

ADVISOR REVIEWS—STANDARD REVIEW

Catholic News Archive

doi:10.5260/chara.22.2.21

Date of Review: September 10, 2020

Composite Score: ★★★★★ 1/2

Reviewed by: Sharon Kabel
Independent Researcher

<sharontherese2@gmail.com>

Abstract

The Catholic News Archive (CNA) is a freely available unique digital collection of American Catholic Newspapers. With an impressive range of advanced search options and search filters, and a friendly mobile layout, it encourages serious research as well as easy browsing. Every newspaper page is scanned with optimal character recognition (OCR). The lack of export options is a serious hindrance to heavy-duty research, which is unfortunate given how extensive and interdisciplinary the content is. The uncorrected OCR is counterbalanced by powerful search options, diverse result display options, and user-corrected OCR.

Pricing

The CNA is open access, with the exception of commercial use which requires permission from the copyright holder(s). At the time of publication, users will be required to create a free account if they wish to download full-text PDFs.

Overview

The CNA is a unique, fully searchable, full text digital collection of American Catholic newspapers, an overlooked and underrepresented content area. Every newspaper page, including the masthead, is scanned with OCR—something that could be useful for newspapers that have gone through several title changes.

The website has been developed with researchers in mind, with easy mobile use, and free accounts where users can save results to private lists (maximum 50 articles/list, 10 lists/user).

Currently, there are 17 newspaper titles (including the Catholic News Service, which more correctly is a news feed) that covered both national (*The Catholic Worker*) and regional/local (*The Voice*).¹

Broadly, the date coverage is October 1831 to December 2016. Each of the 17 newspaper titles is given its own page, with a brief biography of the newspaper and the issues available in the CNA. Most titles do not have complete runs. They range from *Shepherd of the Valley* (which appears to be a complete run, from 1832-1854), to the *Catholic Commentator* (which carries only 45 weekly issues from 1963).

The CNA is part of a broader project within the CRRA: the Catholic Newspapers Project. Digitization projects within the Project, including scanning newspapers for the CNA, are funded by member and partner institutions, fundraising campaigns,² and grants.³

User Interface

The CNA has a clean interface with several options for browsing content, such as by date, by tag, by diocese, and by title. Advanced searching offers several features that more prominent subscription databases do not even carry. Users can do wildcard searching, as well as proximity searching with any number of words in between—two powerful tools to counterbalance uncorrected OCR.

Users can filter results by decade, and can also set a custom date range, down to the day. CNA will facilitate mass downloads for users (see Figure 1).

More detailed tutorials for advanced search are available under the Help/question mark button to the right of the search bar (Figure 2).

The screenshot shows the Catholic News Archive website interface. The header includes the logo for the Catholic Research Resources Alliance and the title 'The Catholic News Archive'. A navigation bar contains links for Home, Search, Titles, Dates, Dioceses, Tags, and Help. A search bar at the top contains the query 'ecclesiastical year'. Below the search bar, the results are displayed in a table with columns for Publication, Category, and Author. The results are sorted by 'best match first'. The first result is from 'The Catholic Telegraph' (86) and 'Catholic News Service - Newsfeeds' (49), dated 24 August 1966. The second result is from 'The Catholic Transcript' (44) and 'The Monitor' (18), dated 14 January 1897. The third result is from 'Late Publications' (ARTICLE), dated 17 June 1886. The interface also includes a 'Donate' button and a 'Search history' link.

FIGURE 1 Catholic News Archive Search Results for “ecclesiastical year” Search

FIGURE 2 Catholic News Archive Search Bar and Advanced Search Buttons

FIGURE 3 Catholic News Archive Advanced Search Options

FIGURE 4 Catholic News Archive Results Page (Showing Images) for “missal” Search

Under the advanced search options, users can choose when, in what field, and in what newspaper to search (see Figure 3). They can customize how they view their results as well. The default is Text, but Images can be extremely useful when sifting through large amounts of results, to see quickly where and how a search term was used.

The search filters are of varying usefulness. Some metadata fields, like author, are not standardized, so one may see many variations of one author name. Category, or type of work, is a curious one. The options are Article, Advertisement, Illustration, the newly added Letter (which does not appear to apply to already-added content), and Page. Page is typically the largest category, because the Catholic News Service is a news feed, and is therefore indexed on a page level, not type of work. However, there are compensations for this. Users can filter by word count (less than 50, 51-1000, and greater than 1,000), which can give a rough idea of what the type of work is, at least between an advertisement and a full-length article. Just as interestingly, under Advanced Search, users can change search summaries (i.e., how

results appear). The default is Text, but one can also select Images, which will show where your search terms appear on the page (see Figure 4).

Users can save search results to a private list, but 1.) the lists have a cap (50 results per list, 10 lists per user), and 2.) saving a link to a list removes valuable (and unique) metadata (see Figure 5).

