literature.[21] The outstanding Silk Road Historical Geography Information Open Platform 丝绸之路历史地理信息开放平台 from Capital Normal University combines extensive data on western China and Inner Asia and a variety of other utilities.[22] Academia Sinica's Chinese Civilization in Time and Space 中華文明之時空基礎架構 includes several animated maps showing change over time but is not an open platform.[23]

Creating, Linking, and Analyzing Chinese and Korean Datasets: Digital Text Annotation in MARKUS and COMPARATIVUS

Hilde De Weerdt*

Leiden University
*Corresponding author. Email: h.g.d.g.de.weerdt@hum.leidenuniv.nl

MARKUS, a multilingual digital text annotation and analysis platform, allows historians and other researchers to construct datasets from primary sources available to them in full-text digital format. Originally designed for those working with pre-twentieth-century Chinese texts, MARKUS has developed into a multifunctional annotation platform that is particularly suited for the automated annotation, referencing, and visualization of named entities in modern and literary Chinese and premodern Korean texts, but many of its additional annotation features can be used to analyze and read texts in any language, as long as the electronic documents are encoded in the most common standard for language encoding, Unicode. Below I discuss the main goals and methodological features of MARKUS and the allied text comparison utility COMPARATIVUS.[1] I will illustrate these with some examples of how MARKUS has been used in Chinese and Korean historical research.

---

[21] http://amap.zju.edu.cn.

[22] www.srhgis.com

[23] http://ccts.sinica.edu.tw/framework.php?lang=zh-tw.

[1] On the algorithm used in COMPARATIVUS, see Paul Vierthaler and Mees Gelein, "A BLAST-based, Language-agnostic Text Reuse Algorithm with a MARKUS Implementation and Sequence Alignment

Automating Text Annotation

Why annotate texts digitally? Historians have been digitally annotating primary sources for a variety of reasons. For some marking up texts is a flexible way to produce a digital edition of a source; annotation is then above all about the structural features of the text (its parts, chapters, sections, etc.). For others digital annotation is equivalent to the notepads and card files of the past: it is a means to collect, organize, and retrieve important topics and passages relevant to particular research questions. The more structural and the more semantic aspects of digital annotation can also be combined, either to produce critical editions in which people, places, time references, and so forth have been indexed, or to allow for a faceted exploration of the annotated topics or passages in which the organization of the original text is maintained. In the latter case the annotations are also regularly aggregated for quantitative analyses and reproduced as data. MARKUS has been primarily designed for the purpose of semantic annotation and text analysis. I will illustrate this using a few examples that also immediately underscore the importance of a clear research question and a well-defined plan for a meaningful digital annotation project, both in terms of the object (text or corpus of texts to be included) and method (the types and procedures of annotation to be used).[2]

Chu Mingkin conducted analyses of the correspondence networks of Song and Yuan Dynasty scholar-officials on the basis of an annotation of the digital corpora of their letters. He used MARKUS for the annotation of all personal names, official titles, and place names and also for the structural division of the digital text into individual letters.[3] Based on a comprehensive analysis of the connections among correspondents, their locations, their offices, and the nature of their correspondence he pieced together the political ties that were being forged in seemingly mundane polite notes collected in individual and collective anthologies of letters. In a different vein, Michael Stanley-Baker collected and mapped the uses of drugs in a broad range of medical texts across time with the automated and keyword markup features in MARKUS. Other examples in literary, intellectual, and art history, in the history of infrastructure, and in other contexts are discussed on the MARKUS forum.[4] In all

Optimized for Large Chinese Corpora," *Journal of Cultural Analytics* March 18, 2019. DOI: 10.31235/osf.io/7xpqe.

