

Using DPLA and the Wikimedia Foundation to Increase Usage of Digitized Resources

Dominic Byrd-McDevitt and John Dewees

ABSTRACT

The Digital Public Library of America has created a process by which rights-free or openly licensed resources that have already been harvested can be copied over into Wikimedia Commons, thus creating a simple path for including those digital collections materials into Wikipedia articles. By meeting internet users where they already are, rather than relying on them to navigate to individual digital libraries, the access and usage of digital assets is dramatically increased, in particular to user groups that might otherwise not have a reason to interact with such digitized resources.

INTRODUCTION

A DPLA-sponsored webinar given by Dominic Byrd-McDevitt, DPLA Data Fellow, and Sandra Fauconnier, GLAM-Wiki Specialist at the Wikimedia Foundation, on April 21, 2020, entitled “DPLA Intro to Wikimedia: Increased Discoverability and Use” introduced a workflow by which records harvested by the Digital Public Library of America (DPLA) could be automatically copied over into Wikimedia Commons with their accompanying metadata.¹ The major benefit of this migration is the ease with which assets can then be added to Wikipedia articles, exposing resources to a large audience of general internet users who might otherwise have no reason to interact with a given repository’s resources. The gains from making digitized resources available in Wikipedia articles is substantial, providing incredibly high usage statistics while requiring very little time commitment to execute the work.

This DPLA project, launched in early 2020, was a result of grant funding provided by the Alfred P. Sloan Foundation and ongoing consultation from the Wikimedia Foundation. DPLA’s interest in designing this system stemmed from an exploration of new ways to increase usage of materials. While previous bulk uploads to Wikimedia Commons by cultural institutions have required technical expertise and steep learning curves in navigating the Wikimedia community, this project was designed to reduce these barriers by taking advantage of DPLA’s role as an aggregator (more information is available at <https://commons.wikimedia.org/wiki/Commons:Partnerships>). With the workflow developed by DPLA’s Technology Team in mid-2020, an authorized bot account on Wikimedia Commons (User:DPLA_bot, https://commons.wikimedia.org/wiki/User:DPLA_bot) uploads assets from DPLA institutions.

Using data provided by contributing institutions, DPLA applies filters to identify eligible items from participating institutions, then for each of these generates wiki markup from descriptive metadata and downloads media files to a server. These files are uploaded by a script that interacts with Wikimedia’s API using the Pywikibot framework (<https://www.mediawiki.org/wiki/Manual:Pywikibot>). By centralizing all of the DPLA network’s Wikimedia Commons uploads, DPLA was able to upload over 2.25 million files (or 2.5 TB of total

Dominic Byrd-McDevitt (dominic@dp.la) is Data Fellow, Digital Public Library of America. **John Dewees** (john.dewees@toledolibrary.org) is Supervisor, Digitization Services, Toledo Lucas County Public Library. © 2022.



storage) from 780,000 items in under a year and a half, becoming the largest single contribution to Wikimedia Commons ever (by more than quadruple the previous record).^{2,3}

This approach to the problem provides a simple on-ramp to participation in Wikipedia for DPLA institutions—especially the many that would otherwise lack the resources or expertise to do so—by requiring of them only those tasks that need their local knowledge, such as describing their own collections prior to aggregation and then making editorial decisions on Wikipedia about them once uploaded.

This project required a chain of partnerships between separate organizations, as well as a variety of metadata and technical requirements that needed to be satisfied: records of digitized resources are created by an organization locally and are then harvested by DPLA. The eligible records in DPLA are then copied over into Wikimedia Commons. Once images are in Wikimedia Commons it is a straightforward process to embed the images in Wikipedia articles, thus achieving the goal of expanded use and access to digitized resources.

John Dewees, Supervisor Digitization Services at the Toledo Lucas County Public Library (TLCPL), was in attendance at the April 21, 2020 webinar and subsequently met with Dominic Byrd-McDevitt on April 30, 2020 to discuss the feasibility of using TLCPL collections as a pilot project for this workflow. The copying of records from TLCPL's repository into Wikimedia Commons was actually started in the course of that first conversation on April 30.

