

12. **Geographic Administrative Hierarchy** (defined in political terms as administrative units)
13. **Longitude and Latitude** (fixed locations in space required for historical comparisons)
14. **Bureaucratic Organization** (the changes in bureaucracy and reporting responsibilities over time)

## ***B. Details of Entities***

NOTE: The database allows one to record the **Source** of information, including the **Pages** in the source from which the information comes, and to add additional **Notes** as seems appropriate. Every item in the database that records information on an individual has the attributes of **Source**, **Pages**, and **Notes**. Therefore I will not note these in the discussions below.

### **1. People**

#### **a. Basic Data: name, male or female, date of birth, and date of death.**

Precise dates of birth and death often are not available, and all we have is a period of **years of activity** (“*floruit*” dates). Sometimes, not even that is available: we simply know the **reign period** (*nianhao*) or **dynasty**. In order to capture the level of precision in the data, the database allows the use of reign period information for all dates. One can give a specific year within the reign period, but one also can simply indicate “beginning,” “middle,” “end,” or “unspecified.” For analytic purposes, the database will algorithmically produce Western dates from the reign period information for birth, death, years of activity, and any other date given in the traditional Chinese *nianhao* designation, but it will preserve the vagueness in the **nianhao** coding.

#### **b. Ethnicity and Tribe Affiliation**

CBDB tracks ethnicity, like Han, Uighur, Tibetan, etc. We have over 465 codes at present. These codes are in the table ETHICITY\_TRIBE\_CODES, which organizes ethnicity and tribe designations by group and subgroup and includes variant forms for ethnicity names.

#### **c. Choronym**

From the Six Dynasties into the Tang, membership in a clan was of central importance in defining one’s social status. From the Song Dynasty onward people did make claims descent from a particular clan from a particular place (like the Cui

clan of Boling) but they carried little social or political weight. The combination of place name and clan name defined a *choronym*. The codes for these choronyms are in the table CHORONYM\_CODES.

### c. **Index Year**

For computational purposes, CBDB needs a single year value to locate a person in time. The *index year* is an artificial value used in analyses. In earlier versions of the database, index year was based on when the person would have turned 60 *sui*. However, starting with the 2021 dataset, the index year has been based on the known or projected **year of birth**. The rules for calculating the value are complex and based on the following assumptions:

- A1:** that a man received a Jinshi (進士) degree at age 30, the Juren (舉人) degree at 27, and the Xuicai/licentiate (秀才/生員) degree at 21
- A2:** that a wife was 3 years younger than her husband
- A3:** the first child was born when the father is at age 30 and a mother at age 27 (per assumption A2)
- A4:** that male children were born 2 years apart
- A5:** that a man died at age 63 and a woman at age 55

#### *Rules Based on a Person's **Birth/Death Dates***

- Rule 1: Ego's index year = ego's birth year
- Rule 2: If we know ego's death year and age at death, then: ego's index year = ego's death year – age at death
- Rule 20: If we know just the ego's death year then: (per A5) ego's index year = ego's death year – 63 (for men), ego's death year – 55 (for women)
- Rule 4W: Ego's index year = husband's birth year +3 (Note: If the woman was a concubine/second wife, then rule 9W precedes rule 4W.)

#### *Rules Based on **Degree Dates***

- Rule 5: Ego's index year = the year he obtained the Jinshi (進士) - 30
- Rule 5W: Ego's index year = the year her husband obtained the Jinshi (進士) - 30+3 = husband's Jinshi year - 27
- Rule 6: Ego's index year = the year he obtained the Juren (舉人) - 27
- Rule 6W: Ego's index year = the year her husband obtained the Juren (舉人) - 27+3 = husband's Juren year - 24
- Rule 7: Ego's index year = the year he obtained the Xuicai (秀才/生員) - 21
- Rule 7W: Ego's index year = the year her husband obtained the Xuicai (秀才/生員年) - 21+3 = husband's Xuicai year - 18

