

Information Discovery in the Chinese Recorder Index

Jieh Hsiang, Jung-Wei Kong, Hao Sung

Research Center for Digital Humanities and Department of Computer Science
National Taiwan University
jhsiang@ntu.edu.tw

Chinese Recorder (CR) is a (first bi-monthly, then monthly) journal published by the Protestant missionaries in China between 1867, a few years after the 1860 treaty that allowed missionaries to enter China, and 1941, when the U.S. became engaged in the Pacific theater of the Second World War. Except for a nineteen-month interruption from June 1872 to the end of 1873, CR was the longest running English missionary journal in China. The period that CR covered was a tumultuous time in Chinese history, when the country went through the Taiping Rebellion, the Boxer Rebellion, various wars with foreign powers, the Republic Revolution of 1911, civil wars among the war lords, the rise of communism, the invasion of Japan, and the great cultural and social transformation during the late 19th and early 20th centuries. Being based within China, CR provided a close look at all spectrum of the Chinese society, not only missionary affairs, but also issues such as Chinese civilization, healthcare, education, political situation, opium, and other social issues of the day. CR is unique in that the articles were written by missionaries in China for the benefit of their fellow missionaries. Being supported by missions and not sponsored by any government, Chinese or foreign, the views presented in CR were more candid and not blurred by political agenda. It, thus, provides an angle unlike any others. Although one cannot say that CR is not biased, at least it is biased in its own special way.

With its 73 volumes, one for each year, and over 50,000 pages in total, CR is difficult to use on its own. In 1986 Kathleen Lodwick published the 2-volume *The Chinese Recorder Index: a Guide to Christian Missions in Asia, 1867 – 1941* (CRI), which made it much easier for scholars to utilize CR in their research.

CRI is more than an index. It consists of 3 indexes and 6 special lists (of Persons by Affiliation, Persons by Location, List of Women, etc). The *Person Index (PI)* includes 8,391 names with 192,149 page records; the *Mission/Organization Index (MI)* 712 mission entries with 34,851 page records; and the *Subject Index* 4,691 entries with 83,636 page records. Altogether, there are 13,794 entries and 310,626 page

records, averaging 22.5 page records per entry.

Unlike a conventional book index that only provides the pages in which an entry appears, CRI assigns tags to PI and MI, thus provides additional information about the nature of the occurrence of the entry on that page. Tables 1 and 2 give the names, nature, and numbers about the tags that we have tabulated.

| Tag name of Person Index | Tag value | number/no value | explanation |
|--------------------------|-------------------|-----------------|--------------------------------------------------------------------------|
| AFF (Affiliation) | Name of mission | 27,678 / 14 | Affiliation |
| ARR (Arrival) | Year | 8,705 / 6,283 | Arrival year (if no value, then it's indicated by volume number) |
| ART (Articles) | | 11,125 | Article written by person |
| ATT (Attacks) | | 1,164 | Attack on the person |
| CHI (Children) | | 7,765 | Children of the person |
| CON (Conferences) | Location, year | 3,449 / 1,793 | Conference attended |
| COR (Correspondence) | | 2,276 | Correspondence to the editor |
| DAT (Dates in China) | Year | 139 / 2 | Dates in China |
| DEA (Death) | Precise year | 2,459 / 2,035 | Death (if no value, then it's indicated by volume number) |
| DEP (Departures) | Year of departure | 5,496 / 5,000 | Departure in the year(if no value, then it's indicated by volume number) |
| ITI (Itinerancy) | Location | 445 / 413 | Itinerancy |
| LOC (Location) | Location | 39,744 / 8 | Location |
| OTH (Other Publications) | | 10,095 | Other publications |
| POS (Positions) | Title and mission | 6658 / 2 | Position of the person |
| SPO (Spouse) | | 13,212 | Spouse of the person |
| UNS (Unspecified) | | 51,739 | Unspecified |