A map from page 96 of the book *Geography of Ohio* (see figures 1 and 2), previously digitized by Dewees, will be used to illustrate the process of how records move through the various tools and platforms discussed.⁴

TLCPL makes digitized resources available through Ohio Memory, a shared CONTENTdm instance for libraries, archives, and museums in Ohio maintained by the State Library of Ohio and the Ohio History Connection.

Home > Toledo Lucas County Public Library Digital Collections > Geography of Ohio

Geography of Ohio

[Show Reference URL](#)

[View Transcript](#)

Use of image is subject to the Terms of Use

☆ Save [Download] [Print] [Share]

Search this record

Page 91
Page 92
Page 93
Page 94
Page 95
Page 96
Page 97

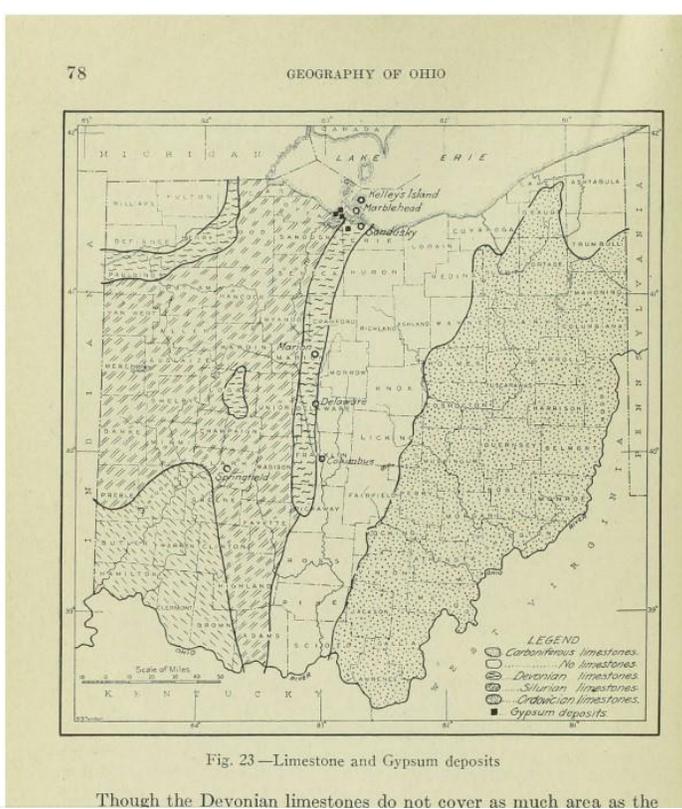


Figure 1. Digitized image of *Geography of Ohio*, page 96, as seen in Ohio Memory.

Object Description

Title	Geography of Ohio
Collection title	Regional History Books
Creator	Peattie, Roderick, 1891-1955
Description	This publication is a Geological Survey of Ohio covering the geography of Ohio. It includes topics about surface features and soil, climate, agriculture, mineral resources, transportation and the development of industries, and the settlement and development for the state of Ohio. This book is part of a series of original research done by the state agency, this one being number twenty-seven in the fourth series of bulletins published. At the time this was published, J. A. Bownocker was the State Geologist of Ohio.
Subject	Physical geography -- Ohio Ohio -- Description and travel Ohio -- History
Time period	1920s
Date of original	1923
Language	eng
Source	R 917.71 Pea (R83G8, b11660922)
Repository	Toledo Lucas County Public Library, Local History and Genealogy
Publisher (original)	Geological Survey of Ohio
Publisher (digital)	Toledo Lucas County Public Library
Rights URI	http://rightsstatements.org/vocab/NoC-US/1.0/
Rights Description	No copyright - United States
Contact Information	Please contact the Local History and Genealogy Department of the Toledo Lucas County Public Library for more information at digitization@toledolibrary.org
Date (digitized)	2018-12-31
Type	Text
Format (original)	Book
Extent (original)	137 pages; 23.5 cm x 15.7 cm
Format (digital)	JPEG
File Name	aalh_geograhyofohio_r91771pea
IIIF Manifest	 https://ohiomemory.org/iiif/info/p16007coll33/115214/manifest.json