A unique feature of the CNA is that the URLs from search results contain the newspaper title, issue date in ISO format (making spreadsheets of links easier to manipulate), your search terms, and any search filters users may have applied. If this link is shared, such as via copy/pasting, the inclusion of the search terms in the URL will mean that the recipient can open the link to find the exact terms highlighted - rather than a bewildering, oversized newspaper page. All of this information, which makes research more transparent, replicable, and shareable, is lost if a result is saved to a private list. In the provided example, the URL is: <https://thecatholicnewsarchive.org/?a=d&d=>

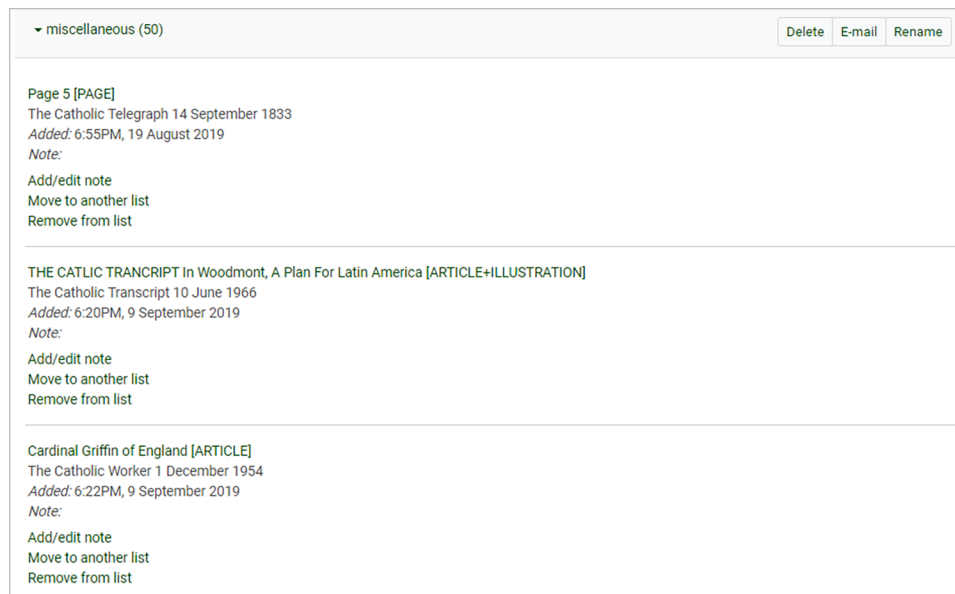


FIGURE 5 Catholic News Archive Example of Results Saved to a User List



FIGURE 6 Catholic News Archive Example of a Newspaper Article with Search Terms Highlighted

ca19650506-01.2.72&srpos=1&e=-----196-en-20--1-byPU-txt-tx-IN-"new+mass"-ARTICLE---1965--. "ca" is The Catholic Advocate, the issue date is 1965-05-06, "byPU" shows that the searcher's results were sorted by publication title, the search terms were "new mass," category was filtered to "article," and date was filtered to "1965" (see Figure 6).

Once users have selected a result, the CNA makes it easy to skim the article in question, with highlighted search terms and with the article itself highlighted. Users can use arrows over the article to browse previous and future issues of the newspaper in question, or to return to their results. Along the right-hand side is a snipping tool, which, if used, takes the user to a new page with their selection, and with easily copy/paste-able citation information (see Figures 7 and 8).

Considering that the CNA claims to be designed for researchers, there is a glaring omission: there is a lack of exporting results. Users can save results to a private list, but 1) the lists have a cap, and 2) saving

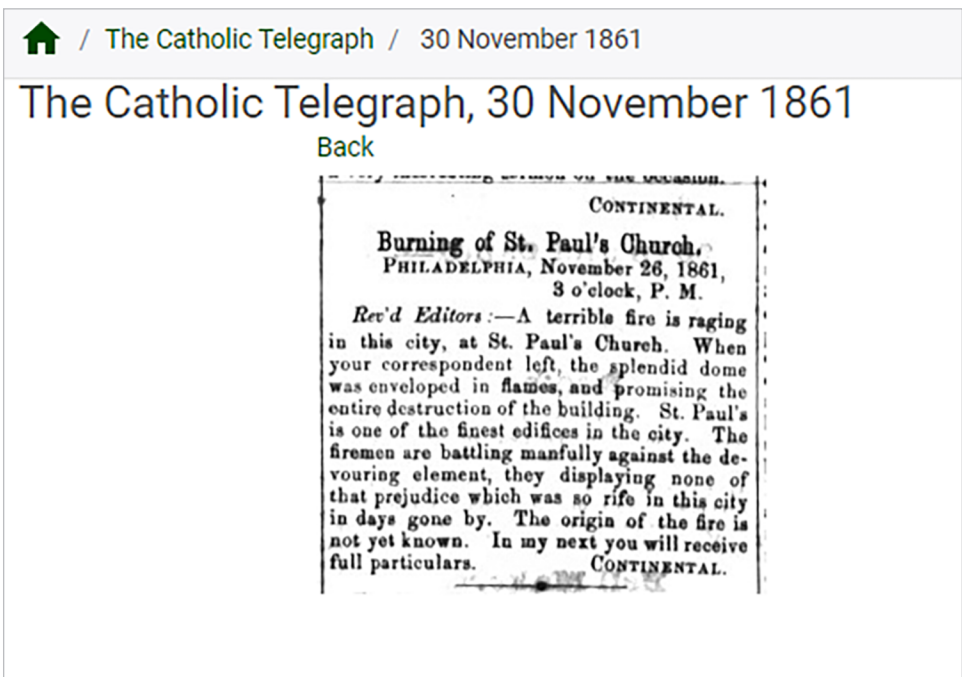
a link to a list removes valuable (and unique) metadata—your search terms and any filters you applied that typically appear in each result will disappear. Users can also share results on social media. However, there is no ability to export the results. If users wish to save results to a reference manager, they will have to do so manually. Zotero offers the option to "capture" web pages, but the result is some random metadata that ends up in the title field, so users would have to manually edit the entire record anyway.

