

SOPAG/CDC Task Force on CDL-Licensed News Resources  
***Report to the Collection Development Committee***  
University of California Libraries

December 13, 2006

Submitted by:

Kay Collins, UC Irvine  
Harold Colson, UC San Diego  
Elliot Kanter, UC San Diego  
Linda Kennedy, UC Davis. Chair  
Wendy Parfrey, CDL  
Lucia Snowhill, UC Santa Barbara

## Table of Contents

I. Executive Summary .....	3
II. Introduction and Charge .....	5
III. Methodology & Analysis .....	6
IV. Findings .....	8
IV. A. Overlap of News Resources .....	8
IV. B Current News Resources: Gaps in Coverage and Future Needs .....	10
IV. C Historical Newspapers: Gaps in Coverage and future needs .....	11
IV. D Open Access/Public Domain Digital Newspaper Collections .....	13
V. The Collection Development Process.....	14
VI. Enhancing Newspaper Access, Cooperative Activities.....	15
VII. Future of Newspapers .....	17

## Appendices

Appendix A, Task Force Charge.....	18
Appendix B, Summary Comparison of Databases (Attached Excel spreadsheet)	
Appendix C, Database Descriptions and Anayses (attached Word document)	
Appendix D, Sample Pages, Serials Solutions Overlap Analysis (Not included, sample printout pages from comparisons available on request from Lucia Snowhill or Committee Chair Linda Kennedy)	
Appendix E, Campus Newspaper Web pages.....	21
Appendix F (not included), Database of All NewsTitles (Not included, extensive spreadsheet available on request from Lucia Snowhill or Committee Chair Linda Kennedy)	

## I. Executive Summary

Note: Findings are numbered F1 through F6 and Recommendations are numbered R1 through R11 to facilitate discussion.

### *Findings:*

- The overlap in coverage among the major current news databases is not nearly as great as expected. There is virtually no overlap in historical newspapers resources. (F1)
- No major gaps in domestic news coverage were identified. However, international news coverage is strongest for Europe, then Latin America and Asia; weak for Africa and the Middle East. (F2)
- Except for the three ProQuest newspapers—*New York Times*, *LA Times*, *Wall Street Journal*—none of the current news databases has comprehensive coverage of any single newspaper title nor includes page image viewing. (F3)
- Major news databases include other types of resources (e.g. legal resources/business resources/journals) that may have been the primary reason for acquiring the database. *LexisNexis Academic* and *Factiva* are two important examples. (F4)
- UC does not have perpetual rights to the current news resources, and perpetual access only to some portion of the historical newspapers from ProQuest. The aggregator news resources do not provide us with content that can replace the need for microfilm as the credible guarantor of the domestic and foreign press record.(F5)
- Because of the mix of tier 1-tier 2-tier 3 databases containing news sources, and because the scope of these resources extends beyond covering news, there are not clear options for reducing licensing of databases. Choices can potentially be made more clearly at the campus level. (F6)

### *Recommendations:*

#### *Overlap and Gaps, Licenses*

- Consider reducing duplication in current coverage of *New York Times* and *Wall Street Journal*. (R1)
- Campuses should address duplication of international and domestic news resources consistent with their local needs. (R2)
- CDL should consider a Tier 1 license for *LexisNexis Academic* to ensure that UC campuses are receiving the lowest consortial price. (R3)
- Several historical domestic news sources are recommended to complement those already licensed. Recommended resources are included in section IV.C. (R4)

#### *Collection Development Process*

- Create a new, separate newspapers bibliographer group to monitor news resources in all formats. (R5) This group could be charged to:
  - Coordinate communication among bibliographer groups regarding news resources;
  - Provide background, analysis and comparisons for newspapers proposals;
  - Make recommendations for newspapers resources;
  - Track newspapers resources in UC, paying attention to overlap and gaps;
  - Coordinate creation of guides, tutorials, and web pages;
  - Recommend mechanisms for improving access to news resources in all formats;
  - Monitor new methods used to distribute, access, and preserve newspapers resources.
- Recommendation: If CDC determines that maintaining a separate comprehensive news list is desirable, we recommend CDL explore the creation of a newspapers title list leveraging existing functionalities. (R6)

### *Improving Access*

- Utilize the Shared Cataloging Program to distribute cataloging records for individual newspaper titles; acquire catalog records from vendors whenever possible. (R7)
- Make availability of cataloging records one of the negotiating/evaluation points in acquiring new databases. (R8)
- Provide cataloging records for LexisNexis Academic news titles; although this database is not technically a Tier 1 resource, in practice it is; every campus subscribes. (R9)
- Employ subject headings that allow online catalog users to locate and limit to news resources, and which allow libraries to extract bibliographic records for news resources for other access mechanisms beyond the catalog. (R10)
- Continue to pressure Factiva to improve its structure and interface in order to be able to provide title-specific links. (R11)
- Encourage creation of a news meta-search portal, either by CDL or by one of the campuses using the tools that CDL develops. (R12)

## II. Introduction and Charge

In April, 2006, the UC Collection Development Committee appointed a Task Force on CDL-Licensed News Resources, noting that

The UC libraries subscribe to a number of online news sources that provide indexing and full text coverage, either selectively or comprehensively, for newspapers in the US and throughout the world. While UC would like to maintain the most comprehensive coverage possible, it is also desirable to minimize duplication of coverage, if possible, to allow resources to be spent on other digital resources. The goal of [the] Task Force is to provide an assessment of UC's online news resources to clarify historical and current coverage and content overlap among resources, highlight features of specific resources that are either barriers or ease use, and provide recommendations for the most comprehensive and economical news coverage for UC.

The Task Force was asked to include the following in its report:

1. Inventory (Excel spreadsheet) showing data, and brief summary/analysis of findings.
2. Statement of whether the current news collection meets UC needs. If not, identify the content gaps and recommend alternative or additional news sources for potential licensing. This might include both commercial and open access sources.
3. Clear recommendations to CDC for options to reduce news coverage overlap, if present, between system-wide resources.
4. If possible, recommendations as to how CDL or UC libraries might better enable end users to discover news resources.
5. If possible, recommend ongoing mechanisms for maintaining an understanding of our news resources and consulting with the appropriate bibliographer groups about changes.