Item Description

Title	aalh_geograhyofohio_r91771pea_0096
IIIF Image	 https://ohiomemory.org/digital/iiif/p16007coll33/115137/full/full/0/default.jpg

Figure 2. Record metadata for *Geography of Ohio*, page 96, as seen in Ohio Memory.

DPLA HARVEST

DPLA is a discovery portal that aggregates records of digitized resources from over 4,000 libraries, archives, and museums from around the United States. This creates a single search interface allowing millions of digital records to be searched simultaneously without having to navigate through a wide variety of different digital libraries. The aggregation of these records is accomplished by working with two different types of partners: Content Hubs and Service Hubs.⁵

Content Hubs are either organizations large enough to contribute to DPLA directly, such as the Library of Congress or Harvard Library, or large digital libraries that work with partner institutions of their own, such as HathiTrust or the Internet Archive. Service Hubs, on the other hand, act as mediators between the national aggregation service and individual organizations in states (such as Ohio and its Service Hub, the Ohio Digital Network) or regions (such as Utah, Idaho, and Nevada, who have collectively formed a Service Hub in the Mountain West Digital Library). Service Hubs ensure that the technical and metadata requirements for harvesting into DPLA are satisfied and act as consultants and facilitators to prospective contributors.

As DPLA has grown over time, the metadata requirements and possibilities have also evolved and have varied depending on which Service Hub a contributing organization is working with. The Ohio Digital Network (ODN) is the Service Hub for our example page from Ohio Memory. ODN's metadata requirements for contributors in March 2021 included a title and a standardized rights statement in the metadata application profile for the contributing collection. More information on the DPLA harvest process for the Ohio Digital Network is available at <https://ohiodigitalnetwork.org/contributors/getting-started>.

The nature of these requirements has also evolved since ODN's first harvest in March 2018. Initially, the standardized rights statement was required to be one of the options from [RightsStatements.org](https://rightsstatements.org) but through the work of DPLA and ODN, now [Creative Commons licenses](https://creativecommons.org/licenses) and the [CC0 public domain dedication](https://creativecommons.org/licenses/publicdomain/) can be utilized as well. Standardized rights statements must be formatted as machine-readable URIs rather than textual descriptions.

Finally, the technical backend that supports the harvest of a digital collection is via an OAI-PMH feed. Other hubs operate in very different ways—such as some that actually host all their contributors' collections in a single domain—but in all cases the end result is providing a data set that DPLA can harvest and ingest. Figures 3 and 4 illustrate this process, showing the *Geography of Ohio* represented as a record in DPLA (available at <https://dp.la/item/aaba7b3295ff6973b6fd1e23e33cde14>) with associated metadata.

HOME BROWSE BY TOPIC BROWSE BY PARTNER EXHIBITIONS PRIMARY SOURCE SETS MY LISTS ABOUT DPLA NEWS DPLA PRO

DPLA DIGITAL PUBLIC LIBRARY OF AMERICA Search the collection Search

All items > Geography of Ohio

Please read DPLA's Statement on Potentially Harmful Content.

Geography of Ohio



Cite this item

Add to a new list

View Full Item [↗](#)



Created Date	1923
Description	This publication is a Geological Survey of Ohio covering the geography of Ohio. It includes topics about surface features and soil, climate, agriculture, mineral resources, transportation and the development of industries, and the settlement and development for the state of Ohio. This book is part of a series of original research done by the state agency, this one being number twenty-seven in the fourth series of bulletins published. At the time this was published, J. A. Bownocker was the State Geologist of Ohio.
Creator	Peattie, Roderick, 1891-1955

Figure 3. *Geography of Ohio* as seen in DPLA, specifically focusing on the thumbnail, link to the original record, and initial metadata fields.