#### *Rules based on **Birth Years of Kin***

- Rule 8: If we know the birth year of ego's **father**, then ego's index year is decided per assumption A3: ego's birth year was 30 years later than father's birth  
(Ego's index year = father's birth year + 30)
- Rule 9: If we know the birth year of a **male's oldest child**, then ego's index year is decided per A3: ego's birth year was 30 years earlier than the birth year  
(Male's index year = oldest child's birth year - 30)
- Rule 9W: If we know the birth year of a **female's oldest child**, then ego's index year is decided per A3: ego's birth year was 27 years earlier than the birth year of her oldest child. (Female's index year = (oldest child's birth year - 27))
- Rule 10: If we know the birth year of ego's **older brother**, then ego's index year is decided per A4: ego's birth year was 2 years later than older brother's birth year. (Ego's index year = (older brother's birth year + 2))
- Rule 11: If we know the birth year of ego's **younger brother**, then ego's index year is decided per A4: ego's birth year was 2 years earlier than younger brother's birth year. (Ego's index year = (younger brother's birth year - 2))
- Rule 12: If we know the birth year of a **male's oldest son-in-law**, then ego's index year is decided per A3 & A4: ego's birth year was 30 years earlier than the birth year of his oldest daughter, and his oldest son-in-law was 3 years older than oldest daughter. (Male's index year = (birth year of oldest son-in-law + 3 - 30) = birth year of oldest son-in-law - 27)
- Rule 12W: If we know the birth year of a **female's oldest son-in-law**, then ego's index year is decided per A3 & A4: female's birth year was 27 years earlier than the birth year of her oldest daughter, and her son-in-law was 3 years older than her oldest daughter. (Female's index year = (birth year of oldest son-in-law + 3 - 27) = birth year of oldest son-in-law - 24)
- Rule 13: If we know the birth year of ego's **grandfather**, then ego's index year is decided per assumption A3: ego's birth year was 60 years later than grandfather's birth year. (Ego's index year = (grandfather's birth year + 60) = grandfather's birth year + 60)

#### *Rules Based on the Index Years of Kin*

(Note: CBDB iteratively uses the derived index years for these values.)

- Rule 14: If we know the index year of ego's **father**, then we use father's index year to decide ego's index year per A3. (Ego's index year = father's index year + 30)
- Rule 15: If we know the index year of a **male's oldest child**, then we use that year to decide ego's index year per A3. (Ego's index year = index year of oldest child - 30)

- Rule 15W: If we know the index year of a **female's oldest child**, then we use that year to decide ego's index year per A3. (Female's index year = index year of oldest child - 27)
- Rule 16: If we know the index year of ego's **older brother**, then we use that year to decide ego's index year per A4. (Ego's index year = index year of older brother + 2)
- Rule 17: If we know the index year of ego's **younger brother**, then we use that year to decide ego's index year per A4. (Ego's index year = index year of younger brother - 2)
- Rule 18: If we know the index year of a **male's oldest son-in-law**, then we use that year to decide ego's index year per A3 & A4: ego's birth year was 30 years earlier than the birth year of his oldest daughter, and his oldest son-in-law was 3 years older than oldest daughter. (Ego's index year = index year of oldest son-in-law +3 -30 = index year of oldest son-in-law - 27)
- Rule 18W: If we know the index year of a **female's oldest son-in-law**, then we use that year to decide her index year per A3 & A4: female's birth year was 27 years earlier than the birth year of her oldest daughter, and her son-in-law was 3 years older than her oldest daughter. (Ego's index year = index year of oldest son-in-law +3 -27 = index year of oldest son-in-law - 24)
- Rule 19: If we know the index year of ego's **grandfather**, then we use grandfather's index year to decide ego's index year per A3. (Ego's index year = grandfather's index year + 60)

The CBDB table that records this basic biographical information is BIOG\_MAIN. BIOG\_MAIN assigns each person a unique ID.

#### d. Floruit years

CBDB gives two years: the earliest and the latest. Often when there is no data for index year or for birth and death dates, texts nonetheless provide datable references to individuals. CBDB gives the earliest and the latest known dates given in the textual sources we have examined so far.

## 2. Kinship

An instance of the **Kinship** relationship for an individual has three components (plus the source information):

```

person
kin
kinship relation

```