The Task Force on CDL-Licensed News Resources members are Kay Collins, UC Irvine; Harold Colson, UC San Diego; Elliot Kanter, UC San Diego; Linda Kennedy, UC Davis (Chair); Wendy Parfrey (CDL liaison); and Lucia Snowhill, UC Santa Barbara (CDC Liaison).

Electronic news resources are invaluable multidisciplinary resources that serve an enormous range of purposes: they are a key resource for ready reference, providing quick access to topical, current and historical information. News sources are an important source of primary material for historical research, and there are increasing numbers of historical newspapers available as a result of commercial and public digitization efforts. Newspapers are an important resource for economists, political scientists, sociologists and other researchers. Newspapers can reveal the daily life of a community; ads, comics and images illuminate popular culture over the years. News resources are a vital resource for area studies. Relatively few print news publications are indexed. Digital access significantly improves access by bringing all this information to users' fingertips.

A substantial amount of UC collections funds are devoted to news resources each year: approximately \$288,000 in annual subscriptions for the titles we examined, excluding *LexisNexis Academic* (news and legal resources, estimated at over \$300,000) and EBSCO *Regional Business News* (paid for as part of *Business Source Premier*). In 2008 we will have an additional \$126,320 in access fees due annually for *ProQuest Historical Newspapers* and \$48,700 for *ProQuest Current Newspapers*. The backfile/perpetual rights/prepaid annual fees investment is about \$1.5 million, excluding Early American newspapers (paid for as part of a package that incorporated *Early American Imprints* and the *U.S. Congressional Serial Set*). It should be noted that some of the titles, such as *Factiva* and *LexisNexis Academic*, considered in among these news resources, are not pure news resources.

The perception has developed that there is a substantial amount of duplication among UC-licensed news resources, and the Task Force attempted to determine if that was true. We should

note that no one set out to purchase duplicative resources. It should be noted that some databases—*Factiva* and *GenderWatch*, for example—were not acquired primarily for their news content. It is difficult to coordinate the purchase of new resources because of their multiple audiences, multidisciplinary nature, and the inclusion of news sources in databases with broader scope. Often multiple bibliographer groups need to be consulted; although sometimes no selector groups other than perhaps the general reference group is consulted.

As originally charged, the Task Force was to look at a suggested list of electronic news resources. After beginning our work and reviewing the suggested databases, we made some changes to the list of databases to be examined. The list below contains the databases we included, and the rationale for excluding some of the others:

Included:

- *Access World News* (NewsBank): included
- *Factiva*: title list reduced to its news content
- *LexisNexis Academic*: title list reduced to its news content.
- *Ethnic NewsWatch*, *GenderWatch*, *AltPress Watch* (ProQuest Diversity Databases). These have been combined into ProQuest's *Diversity Databases*, with the exception of *Ethnic NewsWatch*, a History, which was treated separately.

Not included:

- *Business Source Premier* (EBSCO), including *Regional Business News*: Business Source Premier was eliminated as it is not a significant source of newspapers. We retained but did not analyze *Regional Business News* (a number of its 99 titles are not newspapers but business journals).
- *Expanded Academic ASAP*: eliminated because of lack of comprehensive newspaper content.
- *Left Index* is a bibliographic database, not a source of full-text newspapers.

UC's libraries license significant historical news resources; upon consultation with CDC, we included the analysis of the following databases in our charge:

- *America's Historical Newspapers*, 1690-1876 (formerly Early American Newspapers)
- *ProQuest Historical Newspapers*: New York Times, Los Angeles Times and Wall Street Journal)
- *Accessible Archives Historical Newspapers* (*Pennsylvania Gazette* and early African American newspapers: a Tier 2 for which the majority of campuses have some or all of its segments)

Appendix B is a spreadsheet providing summary data for each of the databases analyzed.

### **III. Methodology & Analysis**

The Task Force created an inventory of the databases to be examined, providing basic comparative data, and examined each of the selected licensed news resources in detail. We conducted an overlap analysis of the titles and created a spreadsheet of the news titles included in the various databases. We briefly surveyed someone on each campus to surface concerns about acquiring and accessing news resources before we developed our recommendations on ongoing mechanisms for handling news resources in UC. We talked with Heather Christenson from CDL about Google News Archive and MetaLib, and discussed a variety of mechanisms for improving accessibility of news resources. We looked at current and future trends in news access, and considered the relationship between ongoing digital news acquisitions, digitization of news resources, and our existing print and microform news resources.

**Database descriptions:** In addition to the spreadsheet included as Appendix B, members of the Task Force prepared analyses of each the major licensed resources; these are included as Appendix C. The analyses address a variety of parameters for each database: value of the database, coverage, nature of content, extent of overlap with other resources, search and display and other user interface characteristics, cataloging issues such as availability of individual catalog records, open URL capability, browser compatibility, and reliability and stability.

**Overlap analysis methodology:** CDL paid for a three-month subscription to Serials Solution (recently acquired by ProQuest) and arranged a one-month free trial of Ulrich's Serials Analysis. Both products were useful but neither offered a perfect solution for newspaper comparison and analysis.

Serials Solutions is a fairly new product built on a knowledge base of 106,000 serials titles from over 1300 database vendors, with a high degree of normalization and standardization of the data fields and frequent updates with publisher data. Its Overlap Analysis tool was easy and intuitive to use. Overlap reports generated ten points of comparison, including number of unique titles, percentage unique titles, full overlap and partial overlap (based on years of coverage for individual titles within each database). We were able to select combinations of two or more newspaper databases to analyze, and the results are presented later in this report. The major flaw in the product was lack of any ability to import/export to Excel (we had to work with Serials Solutions staff to load several manually built Excel files that contained only newspaper titles). The Task Force concluded that the Overlap Analysis "was promising in concept but not in execution." Serials Solutions has informed CDL that the next version of the product to be released in Spring 2007 will have a robust import/export feature and report generator.

