

Implementing the Catholic Newspapers Program: Collaborative Digitization and Discovery within a Nonprofit, Distributed, Online Organization

Patricia A. Lawton

Catholic Research Resources Alliance, University of Notre Dame, Notre Dame, IN USA
plawton@nd.edu

Alexandra M. Budz

University of Illinois Urbana-Champaign, Graduate School of Library and Information Science, Champaign-Urbana, IL USA
ambudz2@illinois.edu



Copyright © 2014 by Patricia A. Lawton and Alexandra M. Budz. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License:

<http://creativecommons.org/licenses/by/3.0/>

Abstract:

The Catholic Research Resources Alliance (CRRA) <http://www.catholicresearch.net> is a nonprofit membership alliance of institutions collaborating to deliver projects and services in support of its mission "to provide enduring global access to Catholic research resources in the Americas." The immediate focus is creating freely available access to those rare, unique and uncommon research materials held by libraries and archives in North America.

In 2011, the CRRA initiated the Catholic Newspapers Program. The goal and scope of the program is to provide access to all extant Catholic newspapers published in North America. Activities include:

- An environmental scan of North American Catholic newspapers. The goal is to determine what newspapers exist, where they are located, and the extent of holdings and formats. A summer 2012 pilot project at the University of Notre Dame resulted in the identification of more than 800 titles.*
- Development of a directory. Extant Catholic papers will be discoverable through the Center for Research Libraries' International Coalition for Newspapers (ICON) database.*
- Digitization. Current work is focused on the collaborative digitization of twelve priority papers and the identification of shared repository options. Twenty-three digitizing partners, comprising both member and non-member archives, libraries, and publishers, are working together to identify complete runs of newspapers and to bring the collection together in a shared repository of newly digitized and born-digital content, with a continuing commitment to providing access to the papers in perpetuity.*
- Catholic Newspapers Online. An evolving list of digitized or born-digital Catholic newspapers.*

Keywords: newspapers, Catholic newspapers, Catholic Research Resources Alliance, collaborative digitization

1. The Catholic Research Resources Alliance (CRRA)

While no online union list of Catholic newspapers currently exists, collaborative efforts combined with advances in technology may facilitate such a listing in the future. --Charlotte Ames, University of Notre Dame, 1997. -- (<http://archives.nd.edu/cathnews/cathhint.htm>)

Some fifteen years later, Charlotte Ames' vision may be realized, and extended. Advances in technology have made it possible for a network of Catholic archivists and librarians to build a portal of Catholic resources and then to strategize and begin construction of a comprehensive Directory of Catholic Newspapers, thus fulfilling Charlotte's vision of a union list of Catholic newspapers. But wait, there is more. In line with the CRRA mission to "provide enduring, global access to Catholic research resources in the Americas," many of these newspapers will also be available in digital form.

Established in 2008, the Catholic Research Resources Alliance (CRRA) is a nonprofit membership organization of forty-one Catholic colleges, universities, archdioceses and congregations collaborating broadly to deliver projects and services in support of its mission to provide enduring global access to Catholic research resources in the Americas.

CRRA accomplishes its goals through a distributed and highly collaborative organization. Individuals with particular roles and expertise are organized into a number of groups with specific foci. The CRRA Board of Directors oversees overall administration, task forces, committees, and working groups handle specific aspects of the organization's mission,, with committee work coordinated with the assistance of two paid staff members. More about the function and organization of CRRA groups can be found here: <http://www.catholicresearch.net/cms/index.php/crra-groups/>.

From the beginning, the CRRA's primary project has been the development of the Catholic portal (www.catholicresearch.net). The portal contains bibliographic records describing more than 101,000 rare and unique member resources and collections, including over 350 newspaper titles.



The screenshot shows the homepage of the Catholic Research Resources Alliance (CRRA). At the top, there is a navigation menu with links for Home, About CRRA, CRRA Groups, The Catholic Portal, News & Events, Join, and Links & Resources. Below the menu is a search bar with the text "Search the Catholic Portal:" and a "Find" button. To the right of the search bar are links for "Login / Create account" and "CRRA Blog". The main content area features a "Featured Member Collection" section with a portrait of a man and the text "University of St. Thomas Holy Cards as Objets Trouves". To the right of this section is a "The Catholic Portal" section with a red header and text describing the portal's mission and features. Below this is a "Catholic Newspapers Program" section with a red header and text describing the program's goals.

www.catholicresearch.net

While the portal serves an important role in collocating Catholic research materials, portal use studies revealed a common refrain among scholars, students, laypersons, and archivists for access to the actual resources—“the things themselves”—not just to records describing them. Rare and unique materials are of course generally not shared through interlibrary lending programs but digitization can make them available, anytime, anywhere, to anyone.