[2]In these examples the data that can be extracted from the annotation (exported from MARKUS in a range of file formats or to other tools and platforms) are a byproduct. Structural and semantic annotation can also be used to produce an enriched edition of historical documents. For example, with Gabe van Beijeren I prepared a digital edition of *The Essentials of Governance from the Reign of Constancy Revealed* (*Zhenguan zhengyao* 貞觀政要) to accompany a translation of the text. The digital edition allows for a very different kind of reading than standard print or even digital editions. Readers can experience the minor differences between manuscript and standard print editions and rearrange the text in a variety of ways. Based on the MARKUS annotations, passages can be filtered by chronological sequence, by speaker, or by those in attendance. Readers can also access linked references to locate further information about any term. More practically, tagging can also be used for editorial purposes. Lists of official titles, place names, personal names, book titles, or key concepts can be readily generated to standardize translations or create indexes. Hilde De Weerdt, Gabe van Beijeren, and Mees Gelein, "Reading *The Essentials of Governance from the Reign of Constancy Revealed* Digitally," 2020. https://chinese-empires.eu/zgzy.

[3]Chu Mingkin, "Secret of Long Tenure: A Study of Zheng Gangzhong's Letters to Qin Hui's Associates," *T'oung Pao* 102.1–3 (2016), 121–60; id. "Jin Yuan zhi ji de shiren wanglu yu xunxi goutong—yi *Zhongzhou qizha* nei yu Lü Xun de shuxin wei zhongxin" 金元之際的士人網絡與訊息溝通——以《中州啟劄》內與呂遜的書信為中心, *Beida shixue* 北大史學/*Clio at Beida* 20 (2016), 286–310. Texts, data, and interactive reading platform available at http://chinese-empires.eu/reference/publications/.

[4]See contributions by Michael Stanley-Baker, Hsu Ya-hwei, Margaret Wan, Xiong Hueilan, Chu Ping-tzu, Hilde De Weerdt and others at https://dh.chinese-empires.eu/forum/category/8/research-blogs.

these and other featured cases the researcher first defined a set of research questions (e.g., What was the social significance of mentioning particular correspondents in individual and collective anthologies in the south and north during the Song and Jin/Yuan periods? How has the use of particular drugs changed over time and spread across space? What kinds of places are associated with what subgenres of novels? How did objects move across collections?). They delimited a relevant body of literature that could vary in size from one text or section within a text to all the poetry or prose produced during a few centuries, or the entire Buddhist or Daoist canon. They also set out ahead of time what types of information needed to be annotated under what types of tags and outlined a procedure to carry out both semantic and structural markup in a systematic fashion. These first steps may appear self-evident to any researcher, but they are key to determining the significance and reliability of any digital research project—there should be some space for experimentation but far too often students are tempted to tag away, hoping to make some sense out of whatever results appear.

Text annotation can be accomplished in regular text editors, so why use MARKUS for this? For its default named entity markup MARKUS uses authoritative scholarly datasets in Chinese, Taiwanese, Korean, and Buddhist studies (Figure 1). I will explain the advantages of this below. In addition, the keyword markup module offers a range of functionality to input term lists, to produce KWIC (Key Word In Context) lists or regular expressions for markup in texts in any language, or to detect relevant keywords based on a similarity test with a term selected from any text uploaded by the user. For large corpora of texts, the batch markup feature can be used to simultaneously tag entities, keywords, or regular expressions in dozens or hundreds of files, as long as those have been uploaded in MARKUS file management. In the allied text comparison utility COMPARATIVUS, the reader can detect text reuse in two or more texts, select passages of meaningful overlap from a table or the texts themselves, and send the selected passages back as markup to the relevant files in MARKUS—the default settings for comparison are optimized for Chinese texts but can be modified. This can be used, for example, to locate and save quotations from particular texts. By default, passages sent back from COMPARATIVUS are tagged with a standard tag type ("comparativus," but tag names can be edited in MARKUS to differentiate between quotations from different texts for example).

In sum, as a first step MARKUS can be used to discover and tag, in individual texts or in collections of texts, a range of Chinese and Korean named entities and keywords, regular expressions, or overlapping passages in any language.

