The Chinese Biographical Database Project: A Model for Collaborative Scholarship

Marilyn A. Levine, Lewis-Clark State College, Lewiston, Idaho USA

Abstract

The Chinese Biographical Database Web site [CBD] at http://www.lcsc.edu/cbiouser is the first scholarly moderated database on biography to appear on the Web. Based on the moderated list server ideal, the CBD is open to the public in a collaborative endeavor. Access to the database to conduct searches and reports is open to all, and anyone can contribute missing information and new biographies, which are considered for inclusion if they are supported by source evidence. Currently, the CBD has over 22,000 records, 13 database tables, 2,500 figures, over three dozen query and report forms, in addition to a capacity for the entry of new biographies. Within the first few months of going online the CBD has been revised to incorporate temporal and spatial information such that GIS methods can be used to locate fields from the database. New entry categories have been made generic to accommodate regional and chronological scope. The Chinese Biographical Database Web site has already demonstrated a new dimension of active learning and sharing. This report will overview the Chinese Biographical Database Project in its first development phase, discuss the initial and ongoing development; and give some comments on the assessment, significance, and future of the CBD.

Overview of the Chinese Biographical Database

During 1997-1998, funded by a research grant from the Idaho Board of Education, I was able to work on the creation of an online moderated database based on Chinese biography. This endeavor and the concept for it arose from two influences. First, for the last 3 years I have been exploring new learning technologies for teaching and professional service on several Web sites that I developed. The dimension that I found most intriguing was the interactive potentials on the Web, as well as the enormous integrative power. Secondly, after serving as an online editor for a Humanities and Social Sciences Online group (H-ASIA) for two years I really appreciated the validity of the foundation idea of scholarly, moderated online discussions. It occurred to me that the effort to duplicate data detracts from the interpretive power of historians. Thus, my conception of the Chinese Biographical Database Web site (CBD) was to have a scholarly moderated web site that would be shared dynamically online as opposed to a downloadable database. The foundation of an interactive scholarly moderated database is twofold:

1. That database entry is open to collaborative endeavor.
2. That all new entries or missing variables entered must be supported by documentation. If information is not supported by a source entry it will not be considered for inclusion.

The CBD Web site has the following characteristics:

1   Over two dozen searches
2   Eleven report forms
3   Over one-quarter more than the anticipated biographical figures were entered for a total of 2,450 figures. About 30% of these figures have between 3-5 sources.
4   The CBD has the capacity for reading and entry in Chinese characters
5   The CBD Web site has an extensive introduction section. Other sections include: Searches, Reports, New Biographical Entries, Feedback, and a Help section.

Among the areas that need expansion are:

1   The need for a Boolean search capacity
2   The need for a linking function for the searches
3   The need for a total report form
4   The need for a statistical and graphical section

As can be seen in Table 1, there were over 22,225 records entered in the database tables. In addition to the records entered, each table had numerous fields. While there is much data that is Not Available (NA), substantial amounts of data were entered. Excluding the NA entries, for example, the Biographical figures table had a total of 18,850 pieces of information, the Sources table had 31,215 data entered, the Alternate Names Table contains 10,039 data, and the Positions table holds 6,080 pieces of data.
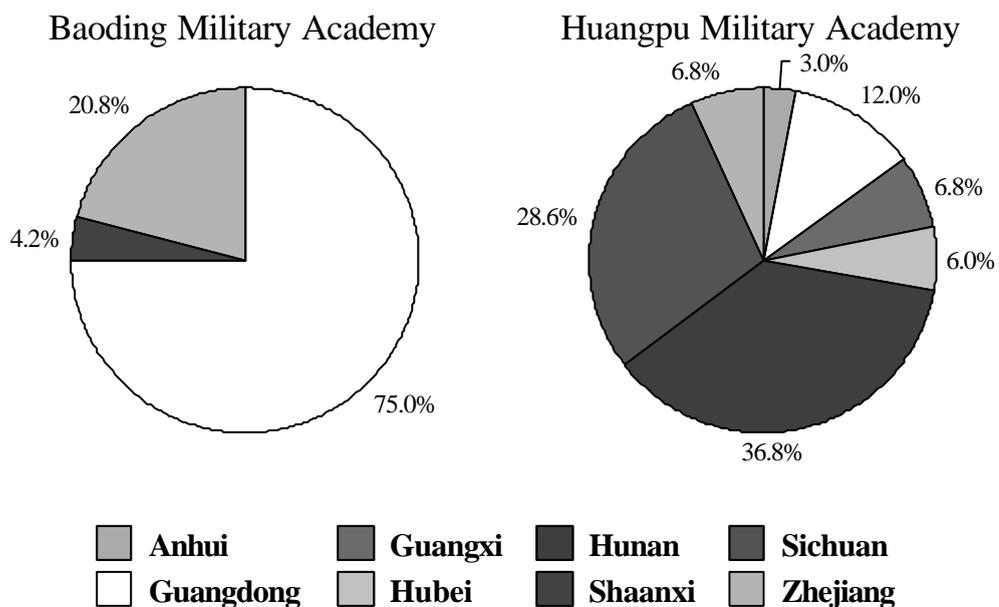
Table Type	Records Entered
Biographical figures	2,450
Career	2,450
Source citations	5,125
Alternate Names	2,750
Family	175
Youth	385
Education	1,700
Affiliations	2,300
Positions	3,040
Events	1,840
Locator	10
Total Records Entered	22,225