Catholic newspapers are not rare in the usual sense of the word. However, they are held in scattered locations with unknown holdings, in fragile condition and vulnerable to disappearance. The Scholars Advisory Committee noted the significance of Catholic newspapers to scholars, the physical challenges of using microfilm as well as the clear advantages of the ability to search across digital content as well as concern that many papers were endangered. They recommended digitizing Catholic newspapers for access, use and preservation. The CRRA had received its marching orders: begin immediately to collaboratively digitize Catholic newspapers. From this charge, the Catholic Newspapers Program was hatched.

2. Catholics and Catholic Newspapers

Scholars noted their extensive use of Catholic newspapers which are relevant to a broad array of topics that span many disciplines and touch on many topics, including immigration and adaptation to new environments, charitable work and social justice, school systems and hospitals. Catholic newspapers, defined as Catholic by purpose, audience or tradition, official and unofficial, document not only actions of episcopal elites but also detail activities of local parishes, societies, schools and organizations in their efforts to preserve languages, maintain holy days, and teach sexual morality. They also capture how Catholics at every level embraced elements of American popular pastimes such as sports and variety shows.

But full access to the papers is complicated and can be frustrating: issues are scattered among numerous archives and libraries, holdings remain hidden, and few papers are available in digital form. With advances in technology and the development of a Catholic archival network, the ability to provide efficient, effective, global, and enduring access to the documents of this rich and influential culture is within reach.

2.1. Catholics in the World

*The Roman Catholic Church is both globally widespread and historically long-lived. There are over a billion Catholics now on earth, almost a sixth of the world's population. According to the Catholic Almanac this global church is divided into 221,740 parishes divided into 2,667 dioceses. In addition, the Church operates many schools, colleges, universities, cooperatives, orphanages and other social service agencies.*¹

The numbers vary yet they add up to the same result: Catholics constitute a significant population in the U.S. and in the world. There are approximately 2,000 different religious groups in the U.S., and of these, most Americans have chosen one of twenty-three variations on traditional Christian denominations.² With more than 69.4 million members, the Catholic Church is the largest denomination in the United States, comprising twenty-two percent of the

¹ Pew Research Center's Religion & Public Life Report. The Global Catholic Population 2013: <http://www.pewforum.org/2013/02/13/the-global-catholic-population/>

² Melton, Gordon J. 2010. American Religion's Mega-Trends: Gleanings from Forty Years of Research (ARDA Guiding Paper Series). State College, PA: The Association of Religion Data Archives at The Pennsylvania State University, from <http://www.thearda.com/rrh/papers/guidingpapers.asp> (accessed June 21, 2014).

population. After Brazil, Mexico and the Philippines, the U.S. has the fourth largest Catholic population in the world³

Which countries have the most Catholics now?

10 Countries with the Largest Number of Catholics, 2010

Countries	ESTIMATED 2010 CATHOLIC POPULATION	PERCENTAGE OF POPULATION THAT IS CATHOLIC, 2010	PERCENTAGE OF WORLD CATHOLIC POPULATION, 2010
Brazil	126,750,000	65.0%	11.7%
Mexico	96,450,000	85.0	8.9
Philippines	75,570,000	81.0	7.0
United States	75,380,000	24.3	7.0
Italy	49,170,000	81.2	4.6
Colombia	38,100,000	82.3	3.5
France	37,930,000	60.4	3.5
Poland	35,310,000	92.2	3.3
Spain	34,670,000	75.2	3.2
Democratic Republic of the Congo	31,210,000	47.3	2.9
World Total	1,078,790,000	15.6	100.0

Population estimates are rounded to the ten thousands. Percentages are calculated from unrounded numbers. Figures may not add exactly due to rounding.
Pew Research Center

Pew Research Center's Religion & Public Life Report. The Global Catholic Population

2.2. Catholic Newspapers, National and Local

Catholic newspapers are of two general types: national and local, or diocesan. There are several national Catholic papers in the U.S. such as the [National Catholic Reporter](#) and [Our Sunday Visitor](#).⁴ The majority of Catholic newspapers, though, are those that emanate from Catholic dioceses.⁵ Numbers are not exact for how many dioceses and diocesan papers exist, but there are 195 American dioceses registered with the United States Conference of Catholic Bishops⁶ and 162 diocesan papers currently listed by Our Catholic Neighborhood.⁷

Diocesan papers generally include a mix of local and national news, with national Catholic news stories gleaned from the [Catholic News Service](#) (CNS) newsfeeds, the Catholic equivalent of Reuters⁸. Most papers typically devote one page to CNS content and the rest of the paper to local content. Locally, the diocesan paper serves as a key communication vehicle for the bishop offering the bishop an opportunity to speak with the faithful on a regular basis.