Table 1 Person Index

| Tag name | Tag value | Number/no value | explanation |
|-----------------------|-----------|-----------------|----------------------------|
| ATT (Attacks) | | 275 | Attacks on the M/O |
| CON (Convert) | | 996 | Converts |
| FIN (Finances) | | 1,585 | Finances of the M/O |
| HIS (History) | | 650 | History of the M/O |
| HOS (Hospitals) | | 639 | Hospital |
| MEE (Meetings) | | 2,340 | Meeting |
| LOC (Location) | Location | 9,533 / 32 | Location |
| OPR (Opium Refuges) | | 0 | Opium refuge |
| ORD (Ordained Asians) | | 349 | Asians ordained by mission |
| ORP (Orphanages) | | 32 | Orphanages |
| PER (Personnel) | | 1,520 | Personnel |
| PRE (Press) | | 1,554 | Press |
| REP (Reports) | | 772 | Reports by the M/O |
| SCH (Schools) | | 1,309 | School run by the M/O |
| STA (Statistics) | | 2,510 | Statistics about the M/O |
| UNS (Unspecified) | | 10,763 | Unspecified |

Table2 Mission/Organization Index

There is a great wealth of information hidden in the tags. For example, if a certain page appears in a person name entry A, and if that page also appears in another person name entry B under an ART (article) tag, then we know that A appears in an article written by B. Using the same page number to check all entries, and we can find all person names, locations, missions, subjects, etc that appear in the same article. If the same page appears in a Subject Index entry indicating a certain event, then we may even “guess” what the article is about without reading the article itself. As another example, the ATT (attacks) tag tells how many attacks of missions and missionaries had been reported in CR, who were attacked, in what years did they occur and where. This information should be valuable to someone studying the attitude of the Chinese society towards Christianity when it was re-introduced to China in the late 19th century. However, since “attack” is not an entry by itself, this information is scattered all over CRI. One has to pore through the 13,000 entries to collect every relevant bit of information.

In this paper, we present a system that incorporates the 3 indices of CRI into one uniform framework, under which the indices are fully integrated and cross referenced. This integration enables the user to explore the rich information hidden within CRI.

Our approach starts with a data structure that decomposes an index entry into a

number of *page records*, each consists of the entry name (*n*), a volume number (*vol*) and page numbers (the start page and the end page, *s* and *e*), a tag (*t*) and a tag value (*v*). (The tag part may not be present if it is not indicated.) Thus, a page record is a tuple (*n*; *vol*:*s-e*; *t*:*v*). If there is only one page indicated, then the start page and the end page will be the same. For instance, the first page record in the George Leslie Mackay entry (Figure 1) is (Mackay, George Leslie; 13:74-74; AFF: CPM) and the second page record is (Mackay, George Leslie; 13:312-312; AFF: CPM). (The George Leslie Mackay entry as shown in this example becomes 46 page records.) This process will decompose the 13,794 entries in the three indices into 310,626 page records. We then designed algorithms to mine the relationships among the page records, mainly using the page numbers and the tags as reference points. While the detail of the algorithms cannot be covered in this abstract, we will use two simple examples to demonstrate the outcome.