Partner	Ohio Digital Network
Contributing Institution	Toledo Lucas County Public Library
Collection	Toledo Lucas County Public Library
Publisher	Geological Survey of Ohio
Subjects	Physical geography -- Ohio Ohio -- Description and travel Ohio -- History
Type	text
Language	English
URL	http://www.ohiomemory.org/cdm/ref/collection/p16007coll33/id/115214 
Standardized Rights Statement	No Copyright - In the United States:  The organization that has made the Item available believes that the Item is in the Public Domain under the laws of the United States, but a determination was not made as to its copyright status under the copyright laws of other countries. The Item may not be in the Public Domain under the laws of other countries. Please refer to the organization that has made the Item available for more information.

Figure 4. *Geography of Ohio* as seen in DPLA, specifically focusing on the remaining metadata fields harvested.

This process achieves the first level of aggregation: harvesting thumbnail images (full-sized images suitable for research are not harvested in this process) and metadata from local digital repositories and making them available for a unified search experience in DPLA. DPLA's aggregation currently contains over 42 million items, with the majority of these containing standardized rights URIs; 18 million items have rights compatible with upload to Wikimedia Commons (as can be seen at <https://dp.la/search?rightsCategory=%22Unlimited%20Re-Use%22>). Once DPLA has access to the records, it is possible for the code authored by DPLA staff to be utilized to then integrate the resources into Wikimedia Commons.

WIKIMEDIA COMMONS HARVEST

Wikimedia Commons is part of the larger network of services and tools under the umbrella of the Wikimedia Foundation. There are a wide variety of different tools available such as Wikidata, a portal for open structured data; Wikipedia, a collaboratively edited open encyclopedia; and

Wikimedia Commons. This last portal uses the same software platform that powers Wikipedia to create an open file and media server that can interoperate with the other tools.

Wikimedia Commons is capable of hosting digital still images, audio files, and video files. Anyone can contribute to this open repository so long as the work is in the public domain or openly licensed. Users may either release a work for which they own the rights under an open license at upload time or may upload any other works by providing evidence in the metadata that the work is out of copyright or openly licensed (more information on copyright and licensing in Wikimedia Commons is available at <https://commons.wikimedia.org/wiki/Commons:Licensing>).

With this in mind, in order for records in DPLA to be eligible for harvest into Wikimedia Commons, they first must have one of the five specific standardized rights statements available at the following links:⁶

- <http://rightsstatements.org/vocab/NoC-US/1.0/>
- <https://creativecommons.org/publicdomain/zero/1.0/>
- <https://creativecommons.org/publicdomain/mark/1.0/>
- <https://creativecommons.org/licenses/by/4.0/>
- <https://creativecommons.org/licenses/by-sa/4.0/>

The URIs above indicate the most recent version of each of the associated copyright descriptions or licenses, though being published under the most recent version is not a requirement for harvest into Wikimedia Commons. While standardized rights statements are not a requirement for contributing to DPLA generally, they are being used as a requirement for Wikimedia Commons upload so that the software has a machine-readable way to determine the compatibility of rights.

Though it is a non-profit educational resource, Wikimedia Commons does not utilize media under fair use or materials only licensed for noncommercial/educational use, in order to ensure its users may reuse the media for any purpose. As a result one thing to keep in mind is that while a given organization may include in their gift or accession agreement a statement that digitized versions of physical resources are allowed to be shared through channels decided by the organization, this does not necessarily extend to Wikimedia Commons users outside your organization, because of the requirement to be able to reuse materials with little restriction past attribution and the need to share alike, depending on the standardized rights statement.