2.3 Reading Catholic Newspapers, Print and Online

"A Catholic Paper in Every Catholic Home" --William Cardinal O'Connor⁹

³ Our Sunday Visitor's Almanac (2014), p. 432.

⁴ The National Catholic Reporter and Our Sunday Visitor are both CRRA priority papers for digitization.

⁵ A *diocese* is a fully organized ecclesiastical jurisdiction under the pastoral direction of a bishop as local Ordinary. Definition from Catholic Online <http://www.catholic.org/clife/terms.php?letter=d> (accessed June 2014).

⁶ <http://www.catholic-hierarchy.org/diocese/lc.html> (accessed June 2014).

⁷ U.S. Diocesan Newspapers <http://www.ourcatholicneighborhood.com/faith/evangelization/media/newspapers/u.s.-diocesan-newspapers> (accessed June 2014).

⁸ CNS newsfeeds are included among CRRA digitization priorities

⁹ A Brief History of the Boston Pilot <http://infowanted.bc.edu/history/briefhistory/#> (Accessed June, 2014)

In 2011, the Center for Applied Research in the Apostolate (CARA) partnered with the Catholic Press Association for a study of Catholic Media Use in the United States.¹⁰ Findings shed an interesting light on the use of the diocesan paper and include:

- The diocesan newspaper in print is second only to the parish bulletin as the most widely used communication tool in the Catholic Church.
- *One in four adult Catholics have read a diocesan paper in print within the last three months* [italics added for emphasis]; only four percent report reading a digital diocesan paper
- Of those who have read a diocesan paper in the last three months, respondents spent an average of seventeen minutes reading the print version and just two minutes reading the digital version (it should be noted that not all dioceses offer a print or a digital version)
- The percentage of Catholics who say they subscribe to a print copy of a daily newspaper for their local area has dropped from 44 percent in 2005 to 36 percent in 2011.
- Seventy-nine percent of readers evaluate their diocesan newspaper or magazine as “good” or “excellent” as a source of news, information, and dialogue about their diocese and the Catholic faith.
- Only sixteen percent of Millennial Generation Catholics say they have a print subscription to a local daily newspaper.
- An anomaly worth noting is CARA’s finding that the most likely readers of the online edition are the oldest generation of Catholics. In this demographic, young Catholics are not moving to online from print as many assume.¹¹

One important factor in the relatively high rate of readership may be that these papers continue to be delivered directly to the home, and subscriptions are generally free. Thus, pushing print product into the home may remain the best way to distribute the paper and may account for this group’s significant use of print vs. online editions. There is some sense that if the paper is there on your kitchen table, one is more likely to read it.

2.4. Catholic Papers of the Past

Of primary interest to the CRRA are the Catholic papers of the past, many of which are scattered, not described or known to the world, and in brittle (or worse) condition. A brief overview of the earliest known Catholic papers and their *raison d’être* reveals a rich legacy of publishing motivated by the desire for spreading the Catholic faith, and attention to social justice, religious freedom, and protection of oppressed minority groups.

CRRA has identified the *Boston Pilot* as a key paper for digitization, the oldest continuing Catholic newspaper in the U.S. The *Pilot* was founded in Boston in September 1829 (as the *Jesuit, or, Catholic Sentinel*) by Bishop Benedict J. Fenwick to serve French Catholic immigrants and a rapidly increasing number of Irish Catholic immigrants to the Boston Area. Anti-Catholic sentiment at this time ran high, and the Bishop felt that the paper could explain

¹⁰ Catholic Media Use in the United States, 2011.

http://c.ymcdn.com/sites/www.catholicpress.org/resource/resmgr/docs/print_study_final_report.pdf. (Accessed June , 2014)

¹¹ Catholic Media Use, 2011, p. 25.

and defend the Catholic Church's teachings while serving as a foundation for newly arrived and uncertain immigrant populations.¹²

When H.J. Deveraux and Patrick Donohue took over the paper in the 1840s, its focus became largely directed toward the Boston Irish and became a major resource for the Boston Irish immigrant community. It presented important information on religious and political matters and provided opportunity for debate on those matters. Early issues of the *Pilot* offer insight into the tumultuous years of a city, a state, and a nation as Americans grappled with the establishment of a government and a political system, changing demographics, and finding a tolerance for new religious and ethnic groups.¹³ The paper also became the vehicle for locating missing loved ones through its “Information Wanted” advertisements.¹⁴

The screenshot shows a web interface for 'BOSTON COLLEGE INFORMATION WANTED'. The title is 'A Database of Advertisements for Irish Immigrants Published in the Boston Pilot'. Navigation links include HOME, ABOUT, HISTORY, RESOURCES, and ADVANCED SEARCH. The main heading is 'Advertisement seeking: Nellie Leary'. The form is divided into three sections: MISSING PERSON'S INFORMATION, SEEKING PERSON'S INFORMATION, and POINT OF ORIGIN INFORMATION. There is also a WORK HISTORY section.