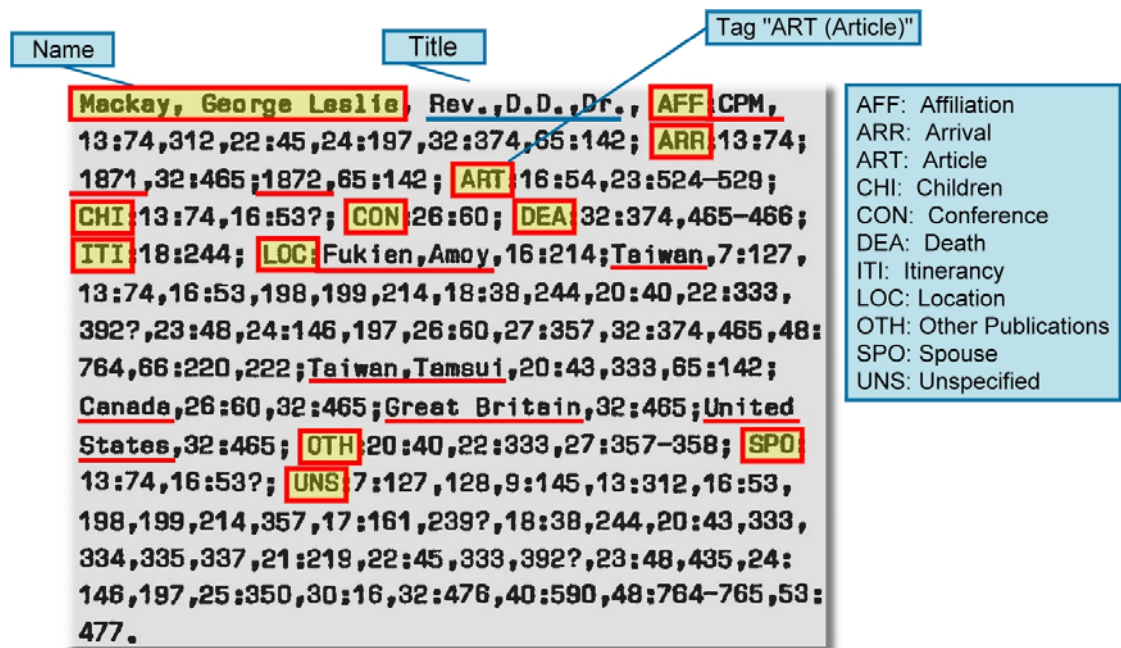


Figure 1 The "George Mackay" entry in Person Index

One of the page records in the Person Index entry of George Mackay is (Mackay, George Leslie; 16:214-214; LOC: Fukien, Amoy). Since in the entry of Thomas Barclay, there is a page record (Barclay, Thomas; 16:214-215; ART), we know that Mackay was mentioned in an article written by Barclay. Other page records show that the same page appeared in the Subject Index of Sino-French War (1884 – 1885), the Person Index of James Maxwell, William Thow, etc. (Figure 2), plus others. Thus, when issuing George Mackay as a query, instead of simply returning Vol 16, p214 as

one of the pages in which Mackay appears, all information in CRI about that page will be organized and returned (Figure 3). (Indeed, it was an article written by Barclay about the returning of the missionaries from Amoy to Taiwan after the lifting of the blockade of Taiwan at the end of the Sino-French War.) Figure 4 shows the webpage resulting from the query “George Mackay”. Note that a chronological distribution is presented and a foldable/expandable classification of the return is on the left.

| | |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mackay, George Leslie LOC(Location): Taiwan LOC(Location): Fukien, Amoy | Barclay, Thomas ART(Article) |
| Mackay, George Leslie, Rev., D.D., Dr., AFF:CPM, 13:74,312,22:45,24:197,32:374,65:142; ARR:13:74; 1871,32:465;1872,65:142; ART:16:54,23:524-529; CHI:13:74,16:53?; CON:26:60; DEA:32:374,465-466; ITI:18:244; LOC:Fukien,Amoy,16:214; Taiwan,7:127, 13:74,16:53,198,199,214,18:38,244,20:40,22:333, 392?,23:48,24:146,197,26:60,27:357,32:374,465,48; 764,66:220,222;Taiwan,Tamsui,20:43,333,65:142; Canada,26:60,32:465;Great Britain,32:465;United States,32:465; OTH:20:40,22:333,27:357-358; SPO: 13:74,16:53?; UNS:7:127,128,9:145,13:312,16:53, 198,199,214,357,17:161,239?,18:38,244,20:43,333, 334,335,337,21:219,22:45,333,392?,23:48,435,24: 146,197,25:350,30:16,32:476,40:590,48:764-765,53: 477. | Barclay, Thomas, Rev., D.D., Dr., AFF:BFBS,64:399; EPM,5:225?,374?,9:279?,15:234,20:290,22:297,24: 50,150,26:338,32:155,37:319,40:358,478,56:608,60: 679,81:65,66:742; ARR:15:234,24:150,32:155,40: 358;1874,64:830; ART:16:214-215,35:103,60:536- 538; CON:5:225?;Shanghai,8:241;Taiwan Christian Youth,66:764;CLS Annual,53:791?; COR:18:362,19: |
| | Wars, Sino-French, 1884-1885 Sino-French, 1884-1885 —16:52-54,64,67,88-95, 105,112-113,146-150,208-213,214-215,234, 279,317-318,318-320,357,429; 17:39,279,360, 434; 18:44,150,366; 21:334; 25:602; 31:476; 38:318,416. |
| | Maxwell, James Laidlaw DAT(Dates in China): 1884-1885 DEP(Departure) |
| | Maxwell, James Laidlaw, Rev., M.D., Dr., AFF:CMEDA, 63:717;CMMA,48:184,54:130,55:700,56:844;EPM,1:6, 65-68,2:59,5:143,8:433,9:398,15:234,16:240,55: 700;IRCS,68:781;MMSC,1:119,64:256; ARR:15:234,56: 844;1863,2:59,21:269; ART:MR:76-77,1:65-68,2:112- 113,55:730-737,57:816-818,63:715-717,65:394-395, 68:532-533; ATT:1:168,258-259; CHI:56:844; CON: CMMA,54:130;CMedA,65:395; COR:1:76,258-259,37: 542-454; DAT:1884-1885,16:214; DEP:16:198,214, 240; LOC:Fukien,Amoy,18:290;Hueoh,Hankow,68:781. |

