



Digital Newspaper Clippings

Kazuko Sakaguchi
Documentation Center on
Contemporary Japan

Newspapers are the handiest news source for information about current problems and their importance. Covering practically everything from factual news to critical thoughts on various subjects, newspapers keep us informed of the current status of the world from various perspectives; they are an amalgam offering a blueprint of the world. Indeed, as Eizo KOYAMA once stated, "every event and matter in the world could be sorted, simplified, annotated, and formalized by newspapers." Compared to other media, newspapers are respected: according to a recent survey, more than 90% of people regard newspapers as useful and reliable. And since current news becomes the topic of later research, the use of historical newspaper articles has steadily increased. For all these reasons I would like to take this opportunity to introduce some notable archives of digital newspaper clippings.

Digital Newspaper Articles
(Kobe University Research Institute
of Economics and Business
Administration)

<http://www.lib.kobe-u.ac.jp/sinbun/>

Since they consist of bulky print materials that accumulate daily, newspapers are quite cumbersome to save for archival purposes. The clipping of newspapers, which physically extracts newspaper articles in various lengths that are then mixed like a mosaic, was invented as a means of archiving to substitute for the storage of whole newspapers. The Newspaper Clippings Collection at the Research Institute of Economics and Business at Kobe University is a unique collection of clippings on economics that has accumulated approximately 500,000 items from 1911 through 1970. The Collection was micro-filmed in the 1970s because it was originally printed on acid paper not suitable for long-term storage and since then about 40,000 clippings have been digitized to prevent further deterioration. However, since raw

digitized images are not searchable, the Newspaper Clipping Collection also has now digitized 25,000 articles in full text format. For this last group one can utilize a keyword search for the entire texts, which is one of distinctive features of this database. The articles were clipped from a great variety of newspapers, including many from the Kansai area, other regional and local newspapers, and foreign newspapers published in the former colonies during the pre-war era, such as *Taiwan Daily*, *Manchuria Daily*, and *Keijyo Daily Report*. A wide range of more than 200 categories was used. Articles can be grouped by chronology, type of industry, newspaper title, and subjects. By typing "sugar industry," for instance, one can view a list of search results that lists up to 500 items, which can be sorted. Also, clippings can be viewed by both HTML and PDF formats.

Asahi Shimbun Clippings
(Institute of Social Science at
University of Tokyo)

<http://www.iss.u-tokyo.ac.jp/>

Any newspaper company maintains clipping files, not mainly for archival purposes but rather to support the current writing and editing of its reporters. The early invention of reduced size editions of newspapers called *shukusatsuban*, containing a month of coverage with indexes to the headlines, provided a useful tool for locating newspaper articles. The Asahi Daily Newspaper Company first published a *shukusatsuban* in 1919. The Asahi Shimbun Clippings at the Institute of Social Sciences at University of Tokyo, a joint project developed with Asahi Daily Newspaper Company, is a database of such *shukusatsuban* headlines. Although highly selective, it has an extensive coverage from April 1922 through September 1997, indexing about 87,000 articles on political science, political party, administration, and social problems. When typing "hikiagesha" (repatriates) in a topic field, for instance, one can see ten hits with bibliographic information on dates, pages, headlines, related people, and related topics. Eventually the system may provide links to full texts of articles; however, only a headline index is available as of this writing. For more comprehensive headline searches, one may use the CD-ROM

version of the *Sengo 50-nen Asahi Shinbun midashi detabasu* (Asahi Headlines for 50 years, 1945-1995). For the Shōwa period specifically, the recent *Asahi Shimbun senzen shimen detabesu* in CD-ROM format provides both a headline index and full texts of articles.

Meiji no Yomiuri Shimbun (Yomiuri
Shimbun in the Meiji Period)
CD-ROM

<http://www.kinokuniya.co.jp/03f/denhan/yomiuri/meitai.htm>

In contrast to newspapers called *o-shinbun*, which emphasized political discussions, the Meiji-period *Yomiuri Shimbun* used to be called *ko-shinbun* because of its focus on ordinary social events. As the targeted readers were primarily "children and women," articles in the *Yomiuri* were written in semi-spoken Japanese, with many kana pronunciations and pictorial illustrations added. The CD-ROM version of the *Yomiuri Shimbun* in the Meiji Period enables one to perform an electronic search of these early, often-ignored materials. The CD-ROM contains an index database, which provides each article with subject headings, and image texts which were digitized from microfilms, so that users can view original headlines in various fonts and original physical layouts for each article. It usually requires some background knowledge to perform an effective search for events and people in the Meiji period; however this CD-ROM's user-friendly system (which has also mounted *The World Encyclopedia* by Heibonsha) enables searches by using links to keywords, dates, a dictionary, an encyclopedia, and chronological tables. Unusually, the CD-ROM includes newspaper advertisements, which provide valuable information on social trends and fashion during the period. With a release of similar CD-ROMs for the Taishō and Shōwa periods (Prewar periods I & II), newspaper articles in the *Yomiuri Shimbun* are now searchable beginning with its inaugural issue in 1875 through the end of 1945. The days of requiring tremendous time and labor to locate articles in the pre-war era are over!