The second tool we tried was Ulrich's Serials Analysis System which contains an Overlap Analysis feature based on Ulrich's Web service. Despite this depth and breadth of available data, including current publisher pricing, the overlap reports are very basic. In comparing databases, only four fields are generated: total number of titles, number of unique titles, number of titles in common, and number of titles not in each database. The lack of any indication of years of coverage by title makes this a shotgun approach. On the pro side, import/export to Excel for further analysis is very easy in Ulrich's product.

The Task Force also discussed the use of ExLibris's SFX Knowledgebase for overlap comparison analysis. Margery Tibbetts confirmed that this is possible only for target databases such as Access World News that consist entirely of newspaper titles. Since there is no title-level identifier in SFX for newspapers, aggregated databases with a mixture of other content (e.g., from EBSCO or Thomson/Gale) cannot be compared for newspaper title overlap.

**Spreadsheet of newspapers titles:** CDC requested that the committee identify all of the titles contained in our existing licenses. The committee created a spreadsheet of all of the titles, recognizing that it represented only the titles available at the point in time that we created the spreadsheet (and the newspapers lists themselves were of varying dates). It was a challenge to identify the titles that represented newspapers only, since titles are not always descriptive of the nature of the resource, and because some of the databases, such as *Factiva* and *LexisNexis* contain thousands of titles that are a mix of newspapers, magazines and journals. Since we could not download titles from Serials Solutions into Excel, we relied on lists (in varying formats) from the various vendors that we then merged into one file. The spreadsheet we created provides an indication of the current situation, but should by no means be regarded as a definitive document. The large file (2.6 MB) is not being distributed but is available from Lucia Snowhill or Linda Kennedy. If CDC agrees, we can mount the database on the CDL password protected site, along with pages from the Serials Solutions' overlap analyses.

The Task Force concluded that manually maintaining such a database of titles as a separate effort on a continuing basis would not be feasible. However, there are other potential options, if

the CDC feels that such a list would be valuable as a separate product on a continuing basis, or if it would be useful to generate such a list periodically or for a special project:

- The SFX KnowledgeBase is able to produce an A-Z list. Per Margery Tibbetts, all major newspapers packages are implemented in SFX (*Readex/Newsbank, Factiva, LexisNexis, EBSCO, and ProQuest*) and linking is activated for participating campuses (as well as possible with the quality of the metadata). The newspapers show up at the title level in SFX's A to Z lists. Margery is still testing and resolving a few problematic linkings with those packages. If titles could be sorted by newspaper or journal or other form of serial, that would be very useful.
- To the extent that the Shared Cataloging Program (SCP) catalogs the individual titles in newspapers packages, Melvyl can output a list making use of subject headings such as "online resources" combined with the "[place]-newspapers descriptor."
- The evolving functionality of Serials Solutions and Ulrich's Serial Analysis System may provide another option for a consolidated newspapers title list in all packages, as may UC's electronic resources management system when it is in place.

*Recommendation:* If CDC determines that maintaining a separate comprehensive news title list is desirable, we recommend CDL explore the creation of a newspapers title list leveraging existing functionalities. (R6)

## IV. Findings

### IV. A. Overlap of News Resources

Our analysis showed that there was far less overlap than anticipated, a level of overlap that seemed justifiable to the Task Force because of the overall richness and diversity of the news offerings available, and the differing characteristics and purposes of the licensed databases that contain news.

Many of us had an expectation that there must be significant overlap among *Access World News*, *LexisNexis Academic* and *Factiva*. Acknowledging there were many non-news titles in *Factiva* and in *LexisNexis Academic*, we attempted to create a subset of news-only titles for *Factiva* and for *LexisNexis Academic* and then compared the three databases.

Comparing the news-only titles using Serials Solutions, we found a majority of the titles in each database were unique:

<i>Access World News</i>	76.8 % unique
<i>Factiva</i> newspapers	76.1 % unique
<i>LexisNexis Academic</i> newspapers	63.8 % unique

Much of the overlap was partial; among all the databases, full holdings overlap was very low. For those titles where there was overlap, *Factiva* usually had longer runs.

*Regional Business News* (via EBSCO) titles, which consisted of specialized business news publications, were 72 per cent unique when compared to *Factiva*, and 89 percent unique when compared to *Access World News*. They were 67 percent unique when compared to *LexisNexis Academic*.

*GenderWatch, AltPressWatch, Ethnic NewsWatch* and *Ethnic NewsWatch: a History* have almost no overlap with the other aggregator newspapers products, and very little among themselves.

*ProQuest Current and Historical Newspapers* were compared separately (apart from Serials Solutions), and the overlap is indicated in the table below. The ProQuest historical coverage, back to the beginning of the papers, is unique, and includes browsable, searchable page images. The coverage overlap with any text version is no more than five years. The current versions of the *New York Times* and *Wall Street Journal* have almost a complete duplication (except for January 1984 of the *Wall Street Journal*, which is included in the coverage of the historical image version).

	ProQuest	LexisNexis Academic	Access World News	Factiva
Los Angeles Times	1881-1985 (image) 1985+ (text)	latest 6 mos	1/5/2006 + (text)	
New York Times	1851-2003 (image) 1980+ (text)	1980+ (text)		
Wall Street Journal	1889-1989 (image) 1984+(text)			2/1984+ (text) Abstracts and selected text 6/1979-1/1984

Among our historical news products, there is almost no overlap. The only duplication we noted was that fifteen years of the *Pennsylvania Gazette* overlap in *America's Historical Newspapers* with the 82 years provided (for the campuses that have licensed it) in *Accessible Archives'* version.

**What could we eliminate?** We are currently receiving a reduced price from *Access World News* for the international content. However, that will probably change at the next renewal, and UC could consider discontinuing the international newspapers module and retaining the California and U.S. newspapers, for which NewsBank's coverage is the most comprehensive among the vendors we examined. We would need to examine the overlap for the *Noticias* portion of Spanish language newspapers. The international and vernacular news coverage of *Factiva* is superior, but not all campuses subscribe to *Factiva* and its search platform needs an overhaul to make it user friendly.