Linking Data

Tagging in MARKUS is more than a process of finding matches in uploaded texts from the linked scholarly datasets or from user-defined lists of terms. A particularly important feature of the MARKUS environment is that default tags are or can be linked to unique identifiers or IDs. A tag consists of the tagged content (a string of characters in the text), a tag type (e.g. person, place, time, plant name, etc.), and an ID (a number or other kind of unique identifier for the particular entity referred to in the tagged content). For example, the historical figure Wei Zheng 魏徵 can be referred to in the text in a number of ways: Wei Zheng 魏徵, Zheng 徵, Wenzhen 文貞 (a posthumous honorary name), and so forth. Because it uses alternative names included in China Biographical Database (hereafter CBDB; see Lik Hang Tsui and Wang Hongsu "Harvesting Big Biographical Data for Chinese History: The China Biographical Database (CBDB)" in

2.a Automated markup

Tag Chinese personal names (full, forenames, and alternate names) and bureaucratic titles based on **China Biographical Database** (CBDB).
Tag Chinese place names based on **China Historical GIS** (CHGIS) and **Taiwan GIS** (TWGIS).

Tag temporal references based on **DILA Time Authority Database** (法鼓佛教學院時間規範資料庫).
Tag personal names from Buddhist sources based on **DILA Person Authority Database** (法鼓佛教學院人名規範資料).
Tag place names from Buddhist sources based on **DILA Place Authority Database** (法鼓佛教學院地名規範資料).
Tag Buddhist terms based on **DILA Glossaries of Buddhist Studies** (法鼓佛教學院佛學詞彙).

Tag Korean personal names (full, forenames, alternate names, etc.) based on **Korean Historical Biographical Information System** (韓國歷代人物綜合情報系統) and **Encyclopedia of Korean Culture** (韓國民族文化大百科事典).
Tag Korean place names based on **Encyclopedia of Korean Culture** and **Map of Dongyeo** (東輿圖).
Tag bureaucratic titles of the Joseon dynasty based on **Veritable Records of the Joseon Dynasty** (朝鮮王朝實錄) and **Encyclopedia of Veritable Records of the Joseon Dynasty** (朝鮮王朝實錄事典).
Tag Korean books based on **Encyclopedia of Korean Culture, Comprehensive Edition of Korean Literary Collections** (韓國文集叢刊) and **Encyclopedia of Veritable Records of the Joseon Dynasty**.

One can select one or more types of markup.

Ready

Figure 1. List of datasets that can be selected in MARKUS automated markup.

this issue), MARKUS will attempt to tag all relevant instances referring to this person and add the relevant CBDB ID (15610 in this case). In this way, all instances referring to this person can be found and exported, regardless of the particular phrasing used in each instance. Tagging in MARKUS can thus be used to normalize data annotated in and extracted from the text. The same applies to place names, time references, bibliographical information, etc. Sometimes the researcher may have to decide between multiple available IDs (the same name can refer to more than one person in the database, or the same person can be included in multiple databases), and sometimes the researcher may opt to add her own IDs if persons of interest are not included in the linked databases.

An ID also establishes a direct link with a corresponding record in external databases containing a range of additional information about the entity to which the tag refers. For example, when MARKUS adds the CBDB ID "15610" to any instance of 魏徵 or 徵, the user can directly consult (in the right pane) the following information about Wei Zheng: dates; places where he lived, worked, or had his ancestral home;

bureaucratic offices he held; family relationships; other social relationships; texts he authored or compiled; and references to other databases, print reference materials, and primary sources with biographies of Wei Zheng. A selection of this information can be directly exported to the Palladio and PLATIN platforms from the VISUS field in MARKUS file management.[5] The same applies to geographical names for which the user can generate an ID by selecting the appropriate location from historical gazetteers shown in the right-hand pane. This ID is linked to longitude, latitude, and other geographical information in associated databases such as TGAZ and the Dharma Drum place name authority dataset;[6] in this way the data annotated in the text can be mapped in linked or standalone geographic information systems.