Nikkei Telecom 21
<http://telecom21.nikkei.co.jp/nt21/service/>

Many newspaper companies offer some kind of free access to online articles at their web sites, but only for a limited time and coverage. A seamless search capability is only available through commercial databases. Database development for newspaper articles accelerated in the 1980s. The Nikkei Telecom 21 is a comprehensive aggregated newspaper database which covers more than forty newspapers, including Nikkei's own four special newspaper editions and some English language publications. (One can undertake a seamless search for major English language newspapers in Japan, such as *The Japan Times*, *Daily Yomiuri*, *Mainichi Daily News*, and *Nikkei Weekly*.) As it was originally intended for users in business, the Nikkei Telecom 21 has a wide coverage of data sets in economics, including journals, company data, financial data, and biographical data. The Nikkei Telecom 21 also offers a digital automatic clipping file service that allows users to electronically save articles that contain keywords set up beforehand. There are some weaknesses: the database is basically in text format (some are in PDF), so charts and tables cannot be retrieved, and photographs and advertisements are unfortunately excluded. The Nikkei service is also expensive; however it might be effectively used in combination with *shukusatsuban* and microfilms elsewhere available in one's library.

Newspaper Directory Database (National Diet Library)
<http://sinbun.ndl.go.jp/>

Newspapers are classified in libraries as serials. However, there are a variety of serial-type, newspaper-like publications that fall into the broad category "newspaper," from nationwide papers to community newspapers with a limited circulation, or multiple versions of newspapers published by the same company (e.g., morning paper, evening paper, regional and local editions). Moreover, one may count *kawaraban* ("roof-tile prints" of commercial newsheets) and *gogai* (extra editions put out by newspapers) as newspapers. In addition, newspapers are published in different formats, such as *shukusatsuban*, microfilms, and reprints. Therefore, it is not so easy to find out which library keeps one's desired "newspapers" and how comprehensive their holdings are. The Newspaper Directory Database, developed by the Bureau of Subjects and Newspapers at the National Diet Library (NDL), is a useful tool for bibliographic information about newspapers, covering over 1,300 libraries' holdings in Japan. One can obtain lists of newspapers arranged in alphabetical order and Japanese syllable order, as well as by dates and formats. For the ease of conversion of the publication date from a specific year in the vernacular calendar to the Gregorian calendar, a link to a conversion table is also provided. However, although a submission of at least one copy of every title published in Japan to the NDL is mandated, one should be aware that not all newspaper titles are included in this Directory. The NDL also maintained newspaper clippings for fifty years after its establishment in 1948. Although not digitized yet, these clip-

pings are now available in the 50 volumes of *Documents of Post-War Japan: Encyclopedia of Social History Through Newspapers* (Tokyo: Oozorasha, 1994-1998).

In general newspaper databases have a distinctive advantage for information searching because of the sheer abundance of information. However, they also commonly have some shortcomings. As previously noted, they typically cannot retrieve charts and tables, and thus some analytical data is lost. Signed articles under copyright protection cannot be included in the databases. And it is impossible to fully accurately read and evaluate newspaper articles without viewing the original headlines and physical layout of each article in the physical newspapers which provided the original visual impact. Moreover, the excessively easy access to information enabled by the databases has also aroused unprecedented questions of violation of privacy and human right issues. New forms of legal dispute and concerns have occurred regarding the correction of misinformation about scandals and criminal investigations, and the fair treatment of convicted persons who have completed prison terms. Losses in privacy lawsuits have made newspaper publishers increasingly cautious about the management of online databases, encouraging them to impose strict licensing requirements for the fair use of these databases. Because of these limitations and concerns of the digital newspaper databases, the Asahi Newspaper Clipping Team, which was established in 1911, still continues its tradition of physical clipping of newspapers, even after the Team's inclusion into the Database Section at Asahi's Bureau of Digital Media.