For relatively current years, the overlap of the *New York Times* between *LexisNexis Academic* and *ProQuest Current Newspapers* suggests that we could discontinue the *ProQuest current New York Times* subscription. Although not a Tier 1 license, all campuses subscribe to *LexisNexis Academic*. The current *Wall Street Journal* coverage in *Factiva* overlaps with the coverage from *ProQuest Current Newspapers*. *Factiva* is also currently including images in addition to text; the images are inserted at the end of the text. Members of the Task Force expressed reservations about canceling either title, first because of the benefits of accessing both historical and current versions with the same interface (although one can't search across current and historical versions simultaneously, they are linked in the same user interface) and secondly (in relation to *Wall Street Journal*), because *Factiva* is not a Tier 1 subscription (7 campuses subscribe at present). Cost savings for the cancellation of one or both of these titles would depend upon whether their current costs represent package pricing.

*Recommendation: We recommend that the CDC consider reducing duplication in coverage of the New York Times and Wall Street Journal, with a view to the reservations mentioned above. (R1)*

Campuses should evaluate their local needs for subscriptions to two tier 2/tier 3 databases: *LexisNexis Academic* and *Factiva*. The news content of *LexisNexis Academic*, which continues to be heavily used, is not as robust as either *Factiva* for foreign news in the vernacular language or as comprehensive as *Access World News* for U.S. and California news. However, *LexisNexis*

*Academic* is easier to use than *Factiva*, contains more translated material, and may be more suitable for undergraduate and general users who would not use *Factiva*. Also, the news resources in *LexisNexis Academic* are bundled with key legal resources. It is up to individual campuses to determine their needs for legal resources and the journal literature contained in *LexisNexis Academic*. Some campuses have considered substituting *Factiva* and *Westlaw* (which allows one to subscribe separately to its legal and news content) for *LexisNexis Academic* (although some campuses will require both). (In July 2007, as part of a complete platform redesign, the base *LexisNexis Academic* service will provide at no additional charge its Country Analysis and Company Dossier modules, now sold only as supplements; this enhancement of the multifaceted base service will further complicate substitution decisions.)

*Recommendation:* We recommend that campuses address duplication of domestic and international news resources consistent with their local needs. (R2)

**Relationship of electronic news resources to other formats:** When surveyed, librarians from across the UC system felt there was a need to maintain some microfilm collections. We do not own the perpetual rights to the full content of the CDL licensed news sources, except for three high-profile titles, *New York Times*, *Los Angeles Times* and *Wall Street Journal*. For those years that we do own the ASCII digital versions, none of the graphics, photographs or classified ads that history researchers find so important are included. Journalism students and others evaluate the whole newspaper layout and content. It will be important to continue to monitor the extent that digital full text images becomes available for newspapers. With rights to perpetual access, it may become possible to rely on electronic versions and a limited number of preservation microfilm copies in the UC system.

#### **IV. B Current News Resources: Gaps in Coverage and Future Needs**

In sheer numbers and at first glance, the news resources available to subscribing UC institutions from our “Big Three” aggregators – *Access World News*, *Factiva*, *LexisNexis Academic* – are quite impressive, perhaps even staggering. A combined title tally reveals that these products alone carry digital runs of some ten thousand newspapers, wires, newsletters, transcripts, and other products of global journalism, published in some two dozen languages from more than 130 countries. And unlike the various free publisher and search engine sites that carry many of these same sources, our licensed versions provide the sophisticated search options, deeper content archiving, and liberal usage rights so essential to academic applications.

Upon closer inspection, however, the imposing façade of digital news power from across the globe fades and weakens. Within the most important current news categories, namely domestic and foreign newspapers, the Big Three do not provide full representations of any dailies whatsoever. Far from being JSTOR-like repositories of complete and browse-ready digital copies of the published press record, the Big Three are actually massive yet hit-and-miss amalgamations of so many millions of separate articles, nicely searchable, but presented in mere text form without pictures, tables, graphics, page layouts, and even some entire stories from the original paper issues. For all their convenient search and display prowess across so many news publications, in the end these aggregators provide us with nothing that can fully replace traditional microfilm as a credible guarantor of the domestic and foreign press record.

There are geographical and language gaps within our content façade as well. As can be expected, the United States is the predominant country of origin for our electronic newspapers and news titles, followed by Europe, then Asia and Latin America, with Africa and the Middle East trailing badly. Many countries in the developing world are represented in our packages only by selected BBC press translations, perhaps supplemented with intermittent feeds from a local English-language daily or two. Although *Factiva* is notable for its vernacular Russian, Chinese, and Japanese sources, most foreign language news content comes in Western European languages (principally Spanish, French, German) only, and Arabic content is completely lacking

in the Big Three services. UC should continue to be alert to finding foreign news resources, particularly with coverage in geographic areas weak in our current resources. In fact, the CALMIDEAST consortium did recommend some Arabic language news resources in their request to JSC.

The Berkeley and San Diego campuses receive additional vernacular news content not found in the Big Three products. These products are candidates for future acquisition:

- The *ISI Emerging Markets* (<http://site.securities.com/>) service specializes in general, economic, and financial news feeds from around seventy countries in Asia, Latin America, Central and Eastern Europe, the Middle East and Africa. Chinese, Russian, Arabic, and other non-Western scripts are supported.
- *NewspaperDirect* and *PressDisplay* ([www.newspaperdirect.com/](http://www.newspaperdirect.com/) and <http://www.pressdisplay.com/>) provides an instant global newspaper room with color image replicas of over 400 dailies published in 66 countries and 37 languages. The service carries a 60-day rolling archive of content that supports issue browsing, keyword searching, and page printing.

ProQuest and EBSCO offer alternative full-text national newspapers products to NewBank's Access World News. ProQuest provides enhanced indexing to selected titles; however, we don't see a significant advantage in a change from *Access World News*, which has a very strong base of California and U.S newspapers.

#### **IV. C Historical Newspapers: Gaps in Coverage and future needs**

Historical newspapers are an essential primary source for research in a wide spectrum of disciplines. They are a staple for scholars and students in history programs, but also for historical research in political science, area studies, public policy, literature and the arts, ethnic, gender and urban studies, science studies, economics. Interest in them is likely to be strong at every general UC campus.