The CNA does not have a public policy or statement on accessibility. According to a report from Veridian Software, which manages the database's content, accessibility to the CNA is consistent with needs and best practices as expressed by the American Foundation for the Blind for people with low or no vision. Newspapers are available as scanned images, PDFs, and HTML.

FIGURE 7 Catholic News Archive Clipping Tool Demo, Before



FIGURE 8 Catholic News Archive Clipping Tool Demo, After



Critical Evaluation

The CNA is a powerful, freely available tool with wide usefulness. Of course students of Catholic history will find much to benefit, Catholic newspapers covered much more than just Church politics. There is

much grist to be made for students of history, religion, theology, philosophy, education, liturgy, and more.

Its advanced search features outshine many more prominent and well-acclaimed digital resources, and its uncorrected OCR is nicely counterbalanced by ongoing user corrections and several ways to view results.

The major drawback of the site is its lack of export options, a serious hindrance for serious researchers. Because the CNA shows continual growth of content and development of its tools, one hopes that this feature will be included one day.

Competitive Products

Catholic newspapers are chronically underrepresented in newspaper collections.

The most obvious competitor is Newspapers.com. which carries none of the seventeen titles in the CNA. Newspapers.com lists only a dozen titles with “Catholic” in the title, none of which is open access.

Chronicling America carries very few Catholic titles. Internet Archive has a few issues of scattershot titles, Readex ARDN carries several Catholic titles, but for most, there are only a few issues. Internet Ar-

Free Text Keywords: newspapers | Catholicism | history | religion | education | interdisciplinary | multidisciplinary

Primary Category: Philosophy & Religion

Secondary Categories: General Reference; History & Area Studies; Multidisciplinary (or interdisciplinary); Sociology, Education, Anthropology, Psychology

Review Type(s): Newspaper or news source, Primary source digital content, Website

Target Audience: Secondary, General public, Undergraduate, Graduate/Faculty/Researcher

Access Type(s): Open Access (OA)



Catholic News Archive Review Scores Composite: ★★★★★ 1/2

The maximum number of stars in each category is 5.

Content: ★★★★★ 1/2

The content is clearly described, neatly arranged, easily browsable, and always growing. Because it is always growing, it would benefit researches for more notice on when new content is added.

User Interface/Searchability: ★★★★★ 1/2

The interface, both desktop and mobile, are robust and user friendly. The one significant drawback is the lack of export options.

Pricing: N/A

Purchase/Contract Options: N/A

chive has a few issues of scattershot titles, Readex ARDN <<https://chroniclingamerica.loc.gov/newspapers/>> carries several Catholic titles, but for most, there are only a few issues. The California Digital Newspaper Collection has no Catholic titles at all. Newspapers.com of course boasts many Catholic titles, but the content is not Open Access.

In terms of topic, scope, breadth, and access, there is no comparable product to the CNA.

Purchase Contract

Usage of and account registration with the CNA is free to the user. Commercial use is not permitted without permission.⁴ CNA will facilitate mass downloads for users upon request.

Authentication

The CNA does not share user search history or account information, and IP addresses remain anonymized.⁵

Author's References

1. "Browse by Title." Catholic News Archive, <<https://thecatholicnewsarchive.org/?a=cl&cl=CL1&e=-----en-20--1--txt-tx-IN----->>>
2. "CRRA Volunteer Recognition." Catholic Research Resources Alliance, <https://www.catholicresearch.org/index.php?option=com_content&view=article&id=133>

Contact Information

Catholic Research Resources Alliance (CRRA)

230 Washington Avenue Extension, Suite 101

Albany, NY 12203

E-mail: <jyounger@catholicresearch.org>

Producer URL: <<https://www.catholicresearch.org/>>

Product URL: <<https://thecatholicnewsarchive.org/>>

3. "Fundraising." Catholic Research Resources Alliance, <https://www.catholicresearch.org/index.php?option=com_content&view=article&id=77>
4. "Terms of Use, Copyright." Catholic News Archive, Updated July 12, 2018, <<https://thecatholicnewsarchive.org/?a=p&p=tou>>
5. "Privacy." Catholic News Archive, Updated December 1, 2017, <<https://thecatholicnewsarchive.org/?a=p&p=craprivacy>>

About the Author

Sharon Kabel is an independent researcher, currently working as a science librarian. She has published research on Catholic data and historical sources in *New Liturgical Movement*, *OnePeterFive*, *New Oxford Review*, and *Catholic Herald*. ■