Figure 2 Some entries containing Vol 16, p 214



Volume 16 Page 214-215 Author: Barclay, Thomas

出現在本頁中的辭條 (共 12 筆) : [檢視所有辭條](#)

1. 在 p.214-215 出現的紀錄 :

Missions, China, Taiwan - [Subject Index](#)

SUB (Subject)

Wars, Sino-French, 1884-1885 - [Subject Index](#)

SUB (Subject)

China, Taiwan - [Subject Index](#)

SUB (Subject)

Barclay, Thomas - [Person Index](#)

ART (Articles in Chinese Recorder)

2. 在 p.214 出現的紀錄 :

Barclay, Thomas - [Person Index](#)

LOC (Location) : Location LOC (Location) : Location

Mackay, George Leslie - [Person Index](#)

UNS (Unspecified) LOC (Location) : Location LOC (Location) : Location

Anderson - [Person Index](#)

LOC (Location) : Location LOC (Location) : Location UNS (Unspecified)

Maxwell, James Laidlaw - [Person Index](#)

UNS (Unspecified) DAT (Dates in Mission Field) : Dates in Mission Field DEP (Departure) : Departure

Thow, William - [Person Index](#)

LOC (Location) : Location UNS (Unspecified)

3. 在 p.215 出現的紀錄 :

Ashmore, William - [Person Index](#)

UNS (Unspecified) CHI (Children) AFF (Affiliation) : Affiliation

Barclay, Thomas - [Person Index](#)

LOC (Location) : Location

Partridge, S.B. - [Person Index](#)

AFF (Affiliation) : Affiliation

Ashmore, William, Jr. - [Person Index](#)

AFF (Affiliation) : Affiliation LOC (Location) : Location UNS (Unspecified)

American Baptist Missionary Union (ABMU) - [Mission/Organization Index](#)

COV (Converts) ORD (Ordained Asians) STA (Statistics)

Figure 3 Information returned about Vol 16, p214 when given query “George Mackay”