MISSING PERSON'S INFORMATION		SEEKING PERSON'S INFORMATION	
Name:	Nellie Leary	Name:	Margaret Walsh
Gender:	Female	Relation to Missing:	Sister
Age:		Gender:	Female
Earlier Name Used:		Residence:	
Alias:		Multiple People:	Y
Description:		Date of Advertisement:	02-25-1860
Other:		Contact Name:	Howland, Daniel, Esq.
Maiden Name:	Walsh	Contact Address:	RI, E.Greenwich
		Second Seeker:	
		Third Seeker:	

POINT OF ORIGIN INFORMATION	
Home County:	Cork
Parish:	Rathbarr
Townland:	Mil town
Barony:	Ibane & Barryroe
Poor-Law:	Clonakilty

WORK HISTORY	
Ireland Occupation:	
U.S. Occupation:	
Company Name:	
Work Location:	
Labor Union Member:	

Buttons: Back to search results, Begin a new search

“Information Wanted” advertisement seeking Nellie Leary, 1860, *Boston Pilot*

Our Sunday Visitor was likewise created to meet a pressing need. Founded in 1912 to combat anti-Catholicism in the United States, “it weathered the 1920s when the Ku Klux Klan was the largest openly anti-Catholic organization, survived the Depression in the 1930s and Al Smith’s defeat to Herbert Hoover, and emerged from the Second World War anti-Communist. By 1961, it was a newspaper with a million readers.”¹⁵

The importance of Catholic papers to the Catholic and the American narrative is clear; yet ironically enough, these papers have not received their due priority among large newspaper preservation or digitization efforts.

2.5. The Dearth of Catholic Papers in Directories or in Digital Form

Many Catholic newspapers are difficult to access: issues are scattered among a variety of locations, including church libraries or small archives; knowledge of where these newspapers

¹² A Brief History of the Boston *Pilot* <http://infowanted.bc.edu/history/briefhistory/> (accessed June 2014)

¹³ <http://infowanted.bc.edu/record/?recid=25732> (accessed June 2014)

¹⁴ <http://infowanted.bc.edu/record/?recid=25732> (accessed June 2014)

¹⁵ Nancy Stout (2008) Catholic Periodical Literature, *The Serials Librarian: From the Printed Page to the Digital Age*, 53:4, 155-172, p. 12. DOI: 10.1300/J123v53n04_11: http://dx.doi.org/10.1300/J123v53n04_11

are located is often local; the full extent of collections is not represented in major bibliographic databases such as [WorldCat](#); and few papers are available in digital form. In many cases, these newspapers are owned by Catholic dioceses and congregations that may not have access to bibliographic tools such as *WorldCat*. Many papers may be under-described or simply not described at all.

Although massive in scope (approximately 140,000 bibliographic title entries and 900,000 separate library holdings records), the Library of Congress' [Chronicling America/Newspaper Title Directory](#) lists just forty percent of the U.S. 857 Catholic titles identified in the scan.

3. Introducing the Catholic Newspapers Program (CNP)

In support of its mission to provide enduring global access to Catholic research resources, the CRRA in 2011 initiated the Catholic Newspapers Program “to provide access to all extant Catholic newspapers in North America.” Program goals include the realization of three components:

- A fully searchable Directory of North American Catholic newspapers for discovery of all extant North American¹⁶ Catholic¹⁷ newspapers
- Digitization of papers identified as top priorities
- Repository services to provide a single search interface for searching and text mining capabilities across digital Catholic newspapers

With the digitization of and broad access to Catholic newspapers, we imagine the possibilities for searching across historical major and minor Catholic newspapers in pursuit of new connections, knowledge, and potentially, new historical perspectives.