Unlike contemporary newspapers in digital form, which are in part a byproduct of their own production process, most historical newspapers have been archived only in microfilm and the process of digitizing them, as OCR text and/or page images, has been labor intensive and very expensive. Except for major papers, in large part the coverage of 18<sup>th</sup> and 19<sup>th</sup> and earlier-20<sup>th</sup> century newspapers depends on often fragmentary paper archives that survived to be microfilmed.

CDL holdings of historical newspapers are almost exclusively for the U.S., and have major and obvious gaps in terms of both period and location. Historical coverage for the rest of the world is virtually nonexistent. Although the Task Force is not in a position to evaluate and prioritize all of the historical newspaper titles and packages currently available, or analyze relative merits of their interfaces, we are prepared to make some recommendations based on content.

*Recommendation:* We make the following observations and recommendations regarding historical domestic news sources to complement those already licensed (R4)

**ProQuest Historical Newspapers:** The purchase from ProQuest of the back files of three "quasi-national" newspapers, from their earliest issues in the 19<sup>th</sup> century until the late 20<sup>th</sup> century--*Los Angeles Times*, *New York Times*, and *Wall Street Journal*--was an important start. UC should consider acquiring four other key regional newspapers in ProQuest's historical inventory: the *Washington Post*, *Chicago Tribune*, *Atlanta Constitution*, and *Boston Globe*. We recommend that acquisition of those four be negotiated (in priority order as listed). Two other historical papers are also currently available from ProQuest: the *Christian Science Monitor* and *Hartford Courant*.

Campus feedback should be solicited on desirability and priorities for adding those titles. It should be noted that the Atlanta, Chicago and Washington papers have **no** duplication by the digital collections from Readex and Thomson Gale mentioned below, while the others overlap only for the 19<sup>th</sup> century.

**Readex America's Historical Newspapers (AHN):** CDL has acquired only Series I, which includes about 750 newspapers from 24 states in the Eastern third of the US, the great majority of which are from states along the Atlantic coast. Coverage begins in the 18<sup>th</sup> century and ends in the 1870s. Many papers have only scattered issues; others span many years. Series II and III, released in summer 2006, expands coverage to several hundred additional titles from the rest of the US, including Alaska and Hawaii and seven papers in California. Dates go through the end of the 19<sup>th</sup> century and for several important cities, all the way to 1922. The advantages of acquiring *AHN* Series II and III include the logic of expanding a "collection" seamlessly using the same interface, as well as the capability of cross searching with our other components of the Readex *Archive of Americana* collection. We recommend that CDL pursue this expansion of our AHN holdings.

**Thomson Gale Nineteenth Century U.S. Newspapers:** This new product (released Summer 2006) is the closest equivalent to Readex *AHN* (*Series II & III*). It also includes several hundred titles, in all states (seven in California), and dates of coverage also range from scattered months to many years. As the name makes clear, this collection is largely limited to the 19<sup>th</sup> century, although some titles do extend a few years into the 20<sup>th</sup>. Comparison with Readex *AHN* (based on examination of a spreadsheet provided by Thomson Gale) shows only about 10 percent overlap. Contents of both collections are therefore essentially unique. The effectiveness of their platform for historical full text/page image content has already been demonstrated in Thomson Gale's *Eighteenth Century Collection Online*. It should be considered for acquisition, either in addition to expansion of Readex *AHN*, or as an alternative to that expansion if an agreement cannot be reached with Readex.

**London Times:** This key newspaper of the English-speaking world spans two centuries, from 1785-1985. CDL acquired a portion through Chadwyck-Healy's *Historical Newspapers Online*, but only 1785-1870, and for those years the database includes only the articles linked to Palmer's Index. Thomson Gale offers the complete historical span through 1985, using the same interface as its *Nineteenth Century U.S. Newspapers*, and, in response to past and current recommendations by bibliographer's groups, CDL has negotiated for it in the past. It would fill an important gap if some or all of this content could be made available.

**Accessible Archives:** This is an early digitizing product which, among other historical content, includes several 19<sup>th</sup> century African-American Newspapers, selected content from Civil War Newspapers, and the 18<sup>th</sup> century *Pennsylvania Gazette*. It was acquired as a Tier 2 by several campuses, and the segments held vary even among those campuses. The interface is very basic (simple keyword search, text only, no browseability by year or issue). The content is unique, but specialized. This would be a good candidate for consideration by an appropriate bibliographers group on whether or not expansion to Tier 1 is advisable.

**Other Commercial Products:** At least two other commercial products devoted to historical newspapers are currently available and would be worth further investigation and analysis. However on initial examination they appear to have a lesser degree of quality control and offer no form of perpetual access. Their criteria for selecting newspapers and dates of coverage are unclear.

- ***Paper of Record* (paperofrecord.com):** This is a project of Cold North Wind, a Canadian company which began with wholesale digitizing of the historical microfilm archives of Canadian newspapers, branched out into extensive coverage of Mexico, and has begun including newspapers from other countries. *Paper of Record* claims more than 8 million pages digitized. Papers can be searched but not browsed. The process of

digitizing has been much faster and cheaper than products such as those by Readex and Thomson Gale. One known tradeoff is that content is scanned across the whole newspaper page, rather than by article, significantly reducing the precision of searching. Subscriptions are available both for individuals and institutions.

- **NewspaperARCHIVE.com:** Claiming almost 50 million newspaper pages online (the home page has a display constantly counting up), this collection is a project of Heritage Microfilm and is marketed first to individuals, with an emphasis on gift pages, birthday front pages, and genealogical research. Institutional subscriptions are also sold and free access is available for K-12 schools and public libraries. It may also be available at no cost for non-profit colleges and universities. A very large number of newspapers from the United States are included, with every state well represented. There is also a much smaller selection of titles from Canada and the U.K.

**Historical Collections from the Rest of the World:** Apart from the London *Times* and *Paper of Record*, we haven't identified any historical newspapers collections with significant content outside of North America. Identifying, evaluating and recommending acquisition of such products would be an appropriate assignment for a bibliographer group that could gather input from all UC campuses.