Analyzing and Visualizing Annotated and Linked Data

One of the great advantages of standard digital markup languages is that they allow texts and other content to be rendered or published in a wide variety of ways. This kind of flexibility also enhances the durability of such an annotated text when compared to a text formatted in commercial software based on proprietary formats. MARKUS uses standard markup languages so that tagged texts and exported data can be read in a range of text analysis and visualization platforms and open and commercial software. Annotated texts can be exported to HTML, XML-TEI, and MARKUS, and COMPARATIVUS data can be downloaded in several tabular formats (CSV, TSV, Excel, HTML).

In MARKUS we also simplified the steps that are typically involved in analyzing data embedded in digitally annotated texts: extracting tagged data, merging tagged data with data from external datasets, and then visualizing and analyzing the combined data. We developed MARKUS into a linked platform in which a large part of the annotation and visualization can be undertaken automatically. By linking files saved in MARKUS to Palladio and PLATIN, researchers can, via the VISUS interface, import biographical information linked to tagged names from CBDB and then explore it, alongside their own data, in maps, network diagrams, tables, timelines, or pie charts. From here they can also export all data for more sophisticated analysis in more specialized spatial, network, or statistical packages. For example, by exporting an annotated MARKUS file comprising the correspondence of the twelfth-century statesman and celebrated author Yang Wanli 楊萬里 (1127–1206) to Palladio, the hundreds of letters in it can be visually explored on a networked spatial map linked to an interactive timeline and topical filters. These were created on the basis of user-generated tags in the file such as recipient name, location of sender and/or recipient, type of letter, main themes covered in the letter, and user-supplied metadata such as the year in which the letter was written.[7]

---

[5]Humanities + Design Research Lab at Stanford University, Palladio, 2014, at https://hdlab.stanford.edu/palladio-app/; Max-Planck-Institute for the History of Science, PLATIN, http://skruse.github.io/PLATIN/.

[6]Huimin Bhikṣu, Aming Tu, Marcus Bingenheimer, Jen-Jou Hung, et al. Buddhist Studies Authority Database Project, 2008, at http://authority.dila.edu.tw/; Peter Bol, Lex Berman, et al., China Historical GIS, 2001, at www.fas.harvard.edu/~chgis/.

[7]See Hilde De Weerdt, "The Uses of Digital Philology in Tang-Song History, Part 2," MARKUS forum: research blogs, January 14, 2017 https://dh.chinese-empires.eu/forum/topic/31/the-uses-of-digital-philology-in-tang-song-history-part-2.

While Palladio and PLATIN work well for the visual exploration of smaller corpora and datasets, the more recently implemented exchange between MARKUS and Docusky (National Taiwan University) enables MARKUS users to export dozens or hundreds of annotated files for further text analysis in Docusky and for spatial mapping in the associated DocuGIS platform—Docusky exports files in XML, which can in turn be reconverted into MARKUS files.[8] Docusky offers MARKUS users a range of extra functionality of which only a few will be mentioned here. First, in Docusky large numbers of files can be aggregated into one textual corpus, and multiple textual corpora can be compared to each other on the basis of word and tag frequencies. Second, Docusky offers metadata services that are linked to MARKUS. Users can supply metadata for MARKUS files or textual divisions within MARKUS files that can be used alongside tags to explore corpora. MARKUS tags can, furthermore, be converted to metadata. For example, volume or chapter headings can be first annotated in MARKUS and then converted to metadata in Docusky so that they can be used to browse the text or search results by volume or chapter.