DPLA locates the asset to upload by using URLs explicitly provided by the Service Hub; the URLs can be provided in one of two ways. One is to provide the IIIF manifest URL (via the [IIIF Presentation API](#)), from which the DPLA-developed software queries the manifest for the list of assets which are listed by the presentation API in the form of [IIIF Image API](#) URLs.

The other way the media location can be identified is by providing a list of direct URLs to the media in the field DPLA calls mediaMaster during the initial harvest process. Unlike the IIIF manifest URL, this is a multivalued field that can accommodate a list of URLs. The reason for this approach is to allow any institution to contribute assets via the pipeline, regardless of whether they actually have implemented IIIF in their repository or not. Not all organizations have adopted the IIIF suite of APIs so it is important to be able to provide more than one avenue for Wikimedia Commons harvest.

However, providing a IIIF manifest when and if it becomes available has benefits over the mediaMaster field. It will always be true when queried, whereas the mediaMaster values are only accurate to the last harvest, which may be a month or more out of date.

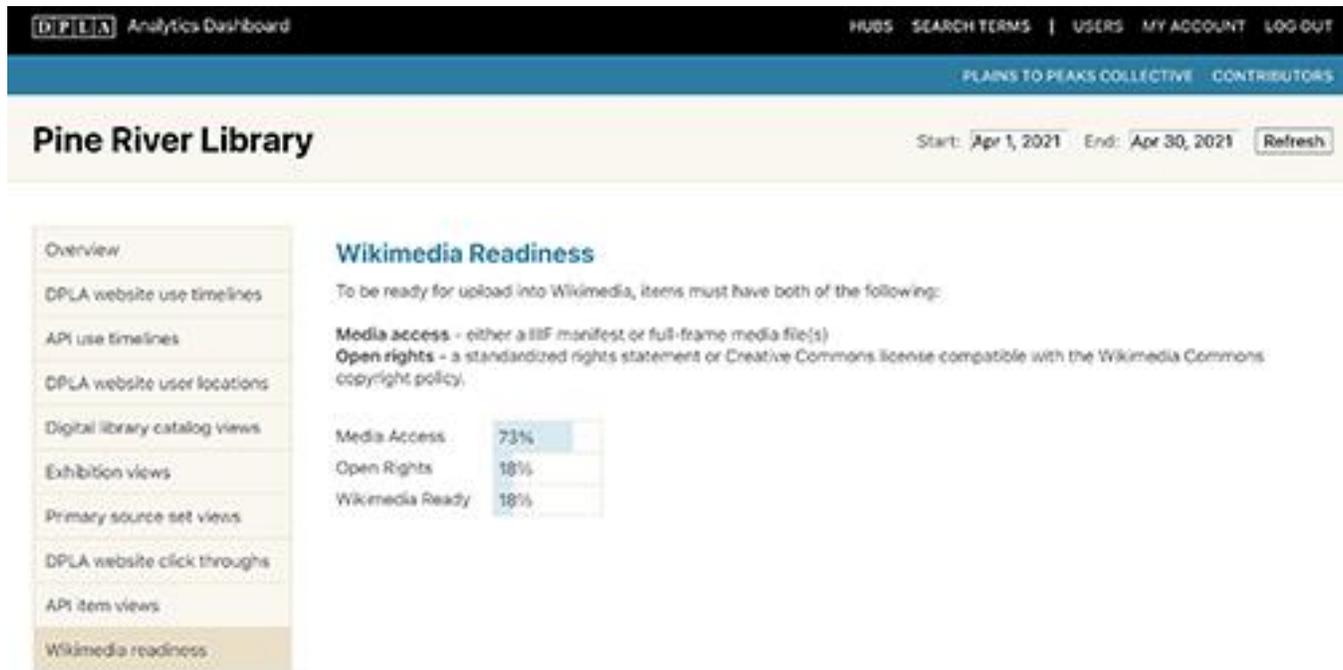


Figure 5. The dashboard developed by DPLA displaying, for Pine River Library, percent of records that have open rights statements and percent of files with media access.

