (42.5%)	419 (41.0%)	4.20 0.82
	search for my own data during the life of the project	5 (0.5%)	37 (3.7%)	163 (16.3%)	469 (46.9%)	326 (32.6%)	4.07 0.82
	search for my own data beyond the life of the project	5 (0.5%)	34 (3.4%)	198 (20.0%)	480 (48.4%)	275 (27.7%)	3.99 0.81
1	creating metadata for my data	5 (0.5%)	36 (3.8%)	240 (25.3%)	485 (51.1%)	184 (19.4%)	3.85 0.79
							78

	Where to Store	RD	
	I store my data in (multiple choices)	Responses	Percent
2-14.	my storage devices	1030	96.3%
11.45-	cloud drives	600	56.1%
	the data repository created by my institution	219	20.5%
0	the data repository commonly used by my subject discipline	153	14.3%
	In the long term, my data should be stored in (multiple choices)	Responses	Percent
	my storage devices	841	78.7%
1	cloud drives	670	62.7%
	the data repository created or endorsed by my institution	362	33.9%
1	the data repository commonly used by my subject discipline	372	34.8%
4 3			



 About 83 	 Requirement / Mandates About 830-900 respondents answered this series of questions 									
DMP Requirement	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Ave SD				
Your major funding agencies have a DMP requirement	75 (8.8%)	205 (23.9%)	211 (24.6%)	282 (32.9%)	84 (9.8%)	3.11 1.14				
Your major funding agencies should have a DMP requirement	52 (5.8%)	103 (11.5%)	252 (28.2%)	364 (40.8%)	122 (13.7%)	<mark>3.45</mark> 1.05				
Your institution have a DMP requirement	88 (10.6%)	241 (29.0%)	213 (25.7%)	222 (26.7%)	66 (8.0%)	2.92 1.14				
Your institution should have a DMP requirement	58 (6.8%)	144 (16.9%)	243 (28.5%)	312 (36.6%)	96 (11.3%)	3.29 1.09				
4.8										
						81				

	Training and Support							
	Formal Training Subjects	Have had a formal Training	Willingness for a Formal Training					
and the	No	741 (69.6%)	NA					
1 noni-	Data Citing Styles	212 (19.9%)	290 (34.0%)					
	Data Management Plan	142 (13.3%)	715 (83.7%)					
	Metadata	134 (12.6%)	375 (43.9%)					
Q I	Consistent File Naming	81 (7.6%)	316 (37.0%)					
	Version Control of Data Sets	76 (7.1%)	340 (39.8%)					
	I need service supports	Responses	Percent					
	and suggestions on							
	Data Management Plan	592	62.8%					
1	Data Storage	440	46.7%					
	Technical Assistance	367	39.0%					
	General Data Management	336	35.7%					
1	Related Law	288	30.6%					
	Research Ethics on Data	286	30.4%					
5	Publication and Citation	264	28.0%					
	Technical Problems	245	26.0%					
			82					

7	Perception of Data Sharing and Reuse						
111	Perception of data sharing and reuse	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Ave SD
Contraction of the second seco	Lack of access to data generated by other researchers or institutions is a major impediment to progress in science	23 (2.3%)	113 (11.2%)	216 (21.4%)	454 (45.0%)	204 (20.2%)	3.7 0.988
	Lack of access to data generated by other researchers or institutions has restricted my ability to answer science questions	27 (2.7%)	171 (17.1%)	268 (26.8%)	392 (39.2%)	141 (14.1%)	3.45 1.017
	I would use other researchers' datasets if their datasets were easily accessible	8 (0.8%)	36 (3.6%)	158 (15.6%)	547 (53.9%)	265 (26.1%)	4.01 0.794
1	I would be willing to share my datasets with other researchers	15 (1.5%)	48 (4.7%)	176 (17.3%)	536 (52.5%)	245 (24.0%)	3.93 0.854
							83

Perception of Data Sharing and Reuse (Cont.)							
Perception of data sharing and reuse	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree	Ave SD	
I would be willing to place at least some of my data into a central data repository with no restrictions		143 (14.2%)	187 (18.6%)	445 (44.1%)	180 (17.9%)	3.56 1.09	
I would be willing to place all of my data into a central data repository with no restrictions	158 (16.0%)	304 (30.7%)	220 (22.2%)	214 (21.6%)	93 (9.4%)	2.78 1.221	
I would be more likely to make my data available if I could place conditions on access	22 (2.2%)	65 (6.5%)	253 (25.2%)	484 (48.2%)	180 (17.9%)	3.73 0.904	
3 3					244	8	