#### **IV. D Open Access/Public Domain Digital Newspapers Collections**

Most of the historical newspapers content being contemplated through commercial vendors is from the 18<sup>th</sup>, 19<sup>th</sup> and early 20<sup>th</sup> centuries, and therefore in principle could be viewed as in the public domain. Naturally, CDL and the UC Libraries are concerned with spending substantial amounts of money to buy resources that are theoretically freely available. Realistically, though, there are still great barriers and probably years of effort before we will see digitization of all this content, on robust and effective platforms, and with consistent and well designed interfaces.

Projects to create online newspapers archives are underway in many states, with a variety of funding and organizational models, and on different platforms. Examples range from the Brooklyn Public Library's implementation of the Brooklyn Daily Eagle (<http://uclibs.org/PID/56847>) to statewide consortial projects such as the Utah Digital Newspapers (<http://www.lib.utah.edu/digital/unews/index.html>) and the Colorado Digital Newspapers (<http://www.coloradohistoricnewspapers.org/>). The site of the U.S. Newspaper Program (<http://www.neh.gov/projects/usnp.html>) would be a good starting point to explore what other digital archive projects are under way. Coordinating access to these and other projects and archives could be an important project for the UC Newspapers Bibliographers Group whose creation we have recommended in the next section.

For our own state, we hope that current pilot projects in California can proceed and believe that CDL could play a strong role among other players in helping move towards a full scale "California Digital Newspapers" archive.

A note on Google News Archive: This latest effort by Google may show promise as a way of discovering both free and licensed news content, complete with open access E-Links comparable to Google Scholar. At this stage, however, it is far too preliminary to judge the results. Google News Archive appears to work most effectively with current content on publisher's websites (including some magazines and other publications, not just newspapers). It has a good navigation structure for moving by period back in time. However almost all of the older content appears to be from the newspapers' own archives, requires direct payment, or involves licensed access to the aggregator or database being searched. Unless this approach changes, Google News Archive will be a discovery tool, not a substitute for purchase or subscription. It will be interesting to see how Google News Archive affects newspaper publishers and news database aggregators.

A significant amount of current news coverage is on the web, frequently from a newspaper's own web site. These resources are suitable for daily browsing but generally do not provide complete coverage of the newspaper. Archiving policies vary as to how much is preserved and how it is searched. Most newspapers do not provide free access to archived articles.

## V. The Collection Development Process

CDC requested that we “recommend ongoing mechanisms for maintaining an understanding of our news resources and consulting with the appropriate bibliographer groups about changes.” We informally interviewed someone on each campus, either the chief collection development officer or another experienced bibliographer. We asked how the campus makes decisions on new newspapers packages, or expansions to existing packages, and how the newspapers package selection process could be improved, whether campus or statewide.

Local decision-making did not seem to be a problem, although a lot of consultation was required. Major priorities were reducing duplication and managing costs and managing the collections for continual cost savings. In making decisions, collection development officers felt they needed more information about the packages: information on duplication and overlap, including how a package might duplicate microform subscriptions. The evaluation process was considered labor intensive. One campus suggested we consider a systemwide subscription to Serials Solutions or Ulrich's serials management software in order to help evaluate and manage subscriptions.

Several campuses mentioned a concern about permanent archives and perpetual access to ensure access to major newspapers in order to help identify subscriptions that we might cancel in both print and microform.

Several suggested that broader consultation on news resources was needed, and that it would be desirable to establish another bibliographer group, or add responsibilities for news resources to an existing group. It has been difficult to consult separately with many groups on overlapping resources, and expertise is required to evaluate the news packages. Having a specific news contact on each campus was recommended by one respondent. We also need a better way to track and evaluate trials.

Based on this input, and our own experience and discussions, the Task Force on CDL-Licensed News Resources recommends the creation of a newspapers bibliographer group. The group would have one lead person with newspapers expertise for each campus, but other campus bibliographers could be part of the group and be subscribed to the list server. The bibliographer group could fill many roles. Communication would be a major function: communicating with bibliographers on a campus, communicating with other bibliographer groups, providing background and comparison information on newspapers proposals, making recommendations on newspapers packages (although other bibliographer groups could also make recommendations and advocate for specific packages of interest). The new group could track newspapers resources to which UC subscribes, coordinate creation of guides and tutorials for inclusion on the CDL instruction site, recommend good examples of web pages and other mechanisms from improving access to news resources.

This new model of a bibliographer group could apply to other multidisciplinary groups—reference bibliographers, for example. The government information librarians already operate in a similar fashion, although they do not have a “lead bibliographer” on each campus.

*Recommendation:* Create a new, separate newspapers bibliographer group to monitor news resources in all formats. (R5) This group could be charged to:

- Coordinate communication among bibliographer groups regarding news resources;
- Provide background, analysis and comparisons for newspapers proposals;

- Make recommendations for newspapers resources;
- Track newspapers resources in UC, paying attention to overlap and gaps;
- Coordinate creation of guides, tutorials, and web pages;
- Recommend mechanisms for improving access to news resources in all formats;
- Monitor new methods used to distribute, access, and preserve newspapers resources.

## VI. Enhancing Newspapers Access, Cooperative Activities

Another Task Force charge directed us to “make recommendations as to how CDL or UC libraries might better enable end users to discover news resources.” In our informal campus surveys, we asked how the campus libraries have enhanced access to news resources via cataloging, web resources, etc.

**Cataloging:** Every campus noted the importance of providing individual cataloging records for newspapers titles in aggregate news databases and including records in the online catalogs. Enhancing newspaper catalog records to bring together all versions of news titles was also mentioned. More descriptive cataloging was suggested, with holdings for microform and paper holdings as well. Including UC e-Links in search engines such as Google News and Yahoo is highly desirable. Some systems, such as UCLA and Irvine, include all cataloged newspaper titles as well as the news databases and major individual titles like the *ProQuest Historical Newspapers*. These lists are often generated via subject headings applied to cataloging records for full-text licensed news resources in all formats.