A dashboard has been developed for DPLA Content Hub and Service Hub administrators to analyze how many records in a given collection conform to the standardized rights statement and IIIF API requirements (see figure 5). Harvest of a collection into Wikimedia Commons from DPLA necessitates that all eligible records in the collection be harvested into Wikimedia Commons; it is not possible for a participating institution to hand-curate which of the eligible items will be included. That is, all records in a given collection with the aforementioned standardized rights statements will be harvested into Wikimedia Commons.

An additional signed agreement or memorandum of understanding has not been required between DPLA and participating organizations due to the open nature of the works being transferred. Since the works have been identified as in the public domain or openly licensed, users can already freely use the resources for any purpose they want, so long as it conforms to the appropriate Creative Commons license.

RESOURCE PRESENTATION IN WIKIMEDIA COMMONS

Each portion of the migration process presents the resource in different ways. The original instance of *Geography of Ohio* is made available in CONTENTdm as a complex digital object: multiple images (or more specifically in this case, pages) associated with a single metadata record. DPLA presents this resource only in terms of its metadata along with a thumbnail image of the resource itself; to view the contents of the resource the user is directed back to the original repository for full access to the digital object. The migration process into Wikimedia Commons actually copies the image assets themselves along with the metadata. In this example, both the

image assets and the metadata are drawn from CONTENTdm. Wikimedia Commons is not able to accommodate complex digital objects, and any that are imported via this process are broken out into discrete simple digital objects in Wikimedia Commons, for example, page 96 of *Geography of Ohio* (see figures 6, 7, and 8; [view page 96 in Wikimedia Commons](#)).