3.1. Populating the Directory: The Environmental Scan

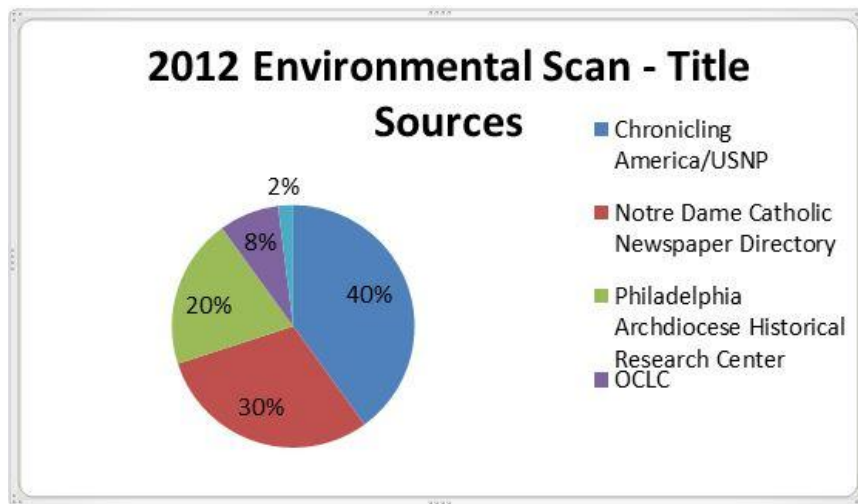
The charge to digitize Catholic newspapers came from our Scholars' Advisory Committee in response to our question of what digital content would be of interest to the broadest spectrum of scholars. Yet we did not know what papers had existed, currently exist, or where they were held. Our first task then, was to conduct an environmental scan of Catholic newspapers to learn where these newspapers are held, how many there are or have been, the integrity of bibliographic records describing them, and metadata elements required for good description.

In the summer of 2012, the University of Notre Dame and the University of St. Michael's College in Toronto, conducted the scan and identified 857 U.S. and 200 Canadian Catholic papers, respectively.

For the U.S. titles, CRRA Digital Projects Assistant, Alexandra Budz mined sources for Catholic newspaper titles, resulting in 857 identified titles, with ninety percent coming from the [U.S. Newspaper Title Directory](#), [Catholic Newspapers in Microform: A Directory of Works at Notre Dame](#), and [Philadelphia Archdiocesan Historical Research Center \(PAHRC\) database](#).

¹⁶ Defined as published in Canada or the United States.

¹⁷ Defined as Catholic by purpose, audience or tradition, official and unofficial.



Sources Consulted for the 2012 Environmental Scan for U.S. Catholic Newspapers

From this initial scan, we learned:

- Many Catholic papers do not have the word “Catholic” in the title or in the description, making simple identification difficult. Examples include the Boston *Pilot*, New Orleans *Clarion Herald* and the San Francisco *Monitor*.¹⁸
- Papers exist in languages other than English, with German topping the list at forty-eight of the 857 titles, followed by twelve in Slovak, and ten each in Polish, Ukrainian, and Spanish.¹⁹
- Title changes are profuse and cataloging practices make it challenging to easily identify the full run of a paper. For example, forty-four different forms of title were identified for just twelve papers. For the casual user looking to identify the complete run, this fact alone could mean failure. The user would need to know all forms of title in order to stitch together a complete run, and cataloging practices dictate that each title has its own bibliographic record, with references to only the previous title, if known.²⁰ To further complicate matters, each title in a different language also requires a separate bibliographic record.

3.2 Identifying the Directory Platform, CRL’s ICON

With newspapers data in hand, we next needed to determine what tool or platform would be best suited to hold the data. In winter 2012, CRRA assembled the Directory Platform Working Group with the charge to identify, explore, and recommend hardware/software solutions for the CRRA Newspaper Directory.

¹⁸ The resource for verification of earlier titles proved to be the indispensable “Catholic serials of the nineteenth century in the United States; a descriptive bibliography and union list” (1968) by Eugene Paul Willging and Herta Hatzfeld. This multi-volume set is a true labor of love. The authors provide rich descriptive information about each of the titles, with equally rich historical context.

¹⁹ See

https://docs.google.com/document/d/1RU453nwUxz4hhRzFgtl4KWqFzjRdO9LMU_wkt0D5gA/edit?usp=sharing for a listing of all identified languages.

²⁰ “There may be multiple bibliographic records to represent a single serial. This is known as successive-entry cataloging; that is, any major change (e.g., some title changes or changes in issuing body in 1xx fields) results in the creation of a new record. These separate records are linked using MARC tags (780/785). “--CMT’s [Serials Training Manual - Yale University Library www.library.yale.edu/catman/serials/training.doc](http://www.library.yale.edu/catman/serials/training.doc)

The Group identified two distinct user groups for the Directory: end-users, particularly scholars, for discovery and access to Catholic newspapers; and librarians, for use as a collection management tool.