Table 1. Tables and Numbers of Data Entered for Online Tables

The CBD is a very exciting humanities resource. For example, the database has at present over 2,500 alternate names (in ten different categories). The Chinese often have several names for different life phases or activities. The CBD has the ability to search by individual name - listing all the alternate names and name types; by the alternate name - listing the name type and the original or most well known name; and a listing of all alternate names in the database. Some of the most important cultural and political leaders have dozens of names that researches can now link with who they are through utilization of the Chinese Biographical Database.

Another measure of the utility of the Chinese Biographical Database is the scope of the knowledge conserved and the diversity of response. For example, the drop down list for the Historical Events contained over 75 historical events, focusing on the post-1850 period (see Table 2). Other Events Drop down lists could be designed for subsets and differing chronological eras in the future temporal and regional expansion of the database.

In the expansion of the Chinese Biographical Database the diversity and quantity of the data will make a significant contribution, especially when an online statistical and graphical section is added along with a Boolean search. For example, the Official Positions table includes more than 40 types of positions of which a sampling of various positions statistics

will provide some basis of both quantitative and qualitative analysis. For over twelve kinds of table types and numerous variables, particular cohorts can be analyzed by linking various

categories. For example, in Figure 1 we can compare a small cohort of the 2,500 figures in the database who went to Baoding and Huangpu Military Academies, through the factor of provincial origin. The statistical and graphical results of this kind of query allows us to consider trends and anomalies such as the intriguing fact that there is a larger percentage of students from Guangdong at the Northern-based Baoding Military Academy rather than the Southern-based Huangpu Military Academy. History based on the reality of what actually occurred will create a foundation for interpretations that will have a reliable breadth.

1911 Revolution	Language Reform
Anti-Christian Movement	Literature Revolution
Anti-Japanese Boycotts	Long March
Anti-Rightist Campaign (1957)	Mass Campaigns (1950's)
Anti-Warlord Activities	Mass Campaigns (1960's)
Anti-Western Boycotts	May Thirtieth Movement
April 12th Coup, 1927	May Fourth Movement
Autumn Harvest Uprising (1927)	Muslim Uprisings
Battle of Taierzhuang	Nanchang Uprising (1927)
Boxer Uprising	Nanjing Decade-Government
Canton Commune (1927)	Nien Rebellion
Canton-Hong Kong Strike	New Fourth Army Incident
CCP-Wuhan split, 1927	Northern Expedition
Civil War, 1946-49	Northern Expedition
Collectivization of Agriculture	Optium War (1839-40)
Consolidation of Republic (1912-1927)	Peasant Movement
Constitution Protection Movement	Qing Army Campaigns
Cultural Revolution	Qing Reforms
Cultural Revolution-Purged/Disappeared	Red Base Activities
Debates on Ideologies and Philosophies	Red Guard Participation
Debate on the National Form of Literature	Religious Liaison Work
Debate on Democracy and Dictatorship	Rural Reconstruction Work
Debates on Chinese and Western Culture	Second Revolution (1913)
December Ninth Movement	Self-Strengthening Movement
Democracy Wall (1979)	Sino-Japanese War, 1937-45
Democracy Demonstrations (1989)	Sino-Japanese War, 1894-95
Economic Modernization, post-1979	Sino-Soviet Relations
Ethnic Liaison Work	Sino-Vietnamese War
First Education Mission to United States	Tao Yuan Movement
Five Encirclement Campaigns	Taiping Rebellion
Five-Year Plans	Taiwan Independence Activities
Four Cleanups Campaign	Taiwan Uprising (2/28/47)
Four Modernizations	Third World Delegations
Fujian Rebellion	Three Antis Campaign
Great Leap Forward	Tongzhi Restoration
Gung Ho Movement	United Front, 1923-27
Hundred Days Reform	United Front, 1937-45
Hundred Flowers Campaign	Wang Jingwei Regime
Labor Movement	Warlordism
Land Reform	Yan'an Base Activities