Third, and very importantly, Docusky corpora tagged in MARKUS can be exported to DocuGIS in which all spatial IDs will be associated with the corresponding longitude and latitude. DocuGIS is a basic geographic information system in which spatial layers can be generated from MARKUS tags and used alongside other topographical and administrative spatial layers. Users can export spatial datasets from DocuGIS in formats that can be easily read in other and more advanced geographical information systems (see also Peter Bol, "The Visualization and Analysis of Historical Space" in this issue). An early example of the analytical potential hereof is a collaborative pilot project to map and compare city wall construction in three provinces across the Ming Dynasty based on city wall inscriptions preserved in local gazetteers. Particular features of walls (construction materials, types of fortification, size), reasons for deterioration, or contributors to and labor force involved in construction projects can be examined over time and in the context of topography, administrative boundaries, historical meteorological layers, or regional clustering.[9] A particular strength of the MARKUS-DocuGIS environment is that any data point on the map remains linked to the original source text, allowing for interactive reading, checking, and even for the editing and correction of the spatial layers.

MARKUS is thus designed and continues to be developed to model existing research flows, allowing for cycles of reading, markup, analysis, and interpretation. To improve the discovery of and access to digital texts, the first step in any digital annotation project, MARKUS is now linked to commonly used open access textual repositories such as Donald Sturgeon's Chinese Text Project (see Sturgeon, "Digitizing Premodern Text with the Chinese Text Project" in this issue) and Christian Wittern's Kanripo, from which texts can be directly imported into MARKUS. Texts from these and other repositories can also be exported to MARKUS through the SHINE API, developed at the Max Planck Institute for the History of Science and the Staatsbibliothek zu Berlin.

---

[8]Tu Hsieh-Chang, Hsiang Jieh et al., Docusky. http://docusky.digital.ntu.edu.tw/.

[9]Hilde De Weerdt, "The Uses of Digital Philology in Tang-Song History, Part 1," MARKUS forum: research blogs, Jan. 14, 2017 https://dh.chinese-empires.eu/forum/topic/30/the-uses-of-digital-philology-in-tang-song-history-part-1.

Curation and Customization

MARKUS was co-designed by humanities researchers and computer scientists with a philosophy of agile software development. Researchers and students were invited at workshops to evaluate MARKUS processes and functionality critically, to raise awareness about the theoretical and methodological implications of digital text annotation, digital reading, and data analysis, and also to contribute towards priorities and revisions in future development. A range of additional features and customization options were added to ensure a close alignment with the interests and research practices in humanities scholarship.

Because humanities research is often an iterative process involving reading, rereading, interpretation, revision, and reinterpretation, we designed annotation modules in MARKUS to allow for a wide range of editorial interventions: correcting the text, custom tags and manual markup, batch deletion and revision of tags, redesign of custom tags, adding comments, and custom settings for the selection of online dictionaries and datasets for consultation. Custom functions require login with a free personal account. An experimental and preliminary machine learning module allows users to generate markup on the basis of machine learning results from a batch of files that have been correctly annotated—in automated markup pre-annotated files can then be selected as the set of files from which rules (regular expressions) for annotation should be automatically generated. This can, for example, be used to detect regularities in particular genres of writing: when annotating a biography of a certain genre (*muzhiming* 墓誌銘) based on dozens or hundreds of pre-annotated biographies, for example, one can expect that, in contrast to the default named entity markup, first names following kinship terms will be detected.

The list of desired functionality and improvements to existing functionality is considerable, and tackling each of these takes time, due to the fact that development for MARKUS requires financial support and a multidisciplinary team. Most recently, we have added a long-anticipated functionality for relational markup, allowing researchers to establish and define a relationship between two tags. Each tag can have multiple relationships as an attribute, and for each relationship the user can add a relationship type and metadata (e.g., external references to primary and/or secondary sources for that relationship). With this feature researchers can generate far better datasets for network analysis than heretofore: relations can be exported as network files including source and target nodes and relationships types as well as other attributes. Relational markup can also be used to establish hyperlinks between passages across multiple MARKUS files.