The purposes for the Directory were becoming clear. It would provide user-friendly discovery and access to full runs of Catholic newspapers; assist digitization and preservation efforts by tracking which issues and formats were held by which institutions; and track both print and digital copies for preservation. In summary, the Directory would be a tool to:

- Facilitate collaborative collection management
- Improve the quality of existing or future digital copies by facilitating identification of where copies are held
- Enhance records for the best possible user experience
- Include member holdings and beyond to express full runs of titles
- Provide global, easy, access to the records and as possible, to digitized copies

The Group identified seven candidates for the Directory platform and evaluated each against four criteria:

1. Capacity to create and/or use clean data (ideally, MARC records) based on international/national standards
2. A hosted solution that does not require additional CRRA technical expertise or capacity to develop or maintain

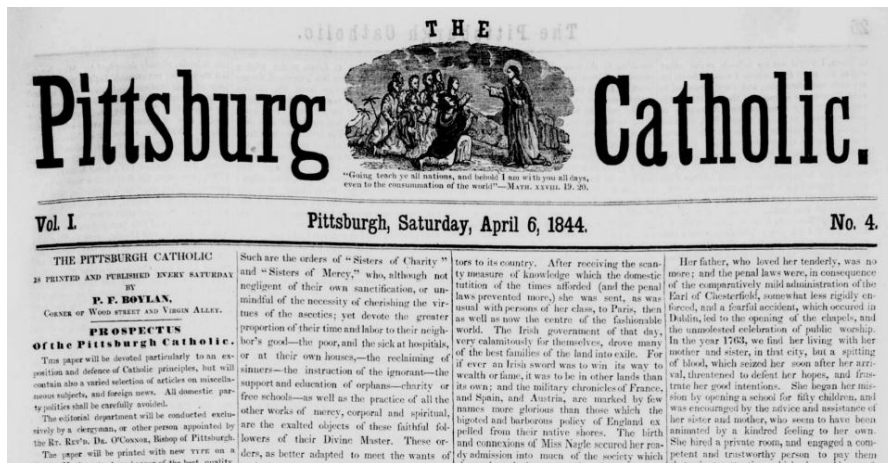
With paid staff of only two, a hosted solution is essential.

3. Ability to contribute to and collaborate with other international and national programs/collections (i.e., a “bigger pond”)
4. Open access to both records and digital content

CRRA values freely available access for all, and strives to ensure this within this program. At the same time, CRRA remains aware of the tension between the time and resources required to complete an ambitious project such as this on our own, and collaborating with groups that have the established infrastructure to expedite the project, on a cost recovery basis.

In spring 2013, the Group recommended the *International Coalition on Newspapers (ICON) database*²¹, hosted by the *Center for Research Libraries (CRL)*, as the home for its CRRA Directory of Catholic Newspapers. Freely available, the *ICON database* is the most comprehensive source of information about significant newspaper collections in print, digital and micro formats. The *ICON database* currently includes more than 300 Catholic newspapers with more added regularly. *ICON database* practices assure that title records will be included to which holdings records from CRRA members (including non-MARC records) can be attached.

²¹ The ICON Database of International Newspapers, a union list of information on the hard copy, microform and digitized holdings of foreign and U.S. newspapers held by major newspaper repositories. It is the largest single database on preservation and digitization of newspapers, containing nearly 30 million issue-level records for 175,000 titles. --<http://www.crl.edu/grn/icon/about>



CRRA Member Duquesne University has digitized the Pittsburgh Catholic, 1844-1981.

4. Digitizing Catholic Newspapers

Answer: 12 newspapers, 23 institutions, 27 individuals, 34 committee members, 2 staff, and 2 consultants.

Question: What does it take to plan the CRRA Catholic Newspapers Digitization Project?

4.1. Determining the Scope

As the environmental scan and Directory work were underway, we were ready to face the central task of digitization.

The Scholars Advisory Committee recommended that priority papers for digitization include major U.S. newspapers and noted the advantage of comparing local or regional perspectives to the national story. Thus, in addition to major city papers, our priority list of twelve includes the *National Catholic Reporter*, *Our Sunday Visitor*, and newsfeeds from the Catholic News Service. Major city papers include: *Boston Pilot*, 1829-1922, *Chicago Catholic New World*, 1852- , *Hartford Catholic Transcript*, 1829- , *Miami Florida Catholic*, 1959-1989; 2009- , *New Orleans Clarion Herald*, 1842- , *New York Catholic News*, 1886-1981, *Philadelphia Catholic Standard & Times*, 1830-2002, *San Francisco Monitor*, 1861-, and the *St. Louis Review*, 1941-²².

While the twelve were deemed a good starting point, all agreed that doing one or more small, manageable “pilot” projects were essential for developing workflows, determining true costs, and establishing timeframes. The pilots will inform the process of working collaboratively across distributed institutions, with a digitization vendor, and Lyrasis, on a small scale. Pilot projects would provide the opportunity to establish and test a framework to ensure a viable, functional framework for a larger digitization project.