*We recommend the following:*

- Utilize the Shared Cataloging Program to distribute cataloging records for individual newspaper titles; acquire catalog records from vendors whenever possible. (R7)
- Make availability of cataloging records one of the negotiating/evaluation points in acquiring new databases. (R8)
- Provide cataloging records for *LexisNexis Academic* news titles; although this database is not technically a Tier 1 resource, in practice it is; every campus subscribes. (R9)
- Employ subject headings that allow online catalog users to locate and limit to news resources, and which allow libraries to extract bibliographic records for news resources for other access mechanisms beyond the catalog. (R10)
- Continue to pressure Factiva to improve its structure and interface in order to be able to provide title-specific links. (R11)

**Web pages:** Each campus has at least one page directing users to news resources (see Appendix E), and have included news databases in their lists of databases and subject guides (a variety of subjects because of the broad applicability of news resources). Maintaining these web pages is highly labor intensive, according to one campus, and another noted the need to coordinate the resources on web pages with cataloging information. There was a concern that by looking in one place or another, users might miss information. The range of campus web pages include:

- Guides to news databases
- Guides to news resources in all formats
- Lists of newspapers by subject area or geographic area
- A-Z lists of newspapers (digital/and/or print or microform)
- Course-specific research guides
- Guides and tutorials in the use of specific databases with news resources

Some of the campus pages are dynamically generated from the library’s program for managing electronic reference resources. Some of the sites do not include all of the major news databases,

we noted, perhaps because the resources are not coded for news. *Factiva* is frequently overlooked, as is *GenderWatch* and *Ethnic NewsWatch*. One of the best campus newspapers web sites is UC San Diego's page at <http://sshl.ucsd.edu/collections/news.html>. Generated from the library's Sage program for managing electronic resources, the main newspapers page covers resources for U.S. News (subcategories: Directories, Aggregated News, National News, Alternative News, Analysis/Opinion, Blogs and Directories), California News, San Diego News, and International News (subcategories for global and regional resources), U.S. Historical News. All of the listings are for online resources, except for the separate list of Early American historical newspapers.

**Library Instruction:** Campuses have integrated news resources into a variety of library instruction activities, such as:

- Introductory library instruction (news resources are often easy to use, topical, with broad subject coverage)
- Classes on electronic news resources
- Course-related instruction as appropriate
- Historical newspapers as primary resources
- Workshops for students on how to find foreign news/news from home

Although some campuses have a well-developed user awareness program that they will want to continue to employ, some campuses may wish to make use of collaborative resources developed with the help of members of a newspapers bibliographer group. Potential activities for members of a systemwide newspapers group include:

- Creating and maintaining a general news resources web page for the UC Libraries, or one that can be adapted for use by the various libraries.
- Collaborating with specific CDL resource librarians on guides and tutorials to the use of licensed news resources for posting on the CDL's Instructional Materials web site: <http://www.cdlib.org/inside/instruct/>
- Creating a template for an instructional guide that can be adapted for incorporation in course-related instruction.

**Federated searching for news articles:** Federated searching (or "metasearching" or "cross-database searching") is under development at CDL using the ExLibris MetaLib software to develop several separate instances of federated searching applications: an Earth Sciences NSDL portal, a European Integration portal, and Undergraduate Portal. The choices reflect the availability of grant funds to develop these specific applications. News resources, as full-text rather than bibliographic databases, present different opportunities for federated searching. We hope news databases will be included in the European integration and undergraduate portals in particular. CDL is working with an API version and X-server and will make the code available for campuses to use.

*Recommendation:* The Task Force strongly supports the creation of a news search portal, either by CDL or by one of the campuses using the tools that CDL develops. (R12)

**Service models:** One trend in reference service is a consolidation of service desks and a sharing of general reference duties. News databases are essential to general reference. Excellent and comprehensive instruction in news resources for all library reference staff is essential. The same guides and tutorials developed for instructing library users can be used to train library reference staff.

We noted one new model for news collections and services. The Pattee Library at Penn State created a News Room (1) to merge the various formats of news. The innovative service includes a wide selection of current newspaper titles, a collection of thirty news magazines, licensed digital news resources via library workstations, the library's historical microfilm collection, and television

news shown continuously on three wall-mounted 50" plasma screen televisions. One set is devoted to international programming, with the sound on, the other sets have the sound muted and headphones are available. See the Pattee Library's web page for more information: <http://www.libraries.psu.edu/newsandmicroforms>.

## VII. Future of Newspapers

The manner in which news is created and distributed is changing and has been the focus of much recent discussion in newspapers, television, radio and on Internet blogs. Print sources are rapidly losing subscribers for the paper but picking up revenue with their electronic products. Revenues to support the news industry are going away from paper and into new electronic forms.

There is increasing use of innovative methods of distributing news that allow people to select the news they want:

- Web-based, PDF, keyword searchable newspapers from around the globe. (The ProQuest Press Display product even offers a spoken word version of newspaper text)
- RSS feeds abound for various specific topics or papers
- Meta-searching across newspapers resources, e.g. Google News plus publisher arrangements to deliver full text for a price to the user
- News offered by Internet service providers
- Podcasts
- News via Cell phone
- TV and radio news with podcast and digital versions, e.g. PBS which offers pod-cast and digital versions—via its own website, and via *LexisNexis Academic*

At the same time, newspapers are starting to collaborate with some of their Internet rivals. Recently Yahoo entered into a partnership with seven newspapers publishing companies (176 newspapers) to collaborate on classified ads. A few weeks earlier, Google announced a pilot project to sell newspaper ads on behalf of 50 newspapers.

For researchers, there is a significant value in the local community coverage provided by newspapers and the preservation of these publications in microform and via aggregator databases. The free-ranging news resources mentioned above are incomplete, scattered and preserved only erratically. Obtaining current news will probably not be a problem. Preserving quality news resources for future researchers may become more difficult. UC will be in the best position to anticipate and respond changing resources and user needs by establishing a news sources bibliographer group that can monitor and advise in the area.