10 A	Data Sharing Pr			
a the	Current Data Sharing Practices	Responses	Percent	
1/100-	My data is available openly upon request	312	30.5%	
R. Contraction	My data is open available only to my team	271	26.5%	
	My data has restricted access	234	22.9%	
	My data is open available to everyone	150	14.6%	
	My data is not available to anyone else	57	5.6%	
	Data is shared (Multiple Choices)	Responses	Percent	
	Through a national data repository	420	41.9%	
	On the PI's Website	406	40.5%	
	On my Institution's Website	395	39.4%	
	On my personal Website	339	33.8%	
	Through a global data repository	321	32.0%	
No. Con	Through a regional data repository	247	24.7%	
40				85

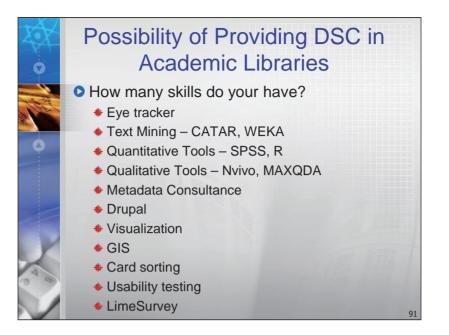
	Data Sharing Con	cerns	
1 har	Data Sharing Concerns (Multiple Choices)	Responses	Percent
	Lack of appropriate policies and rights protection	519	72.7%
Q	Misuse of data	514	72.0%
	Misinterpretation of data	483	67.6%
	Legal and ethical issues	338	47.3%
1	Fear of losing the scientific edge	259	36.3%
	Lack of resources (technical, financial, personnel, storage space, etc.)	207	29.0%
	Nobody needs my shared data	59	8.3%
	Should not share data	32	4.5%
A S			
1			

× •	Data Reus	е	
214	Criteria to reuse other researchers' datasets (multiple choices)	Responses	Percent
2100	Data quality	509	92.0%
(TR)	Data collection process	318	57.5%
~	Potential to publish	308	55.7%
Y	Data up-to-date	268	48.5%
	Data accessibility	232	42.0%
	Reasons (multiple choices)	Responses	Percent
	Expansion of research topics	370	67.0%
	Inability to collect the data by myself	367	66.5%
	Credibility of data from authorities	320	58.0%
	Explore potential research topics	315	57.1%
	Encourage or restriction of discipline cultures	145	26.3%
4	Exemption of IRB liability	67	12.1%

	Fairness of Da	ta Rei	JSe		
3 Miles	Fairness of data reuse		M	SD	р
11/25-	Co-authorship on publications resulting	luse	3.10	1.157	.201
	from use of the data	others use	3.08	1.151	
	Formal acknowledgement of the data	luse	4.39	0.671	.001
•	providers and/or funding agencies in all disseminated work making use of the data		4.36	0.688	
	Formal citation of the data providers	luse	4.49	0.612	.005
	and/or funding agencies in all disseminated work making use of the data		4.46	0.642	
	The opportunity to collaborate on the	luse	4.11	0.697	.071
	project	others use	4.09	0.712	
	At least part of the costs of data	l use	3.16	0.950	.002
A G	acquisition, retrieval, or provision must be recovered	others use	3.12	0.969	

	Fairness of Data R	leuse	(Co	nt.)	
	Fairness of data reuse		M	SD	р
11100-	Reprints of articles that make use of	l use	3.47	0.990	.696
	the data must be provided to the data provider		3.47	1.000	
	The data provider is given a complete	l use	3.36	0.986	.486
	list of all products that make use of the data	others use	3.36	0.996	
	Legal permission for data use is	l use	3.86	0.908	.353
	obtained	others use	3.85	0.927	
	Mutual agreement on reciprocal	l use	3.69	0.916	.406
	sharing of data	others use	3.69	0.944	
	The data provider is given and agrees	l use	3.97	0.839	.503
X	to a statement of uses to which the data will be put	others use	3.98	0.860	
	It is appropriate to create new datasets	l use	3.80	0.826	.071
N G	from the shared datasets	others use	3.82	.840	







Further Information

1. Registration:

09:00-09:30, October 28th, 2019; 08:30-09:00, October 29th~31st

Please go to 3rd Floor, Cluture & Education Zone, National Central Library to collect your badge and conference handbook. Please continue to register for the rest of the days, and collect lunch vouchers.

- 2. All the lectures would be delivered in English. Q&A session will be arranged after the lecture finishes.
- 3. The content of course is protected by copyright act and intellectual property law

-Individual video and audio recording are prohibited

- -Courses that have the authorization of the lectures will be uploaded into the "Digital Audio Visual Archive System" (http://dava.ncl.edu.tw). All are welcome to view online.
- 4. Meals & Refreshment:

-October 28th~31st: Lunch at Conference Room 301 for lectures and international participants; 4th Floor for Taiwanese participants.

-Tea served every day; Fruits & refreshments served in the afternoon.