Conclusion

MARKUS originated from the generalization of a methodology that was first used in one particular project: the systematic digital annotation of sources of information in notebooks (*biji* 筆記) in order to map the temporal, geographic, and social distribution of informants in communication networks as they are reflected in these sources.[10] This generalization in turn resulted from the interest shown in such a mapping of sources by scholars in various fields in the humanities and social sciences. Since a first version went

---

[10]Hilde De Weerdt and Brent Ho. *Information, Territory, and Networks: The Crisis and Maintenance of Empire in Song China*, Accompanying data and visualization site, 2015. http://chinese-empires.eu/reference/information-territory-and-networks/.

live in the summer of 2014 MARKUS, which currently only runs in Google Chrome, has been used by 14,680 unique users (figure as of October 4, 2019). The MARKUS site includes a forum with research blogs and tips (e.g., how to redesign custom tags, or when to use batch markup or keyword markup rather than automated markup), short instructional videos, bug reports, and announcements. The site and many of the instructional materials are available in three languages and four scripts (English, traditional Chinese characters, simplified Chinese characters, and Korean). MARKUS was originally developed as an open source tool; the code of the original version and COMPARATIVUS can be used and modified for non-commercial purposes.

Recommended Links

Software

De Weerdt, Hilde, and Brent Ho. MARKUS: A markup, reading, and visualization platform for classical Chinese texts 2014. https://dh.chinese-empires.eu/markus/. Code: https://github.com/dHumanities/markus

De Weerdt, Hilde, Mees Gelein, and Brent Ho. COMPARATIVUS: A text comparison platform 2017. http://dh.chinese-empires.eu/comparativus/

Code: https://github.com/dHumanities/markus/tree/master/comparativus

De Weerdt, Hilde, Hu Jing, Mees Gelein, Brent Ho, Kim Hyeon, and Kim Baro. K-MARKUS: Korean text analysis and reading platform 2019. https://dh.chinese-empires.eu/markus/beta/

Tu Hsieh-Chang, Hsiang Jieh *et al.* Docusky. http://docusky.digital.ntu.edu.tw/ Instruction manual (including basic MARKUS instructions). https://docusky.digital.ntu.edu.tw/DocuSky/ds-11.instructions.html

Recent slide presentations:

De Weerdt, Hilde. Creating, Linking, and Analyzing Chinese and Korean Datasets in the MARKUS Environment. Cyberinfrastructure and Platforms, Digital Technologies Expo, Association for Asian Studies Annual Conference, Denver, March 22, 2019. www.eventscribe.com/uploads/eventScribe/PDFs/2016/7155/875394.pdf

De Weerdt, Hilde. Developing a Text Analysis Infrastructure for East Asian Languages: MARKUS, VISUS and COMPARATIVUS. Council on East Asian Libraries annual conference plenary session, Denver, March 20, 2019. https://drive.google.com/file/d/1c4QA13MaArvBnU-j-t_IBUlIfPNQIBM4/view

Videos

De Weerdt, Hilde, Brent Ho, Daniel Stumm, Hu Jing, Gabe van Beijeren, Xiong Hueilan, Qiao Jiyan, Monica Klasing-Chen, Liu Jialong, and Feng Ying. MARKUS instructional videos. https://dh.chinese-empires.eu/markus/beta/video.html (English) https://dh.chinese-empires.eu/markus/beta/video_zhtw.html (traditional Chinese) https://dh.chinese-empires.eu/markus/beta/video_zhcn.html (simplified Chinese) 2014-

De Weerdt, Hilde. MARKUS, VISUS & COMPARATIVUS: Developing a Text Analysis Infrastructure for East Asian Languages. *Digital Archives: New Fields in East Asian Cultural Studies*. Kansai University Open Research Center for Asian Studies, Osaka, February 17, 2018. www.youtube.com/watch?v=I1WAS-_faGk (1.15–2.00)