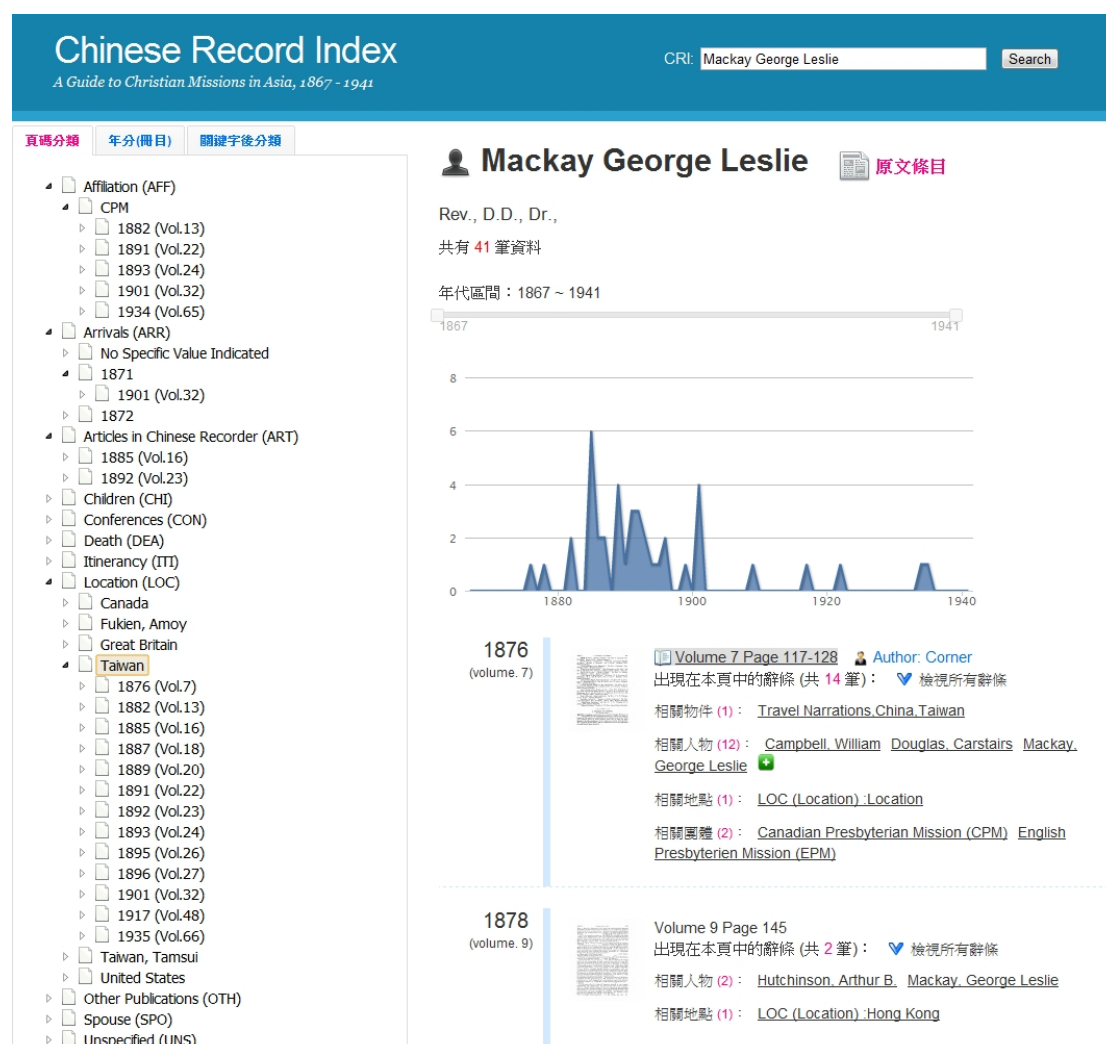


Figure 4 Query result of “George Mackay”

Figure 5 shows the return of the query “attack”, for which a total of 574 different pages in the entire CR are retrieved. (There are all together 1,439 page records with the ATT tag. But some referred to the same pages.) The peak occurred in 1900, during the Boxer Rebellion, which may not be surprising. However, the post-classification of the query result (Figure 6), an important feature of our system, shows the lists of authors, persons, missions, which might provide interesting information for scholars to pursue further.

The method and system that we have developed provide a more global view of both CR and CRI. It integrates the three indices of CRI and thus reveals relations among them and information implicitly embedded. Our work should make CR and CRI more accessible to the research community. The approach that we have taken is also a general one that can be applied to any index with a similar structure.

References

Kong, J.-W., (2011) *Design and Implementation of a Retrieval System for the Chinese Recorder Index*, MS Thesis, National Taiwan University

Lodwick, K., (1986) *The Chinese Recorder Index: A Guide to Christian Missions in Asia, 1867-1941*, Scholarly Resources Inc. Washington D.C.