5. Wi-Fi service:

-Username: meeting

-Password: ncl2018s4

6. Questionnaire Survey:

-At the end of the last session, please fill out the "International Joint Research and Training Program for Development and Application of Digital Humanities" questionnaire survey, and give it to the staff.

- 7. Course Certificates will be issued on October 31st to international participants who attend the whole courses.
- 8. Civil servant learning hours for Taiwanese participants are 32 hours.

注意事項

- -、108年10月28日上午9:00~9:30請至國家圖書館文教區3樓報到,領取國際研習 班識別證及手冊。10月28日至31日上午8:30~9:00,亦請報到並領取餐券。
- 二、所有課程皆以英文進行,於主講人授課結束後,安排Q&A時間。
- 三、課程內容涉及著作財產權,請勿自行錄影、錄音;經講師同意授權錄影之課程,由本館於研習會結束後,將錄影課程上傳本館「數位影音服務系統」(http://dava.ncl.edu.tw),歡迎上線點播觀看。
- 四、膳食供應:
 - (一)10月28日~31日午餐:主講人及國際學員用餐地點安排於301會議室;國內學員用餐地點安排於4樓。
 - (二)會議全日供應茶水,下午提供水果及茶點。
- 五、國際會議廳無線網路名稱:meeting,密碼:ncl2018s4。
- 六、研習課程結束後,國際學員請填寫「數位人文的發展與應用國際圖書館專業館員研 習班」滿意度調查表,並交給現場工作人員。
- 七、全程參與課程之國際學員,本館於10月31日核發研習證書。
- 八、全程參與課成之國內學員,本館核發32小時之終身學習時數。

List of International Participants 國際學員名單

#	Name	Position	Institution
	Denmark 丹麥		
1	BLOMQVIST, Inga-Lill	Librarian	NIAS-Nordic Institute of Asian Studies
	France 法國		
2	LAU, Dat-Wei 劉達威	librarian in charge of the chinese collections	Ecole Française d'extrême-orient 法國遠東學院
	Germany 德國		
3	TABERY, Thomas	Curator of the East Asian Collection	Bavarian State Library
	Hong Kong 香港		
4	CHAN, Yuen Yee	Special Collections Librarian	The Chinese University of Hong Kong
5	CHU, Chun Wang	Librarian	Vocational Training Council
6	CHUI, Wai Lung	Computer Officer	The Education University of Hong Kong
7	HO, Kai Sum	Assistant Librarian	Education University of Hong Kong
8	KO, Angela 高玉華	Assistant Librarian	Fung Ping Shan Library, University of Hong Kong
9	LAU,Ching Tin	Library Assistant	The University of Hong Kong Libraries
10	LEE, Chi Hang (Ryun) 李智恒	Head of Cataloguing	The Chinese University of Hong Kong
11	LEE, Yuet Yu	Teacher Librarian	Tak Nga Primary School
12	LI, Cheuk Wai	Assistant Librarian	The Education University of Hong Kong
13	LUM Vincent	Digital Technologies Librarian	Chinese University of HK Library
14	LUNG, Sze Man	Assistant Librarian	VTC, Hong Kong
15	MAHBOOBANI,	Library Assistant II	The University of Hong Kong
	Vanessa Ramesh		
16	MO, JingYi	Student	HKU SPACE
17	NG, Ka Ling	Library assistant II	Chinese University of Hong Kong
18	SHEA, Yuen Ying	Student	HKU SPACE
19	TAM, Man Lik(Owen)	Librarian	The Open University of Hong Kong
	譚文力	圖書館長	香港公開大學
20	TSANG, Lau Foon	University Archivist	The Chinese University of Hong Kong
			Library
21	TSUI, Diana	Arts and Architecture Faculties Librarian	The University of Hong Kong Libraries

International Workshop for Professional Librarians : Development and Application of Digital Humanities