Table 2. Historical Event Types

Development and Assessment of the CBD in the Initial Phase

In terms of the Project creation, during the first phase, several months had been allocated in the original plan to the design phase of the project. By the end of the fall, it was decided that because of compatibility issues and the need for Chinese characters, that the program *Access 97* would be utilized. The design was created by having 11 tables that had a one-to-many relationship such that multiple entries could be made for individual figures in categories such as names, family, education, historical events, affiliations, positions, and sources. By late December 1997 the private entry Web site had been designed, tested and redesigned and was ready for data entry. The data entry was performed from January - June 1998 by two students, who did a commendable job in their attention to detail. Overview of their work was performed through weekly downloading from the server of the database.

In March 1998 the CBD went public and within one week more than 500 people had looked at the site. Within 24 hours the CBD Web site was rated *Essential* by T. Matthew Ciolek, who heads the WWW Virtual Library, the major assessor of academic Web sites in the world. This rating is only given to 7% of the Web sites viewed. Within two weeks the CBD Web site was cited in the Scout Report, appeared in the first electronic newsletter pertaining to online resources, AsianDoc, which was in turn highlighted in the Chronicle of Higher Education online. Since then the CBD appears in most search engines and within 12 weeks over 80 Web sites have linked to the CBD as a resource. After being approached by scholars from major institutions - there may be possibilities of merging other datasets into the CBD, and the CBD will be a partner in several Web-based endeavors.

In addition to high public and private evaluative feedback, one measure of evaluating the Web site is by examining the data from a Web tracking tool. The statistical summary and other indicators suggest the profile of a Web site that has several pages visited per visit with actual requests for searches and reports. One page that should have had more interaction was the Introduction, which means that the viewers may not be optimally understanding or utilizing the site.

These are the summary results of that tracking data:

Data is from February 1, 1998 - May 22, 1998
Total Number of Requests   18,693

Total Number of Visitors   1,920
Total Number of Visits   2,838
Average Visits Per Day   26.0
Average Requests Per Visit   1.6

Lessons Learned and Future Goals

Though off to a good beginning the CBD has a long way to go to be of optimal utility to viewers. Aside from expanding the technical capability of the site, the biggest challenges are philosophical. In particular two areas that need reflection are:

1. The issue of encouraging scholars to share an online resource.   This was the most difficult and surprising result of the CBD. I had expected many more colleagues would contribute to the site than actually did contribute.   My strategy is to first increase the technical capacity of the site and then really consider how to encourage an environment of such a shared endeavor.

2. There are several issues surrounding the admission of sources, published, non-published and archival. For example, should criteria be developed for the uploading of contributions that can assess the validity of the sources? The rationale behind this is a database is only as reliable as the data entered.

Future plans for expansion include five different dimensions:

Expansion of Technical Capabilities to Enhance Utility
>   Migration to a Unix Server
>   Linked Search Results
>   Boolean Search
>   Total Report Form
>   Statistical / Graphical Section
>   Multimedia Section
>   Longitude and Latitude and Time for Spatial-Temporal Analysis

Expansion of Chronological Scope

Expansion of Geographical Scope to a Global Biographical Database

Expansion of Database to Server Pedagogy
>   The expansion of the database as a foundation for K-12 Education Web sites

The expansion of the database as a foundation for Post Secondary Web sites, with particular attention to graduate research

Philosophical and Organizational Development of a World Biography Center

Because of the belief in the "active" nature of technology, I would eventually like to develop a Center for World Biography, with the database and attendant pedagogy sites as a foundation.

The theory and practice of biographical research and its implications would be a major focus of such a Center

In conclusion, there is a growing sense of cooperation worldwide in utilizing the World Wide Web to preserve and share knowledge. The Chinese Biographical Database as a pilot project that seeks to be a part of that movement. The CBD has gained significant restructuring and insights in terms of issues such as developing chronotopically and with more expanded, versatile fields, by the input of such online collective projects as the Electronic Cultural Atlas Initiative and the Humanities and Social Sciences Online (the latter pioneered the scholarly moderated listserve model). The CBD is now in Version 2.0 and has online discussions available for suggestions on the technical and scholarly issues of the database structure. The CBD aspires to provide a model of scholarly collaboration that is interactive and expansive. It is an attempt to provide a new ethos of possibility and cooperation within the rapidly changing new technologies.