We conceptualized the pilot projects as samples of the larger project and therefore identified pilots that will involve more than one institution holding content with a mix of print and microfilm originals. We also looked at for projects where funding opportunities looked promising. A handful of smaller, more manageable (“shovel-ready”) pilot projects were identified, including early years of the *Boston Pilot* (affectionately called “The Pilot pilot”), newsfeeds from the Catholic News Service, and runs of the twelve from various years such as

²² The full listing of the Scope is viewable here:

http://www.catholicresearch.net/cms/files/6113/9534/6031/Scope_of_the_Newspapers_Digitization_Project_website.pdf

the Civil War, pre-1923 content in the public domain, and the pre and post-Vatican II years, 1958-1970.

4.2. Determining Cost Estimates

The scope of the project was clear. Next, how do we do this? Identifying sources of funding was an essential preliminary step and we needed a one-two page prospectus describing the project. Anticipating prospective funders' interests, the key question became: how much will it cost to digitize the twelve priority papers? From this seemingly benign question, Pandora's Box was unhinged.

Where is the content? What institutions hold the "best versions" of film or print from which to digitize?

Having defined the scope of the project, it was then necessary to identify institutions that held the selected years of the twelve priority papers. If we were to digitize, we had to have something from which to digitize, and, we needed to get page counts from "best versions" so we could arrive at estimated costs for digitization. In less than six weeks' we had identified twenty-three different institutions that collectively held the identified content, were interested in participating in the project, and devoted themselves to the tedious task of counting microfilm frames and page counts in order to arrive at cost estimates for a grant proposal. The institutions and individuals stepping forward in this endeavor are the CRRA "Digitizing Partners." The Digitizing Partners include twenty-seven individuals from twenty-three CRRA and non-CRRA institutions including archivists, librarians from colleges, universities, archdioceses, newspapers, seminaries, and the Catholic News Service (The full list of Digitizing Partners is viewable here:

http://www.catholicresearch.net/cms/files/1814/0088/0111/Digitizing_Partners_list_of.pdf)

In the winter of 2013-14, Partners demonstrated commitment to and faith in the project, as well as great fortitude in assessing and estimating over 1.4 million frames/pages of microfilm and print. Our Lyrasis friends along with our many committees and advisors provided gentle guidance through the minutia of the newspaper digitization processes. When it comes to digitizing newspapers, it does indeed take a village.

CRRA is committed to high quality imaging, with the intent that the imaging will be done well and for all time. The Directory Working Group had recommended that imaging and metadata be done in accordance with the [National Digitization Newspaper Program](#) standards, and additionally, with article segmentation and headline correction. This high standard resulted in a cost estimate of nearly two million dollars²³ to digitize and assign metadata to the 1.4 million frames/pages of microfilm and print.

4.3. Online List of Catholic Newspapers

We soon learned that digitizing the twelve would not only be expense, but would take some time. Under the guidance of Noel McFerran, then-chair of the Catholic Newspapers Task Force, the need for immediate access to the full text of Catholic newspapers was addressed with the [Catholic Newspapers Online](#), an online list of Catholic newspapers in digital form. Data from this list will be included in the *ICON* database and continues to grow through suggestions from members and non-members alike. There are currently over seventy-five titles listed, including both born-digital and digitized content from the U.S. and Canada.

²³ Conversations continue with vendors and anticipate lower cost estimates yet to come.

4.4. Building a Shared Repository

Cost estimates for the prospectus would be incomplete without cost estimates for how and where the digitized content would be stored and accessible. Repository solutions and costs were needed, aligned with the goal of a single search interface for searching with text mining capabilities across the digital Catholic newspapers.

Under the leadership of Betsy Post as Chair, The Repository Working Group was convened with the goal to identify and recommend strategies, options, and cost estimates for repository options for Catholic newspapers in digital form. The group's work took place during January 2014 in a series of four teleconferences with a recommendation ready for the February 2014 CRRA Board Meeting.

Five solutions were reviewed ([NDNP software](#), [CONTENTdm](#), [VuDL](#), [Islandora](#), and [Veridian](#)). The Repository Working Group recommends further exploration of the two solutions that most closely met the functional requirements adopted by the Group.

- DL Consulting's [Veridian](#) software is a best-in-class platform for the delivery of digital newspapers. It is a turnkey solution that can be hosted locally or by the vendor; and consortial pricing options are available.
- [Islandora](#), a robust open-source framework, was recommended because task force members favor open source solutions. Islandora has a beta [newspaper solution pack](#) which would need significant development by Lyris, CRRA, [Discovery Garden](#), or the open source community if it were to match Veridian's functionality.