#	Name	Position	Institution
22	WONG, Sin Ting	Library Assistant II	The Education University of Hong Kong
	(Paula)		
23	WONG, Shun Han	Senior Assistant Librarian	Hong Kong Baptist University
	(Rebekah)		
24	YU, Ling Ling	Senior Programme Director	HKU SPACE
25	唐錦亮	Library Assistant	HKBU Library
26	陳偉明	Cataloguing Librarian	Fung Ping Shan Library, University of Hong Kong
27	黎浩賢	Library Assistant	Caritas Institute of Higher Education
	Japan 日本		
28	YAMANAKA Setsuko	Head of Academic Support Division	Kyoto University Library
	Netherlands 荷蘭		
29	GILBERT, Marc	Subject Librarian for Chinese studies Curator of Chinese special collections	Leiden University Libraries
	Philippines 菲律賓	•	
30	ABERASTURI,	Library Assistant	Paranaque City Public Library
	ABERASTURI, Ma. Cristina		
	ABERASTURI,	Library Assistant	Paranaque City Public Library University of the Philippines
31	ABERASTURI, Ma. Cristina BAGUNU,	Library Assistant	
31 32	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony	Library Assistant Head Librarian	University of the Philippines De La Salle Medical & Health Sciences
31 32	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES,	Library Assistant Head Librarian Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute
31 32 33	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony	Library Assistant Head Librarian Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences
3132333334	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES, Bryan Paolo ISUGA, Angelo Luis	Library Assistant Head Librarian Librarian Library staff	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences Institute
3132333435	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES, Bryan Paolo ISUGA, Angelo Luis MARIFOSQUE, Myzel	Library Assistant Head Librarian Librarian Library staff Librarian Head Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences Institute University of the Philippines Diliman
31 32 33 34 35 36	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES, Bryan Paolo ISUGA, Angelo Luis	Library Assistant Head Librarian Librarian Library staff Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences Institute University of the Philippines Diliman National Library of the Philippines
31 32 33 34 35 36 37	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES, Bryan Paolo ISUGA, Angelo Luis MARIFOSQUE, Myzel RAMIREZ, Melanie	Library Assistant Head Librarian Librarian Library staff Librarian Head Librarian Reference Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences Institute University of the Philippines Diliman
31 32 33 34 35 36 37	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES, Bryan Paolo ISUGA, Angelo Luis MARIFOSQUE, Myzel RAMIREZ, Melanie RICARDO Cissette	Library Assistant Head Librarian Librarian Library staff Librarian Head Librarian Reference Librarian Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences Institute University of the Philippines Diliman National Library of the Philippines Paranaque City Public Library
 30 31 32 33 34 35 36 37 38 39 	ABERASTURI, Ma. Cristina BAGUNU, Jessie Rose BAHIA, Anthony BRILLANTES, Bryan Paolo ISUGA, Angelo Luis MARIFOSQUE, Myzel RAMIREZ, Melanie RICARDO Cissette TORRES, Ethel	Library Assistant Head Librarian Librarian Library staff Librarian Head Librarian Reference Librarian Librarian	University of the Philippines De La Salle Medical & Health Sciences Institute De La Salle Medical and Health Sciences Institute De La Salle Medical and Health Sciences Institute University of the Philippines Diliman National Library of the Philippines Paranaque City Public Library

List of International Participants 國際學員名單

#	Name	Position	Institution
	Thailand 泰國		
41	WEEWORAWONG, Warawan	Head of Thailand Information Subdivision	Chulalongkorn University , Office of Academic Resources
	U.S.A. 美國		
42	CHOW, Susan	Metadata Librarian	Getty Research Institute
43	NI, Dongyun	Dept. Chair & Chinese Stuides Librarian	University of Hawaii at Manoa
44	SU, Meng-fen	East Asian Studies Liaison Librarian	University of Texas Libraries
45	TU, Feng-en	Fung Fellow	Harvard University
46	WANG, Qiuli	Instructional Media Specialist	Colby College
47	XUE, Susan	Head of Public Services Electronic Resources Librarian	University of California, Berkeley
48	ZHANG, Ying	Research Librarian for Asian Studies	University of California, Irvine
49	蔣樹勇	Associate professor, Chinese studies librarian	University of Illinois at Urbana-champaign
	Vietnam 越南		
50	HOANG, Phung Mai 黃芳梅	Deputy Librarian	Institute of Hán-Nôm Studies
51	HUONG, Ngo Thi Thanh	Head of Technical Dept	VNUHCM; University of Technology
52	NGO, Huyen Thi	Head of Department	University of Social Sciences and
		of Information and Bibliography	Humanities, Vietnam National University - Ho Chi Minh City

International Joint Research and Training Program for Development and Application of Digital Humanities

Issuer	Tseng Shu-hsien
Publisher	National Central Library
	No.20, Zhongshan South Road,
	Taipei 10001, Taiwan , R.O.C.
	+886-2-23619132 https://www.ncl.edu.tw/
	https://www.ncl.edu.tw/
Editor	International Cooperation Division,
	National Central Library
Printing	Art Co., Ltd.
	No. 8, Lane 315, Section 2, Zhongshan Road,
	Zhonghe District, New Taipei City 235, Taiwan , R.O.C.
	+886-2-22477615
Publication Date	October, 2019

數位人文的發展與應用研習會

發行人	曾淑賢
出版人	國家圖書館
	10001 臺北市中山南路 20 號
	02-2361-9132(代表號)
	https://www.ncl.edu.tw/
編輯	國家圖書館國際合作組
設計印刷	漢藝有限公司
	235 新北市中和區中山路 2 段 315 巷 8 號 2 樓
	02-2247-7615
出版日期	2019 年 10 月

ISBN: 978-957-678-676-1 (平裝) GPN: 1010801705