---

(1) Cheney, Debora, Jeffrey Knapp, and Robert Alan, et. al. "Convergence in the Library's News Room: Enhancing News Collections and Services in Academic Libraries." *College and Research Libraries*, 67(2) (September 2006): 395-417. Link via Library Literature: <http://vnweb.hwwilsonweb.com/hww/jumpstart.jhtml?recid=0bc05f7a67b1790ee8f5e118a2d98293a5a89d55f8b9ebd70f97dc64ec3f444696f0c03d153da10e&fmt=P>

SOPAG/CDC Task Force on CDL-Licensed News Resources  
**Report to the Collection Development Committee**  
University of California Libraries, December 13, 2006

**Appendix A: Committee Charge**

April 17, 2006

**TO:** Linda Kennedy, UC Davis  
Harold Colson, UC San Diego  
Elliott Kanter, UC San Diego  
Kay Collins, UC Irvine

**FROM:** Lucia Snowhill, Chair, UC Collection Development Committee (CDC)

**RE:** Charge for the SOPAG/CDC Task Force on CDL-Licensed News Resources

SOPAG and CDC are appointing you as members of the SOPAG/CDC Task Force on CDL-Licensed News Resources. Linda Kennedy will chair the Task Force. I will also serve as a member of the Task Force as the JSC representative, and Wendy Parfrey will be the CDL contact for comparative licensing and business issues. The Task Force charge follows below.

Thank you in advance for your participation on this Task Force. Please let me know if there are any questions. We look forward to your reports.

**VII. Charge**

**VIII. SOPAG/CDC Task Force on CDL-Licensed News Resources**

The UC libraries subscribe to a number of online news sources that provide indexing and full text coverage, either selectively or comprehensively, for newspapers in the US and throughout the world. While UC would like to maintain the most comprehensive coverage possible, it is also desirable to minimize duplication of coverage, if possible, to allow resources to be spent on other digital resources. The goal of Task Force is to provide an assessment of UC's online news resources to clarify historical and current coverage and content overlap among resources, highlight features of specific resources that are either barriers or ease use, and provide recommendations for the most comprehensive and economical news coverage for UC.

**Methodology:**

- Identify all CDL-licensed news resources, online only (not print or microfilm)
- Identify which resources have indices, abstracts, full text content
- Identify individual titles or sources contained within each resource
- Identify dates of coverage for each source

- Assess completeness of content (e.g. does it include all editions; is it cover-to-cover including advertisements; variation in coverage—some comprehensive, some selective)
- Assess the reliability/stability of content included (do titles move in and out of the resource regularly, or is the title list stable?)
- Assess licensing features, e.g. perpetual access
- Assess method of presentation (e.g. simple pdf, HTML, ranges of pages, extent it is searchable)
- Assess unique platform features where there is overlap of content across vendors
- Highlight any barriers to use, e.g., do users need special training to print a single item; is the item you print easily identified as to paper, page, date; if the vendor provides multiple papers, can these be searched simultaneously?; limits to the number of hits? It would be helpful to have these ranked in some way, even something as simple as major problem vs. minor annoyance.

**The Task Force should include in its report:**

6. Inventory (Excel spreadsheet) showing data, and brief summary/analysis of findings.
7. Statement of whether the current news collection meets UC needs. If not, identify the content gaps and recommend alternative or additional news sources for potential licensing. This might include both commercial and open access sources.
8. Clear recommendations to CDC for options to reduce news coverage overlap, if present, between system-wide resources.
9. If possible, recommendations as to how CDL or UC libraries might better enable end users to discover news resources. (e.g., Recommendations of “best practices” for campus library web pages that list news sources? Are there instructional materials that may be useful? Do we need a news metasearch?)
10. If possible, recommend ongoing mechanisms for maintaining an understanding of our news resources and consulting with the appropriate bibliographer groups about changes.

**Selected CDL licensed news resources, including Tier 1 and Tier 2 resources, current and historic, national/international/local news, English and other languages:**

- Access World News (NewsBank)
- Business Source Premier (EBSCO), including Regional Business News
- Ethnic NewsWatch, Left Index, etc.
- Expanded Academic ASAP
- FACTIVA
- LexisNexis Academic (NOTE: locally licensed by campuses)

**Task Force members:**

Kay Collins, UC Irvine  
Harold Colson, UC San Diego  
Elliott Kanter, UC San Diego  
Linda Kennedy, UC Davis (Chair)  
Lucia Snowhill, UC Santa Barbara (JSC Representative)  
Wendy Parfrey (CDL Contact)

**Timeframe:** Progress report due July 1, 2006.  
Final report due September 1, 2006.  
Deadline extended to December 2007

Please submit your reports to CDC.

SOPAG/CDC Task Force on CDL-Licensed News Resources  
**Report to the Collection Development Committee**  
University of California Libraries, December 13, 2006

**Appendix E: Campus Newspapers Web Pages**

UC Berkeley  
News Databases  
<http://sunsite2.berkeley.edu:8088/ERF/servlet/ERFmain?cmd=searchResType&resTypeId=17>

UC Davis  
Newspapers and Current Events  
<http://www.lib.ucdavis.edu/ul/research/subjects/index.php?heading=39>  
Newspaper Collection  
<http://www.lib.ucdavis.edu/dept/govinfo/newspaper/>

UC Irvine  
News & Newspapers  
<http://www.lib.uci.edu/online/subject/subpage.php?subject=news>

UC Los Angeles  
News & Newspapers  
<http://www2.library.ucla.edu/search/1584.cfm?ty=40>  
Reference E-resources  
<http://www2.library.ucla.edu/search/960.cfm>  
Reference collections Online (YRL) Newspapers  
<http://www.library.ucla.edu/yrl/referenc/rco/newspap.htm>

UC Merced  
Database and Article Searching  
<http://library.ucmerced.edu/databases/index.html>

UC Riverside  
Newspaper Indexes & Databases  
<http://library.ucr.edu/?view=find/newspapers.html>

UC San Diego  
News  
<http://sshl.ucsd.edu/collections/news.html>

UC San Francisco  
No separate news page; can search by the word news or newspaper in the GALEN databases list. <http://www.library.ucsf.edu/db/>

UC Santa Barbara  
Newspapers  
<http://www.library.ucsb.edu/eresources/epubs/newspaper.html>

UC Santa Cruz  
Newspapers  
<http://library.ucsc.edu/Zope/eresources/bytool/Newspapers>