Two additional recommendations were that we digitize for preservation purposes (preservation quality images), and identify and attain one of the four [NDSA levels of preservation](#). Cost estimates for repository solutions are determined primarily by the number of pages processed and hosted. The average annual cost over a period of five years with up to one million pages of content was estimated at approximately \$25,000 per year.

A recap of activities surrounding the development of the project, prospectus, and cost estimates is viewable here:

http://www.catholicresearch.net/cms/files/4913/9896/1175/Timeline_for_Digitizing_Catholic_Newspapers_20140113_rev20140327.pdf

4.5. Building the Infrastructure for Member Digitization

Concurrent with work on digitizing the twelve papers, CRRA encourages and supports members and non-members to seek funding to digitize locally-held Catholic papers. CRRA member Duquesne University was years ahead of our project and serves as a stellar model and inspiration. Duquesne has digitized 1844-1957 of the [Pittsburgh Catholic](#), with 1957-1981 currently being digitized. The Gumberg Library's operating budget funded most of the project, and [Library Services and Technology Act](#) (LSTA) funds coupled with the discounted rate from the [Lyris grant](#) funded a large portion of the project, years 1866-1925. At this writing, Duquesne is working locally to identify funding to digitize the remaining years, 1982-2000.

To assist members in digitizing their local Catholic newspapers, the Newspapers Task Force identified essential documentation that would be valuable to institutions looking to digitize their Catholic newspapers: digitization specifications and digitization agreements. In the 2011 [Survey of Digitized Rare Catholic](#), which queried 230 American Catholic Colleges and

University (ACCU) member libraries, we learned that only one third of ACCU member libraries had digitized any Catholic materials, thus most of the digitization work would likely be a first-time experience. With the intent to make the process simpler, CRRA has convened working groups to draft both digitization agreements and digitization specifications. At this writing, the work of these groups continues.

Digitization Agreements

In a collaborative such as this, where members have content, dioceses and publishers hold copyright, and CRRA is the eventual content provider, drafting a digitization agreement among the various parties has become an interesting endeavor.

CRRA is developing model digitization agreements that identify the responsibilities of involved parties. These include 1) CRRA as a primary agent in fostering digitization, dissemination, and discoverability of digitized content, 2) the copyright owners, 3) the institutions providing the content to be digitized, and 4) digitization vendors. Through these agreements, CRRA will ensure the appropriate permissions are in place for digitizing; for making the content accessible via institutional and CRRA websites; for depositing (publishing) the content in multiple repositories, including but not limited to the institutional and CRRA repositories; for developing a “terms of use” policy on use of CRRA digital collections; and for making the digital content available for other uses in future. CRRA will ask for a copy of the digital content and metadata, permission to provide nonexclusive, perpetual access via the CRRA website and/or repository, indemnification against copyright violations and assurance that holding institutions have similar permission to provide access via their websites and/or repositories and indemnification against copyright violations. The permissions will be developed in the context of the Creative Commons copyright licenses, which provide ways for copyright owners to grant permission to others to use their works in various ways.

5. The Journey of a Thousand Steps Continues

Alone we can do so little, together we can do so much -- Helen Keller

We have made great strides in this process of digitizing Catholic newspapers and have learned about working collaboratively across institutions, and about what is entailed in a digitization program of this ambition. Yet, many questions remain and much remains to be done. We have learned that this is a very expensive proposition, but a worthwhile one. We have a broad community of scholars, librarians, archivists, students, supporting the project, with institutions and individuals willing to invest time and energy. We are well on our way to having an infrastructure in place that will support interested institutions’ efforts to digitize newspapers according to standards and have a place to store and access the content.

Questions abound and mark the path for our next steps. For example, how will having a repository affect our records in ICON? How will the two systems interact and complement one another? What, if anything, about the Directory will change with the implementation of a repository? Will we continue to provide access to the twelve priority papers and how will we continue to make available future issues? How will the repository be maintained and sustained in perpetuity? These questions drive and guide us, and with continued support of this community we will take the next steps in digitizing Catholic newspapers for all. The heart of our work is, after all, to ensure that the history of Catholic life and thought is not lost to future generations, but will live on, for all time.

Acknowledgements

Implicit throughout this project and underscoring and informing every step of the journey is a commitment to service through the provision and preservation of these important resources, guided by a shared sense of collaboration.

It is a cliché, but here, it very true - it takes a village. Deep and heartfelt thanks to all the many participants, too numerous for all to be mentioned. Without each of you, none of this would be possible and for your continued support, time and faith in this endeavor, thank you.