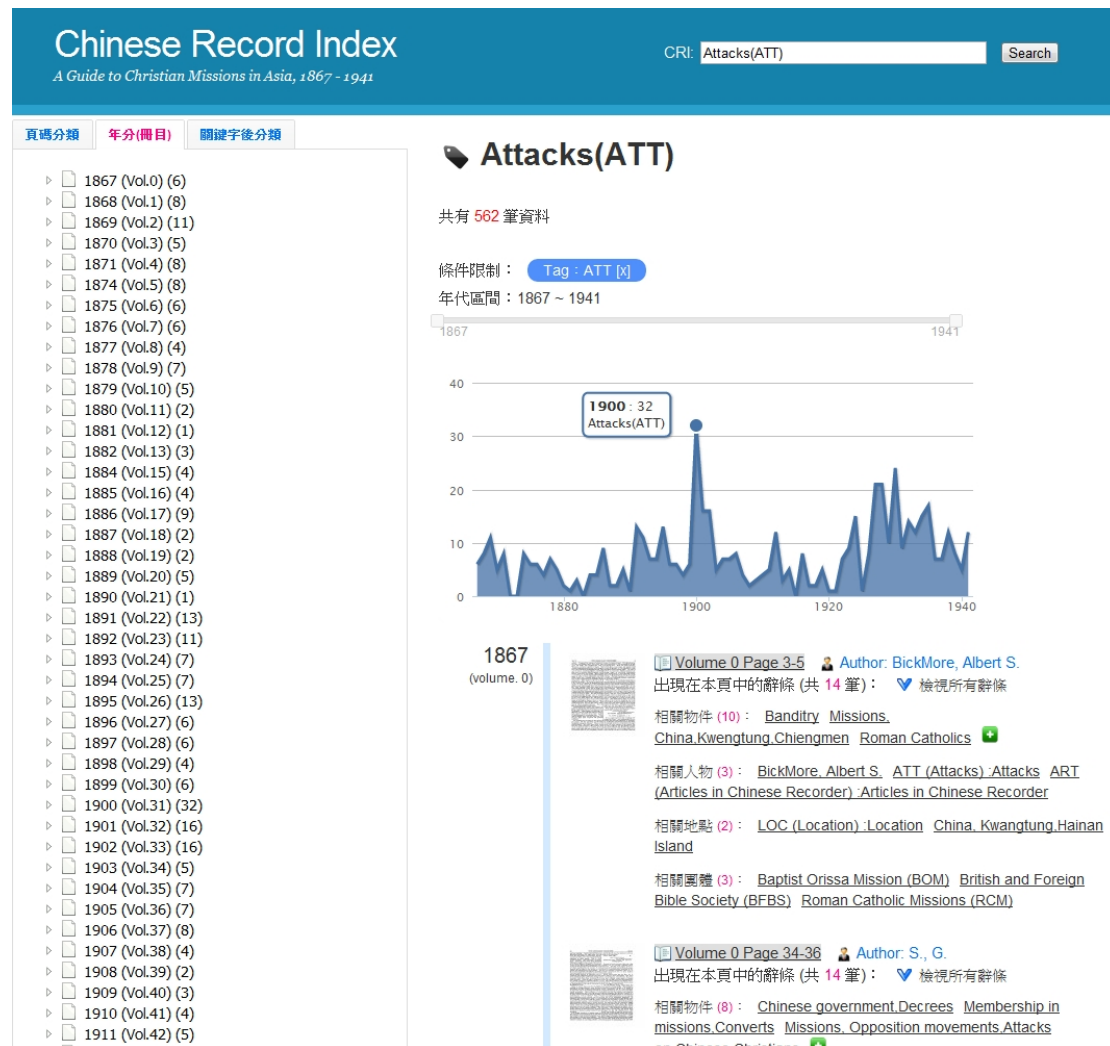


Figure 5 summary of query “attack”

Figure 6 Classifications of the query result of “attack” according to Author, Person, Mission, and Subject.