De Weerdt, Hilde. Digital Perspectives on Middle-Period Political History. University of Michigan, Ann Arbor, Lieberthal-Rogel Center for Chinese Studies, April 5, 2016. https://youtu.be/2oxHTEFEa38 (different versions of this talk: Gothenburg http://media.hum.gu.se/filedb/index.php?cdir=TmpVNU1qZz0%3D&c_hash=62c1644730fcf46086164dc08fdcf5e8 and http://media.hum.gu.se/filedb/?cdir=TmpVNU1qYz0%3D&c_hash=41c9fba5f4490c7c0501ac047752a02b and Stanford https://vimeo.com/168242706)

De Weerdt, Hilde. Humanities Tools for Library Resources. University of Michigan, Ann Arbor, University Library, April 4, 2016. http://leccap.engin.umich.edu/leccap/viewer/r/azO7QY

De Weerdt, Hilde. Wenben biaoji yu lishi yanjiu 文本標記與歷史研究 (Textual Markup and Historical Research). Academia Sinica, The Institute of History and Philology, Taipei, April 29, 2015. https://www.youtube.com/watch?v=NltG3EjC9_A

De Weerdt, Hilde. Songdai xin zixun jiegou de xingcheng 宋代新資訊結構的形成 (The Development of a New Information Regime in Twelfth-Century Song China). National Taiwan University, Chinese Department, Taipei, April 27, 2015. www.youtube.com/watch?v=1Xd_mJ9eJHk

Podcast and media

Liu, Jialong. "访谈｜魏希德：如何将数位人文工具Markus用于历史研究." Interview on Digital Humanities, *The Paper*/澎湃, February 10, 2017. http://m.thepaper.cn/newsDetail_forward_1611410

Shuman, Amanda. "Hilde De Weerdt on MARKUS." DH East Asia Podcast, July 31, 2016. http://www.dheastasia.org/2016/07/31/podcast-3-hilde-de-weerdt-on-markus/

Blogposts

De Weerdt, Hilde. "The Uses of Digital Philology in Tang-Song History, Parts 1 & 2." MARKUS forum: research blogs, January 14, 2017 https://dh.chinese-empires.eu/forum/topic/30/the-uses-of-digital-philology-in-tang-song-history-part-1 & https://dh.chinese-empires.eu/forum/topic/31/the-uses-of-digital-philology-in-tang-song-history-part-2

De Weerdt, Hilde. "Collaborative Innovation and the Chinese (Digital) Humanities." University of Nottingham China Policy Institute Blog, June 9, 2016, https://blogs.nottingham.ac.uk/chinapolicyinstitute/2016/06/09/collaborative-innovation-and-the-chinese-digital-humanities/

De Weerdt, Hilde. "Isn't the Siku quanshu enough? Reflections on the impact of new digital tools for classical Chinese." Communication and Empire: Chinese Empires in Comparative Perspective, February 20, 2014, http://chinese-empires.eu/blog/isnt-the-siku-quanshu-enough-reflections-on-the-impact-of-new-digital-tools-for-classical-chinese/

De Weerdt, Hilde. "Digital Interpretations." Communication and Empire: Chinese Empires in Comparative Perspective, February 5, 2014, http://chinese-empires.eu/blog/digital-interpretations/

MARKUS on social media

https://twitter.com/did_acte
http://tw.weibo.com/5822830227

Related tools

Brent Ho. Open access text finders for MARKUS (http://dh.chinese-empires.eu/markus/beta/index_ctext.html) and https://dh.chinese-empires.eu/markus/beta/index_k.html)

Sturgeon, Donald. Ctext MARKUS plugin https://ctext.org/tools/plugins/list

Brent Ho. CBDB lookup tool (http://dh.chinese-empires.eu/markus/beta/cbdbLookup.html)

Brent Ho. CBDB dictionary (http://chinese-empires.eu/analysis/tools/cbdb-mac-dictionary/

Brent Ho. The Sieve Online. https://dh.chinese-empires.eu/markus/beta/sieveOnline.html (based on Joshua Day and Sarah Schneewind, The Sieve, 2013, http://ctext.org/tools/literacy-sieve)