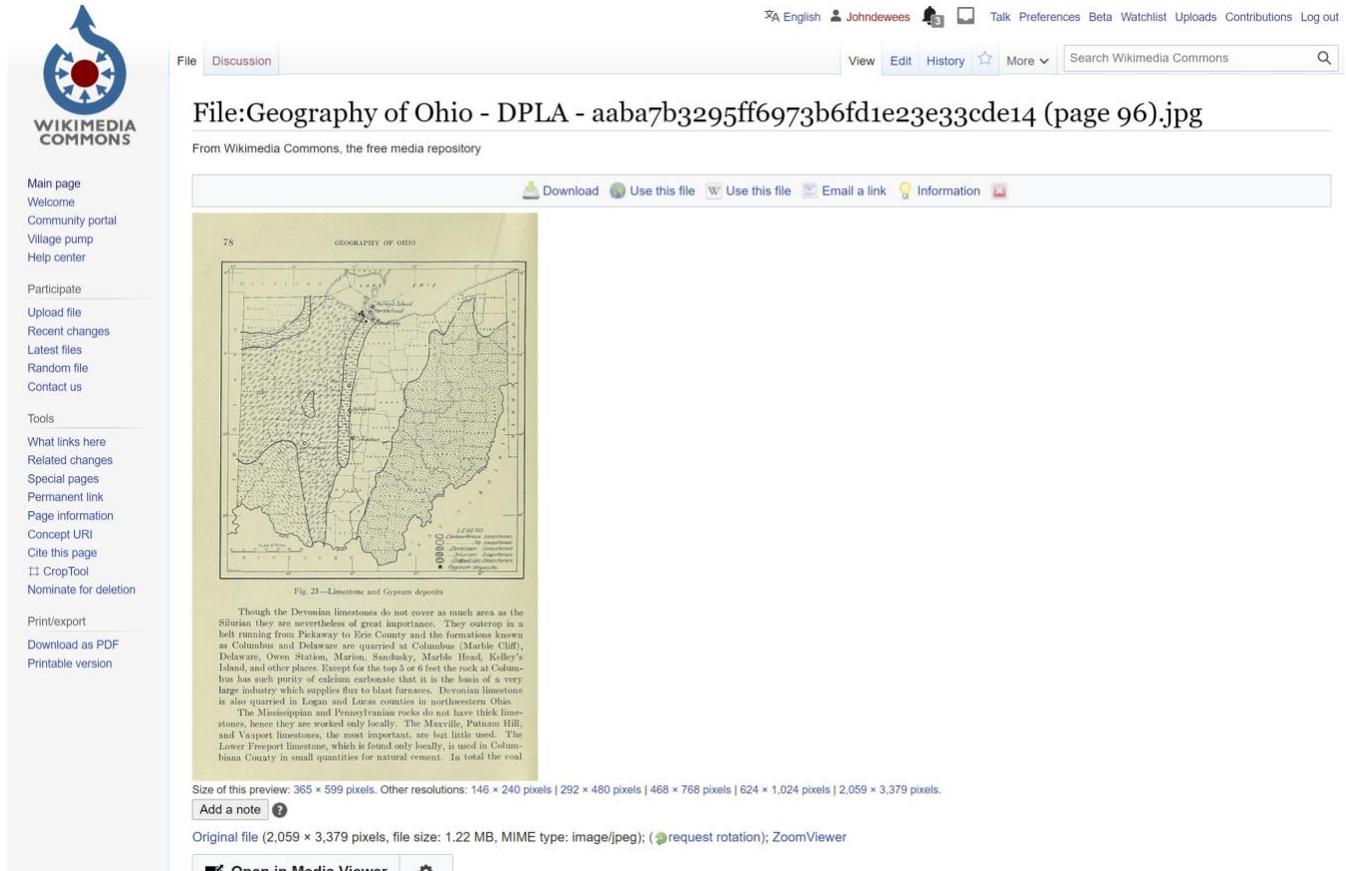


Figure 6. *Geography of Ohio*, page 96, as seen in Wikimedia Commons, with a focus on the file name, image, and viewing options.

Summary [\[edit\]](#)

Geography of Ohio

Creator	Peattie, Roderick, 1891-1955
Title	Geography of Ohio
Description	This publication is a Geological Survey of Ohio covering the geography of Ohio. It includes topics about surface features and soil, climate, agriculture, mineral resources, transportation and the development of industries, and the settlement and development for the state of Ohio. This book is part of a series of original research done by the state agency, this one being number twenty-seven in the fourth series of bulletins published. At the time this was published, J. A. Bownocker was the State Geologist of Ohio.
Date	1923
Collection	Toledo-Lucas County Public Library
Source/Photographer	<div style="display: flex; align-items: center;"> <div style="border: 1px solid #ccc; padding: 5px; margin-right: 10px;"> </div> <div> <p>This file was contributed to Wikimedia Commons by Toledo-Lucas County Public Library as part of a cooperation project. The donation was facilitated by the Digital Public Library of America, via its partner Ohio Digital Network.</p> <ul style="list-style-type: none"> • Source record: http://www.ohiomemory.org/cdm/ref/collection/p16007coll33/id/115214 • DPLA identifier: aaba7b3295ff6973b6fd1e23e33cde14 </div> <div style="border: 1px solid #ccc; padding: 5px; margin-left: 10px;"> </div> </div>
Permission <small>(Reusing this file)</small>	<div style="display: flex; align-items: flex-start;"> <div style="margin-right: 10px;"> </div> <div> <p><i>This media file is in the public domain in the United States. This applies to U.S. works where the copyright has expired, often because its first publication occurred prior to January 1, 1927, and if not then due to lack of notice or renewal. See this page for further explanation.</i></p> </div> <div style="margin-left: 10px;"> </div> </div> <p><small>This image might not be in the public domain outside of the United States; this especially applies in the countries and areas that do not apply the rule of the shorter term for US works, such as Canada, Mainland China (not Hong Kong or Macao), Germany, Mexico, and Switzerland. The creator and year of publication are essential information and must be provided. See Wikipedia:Public domain and Wikipedia:Copyrights for more details.</small></p>
Standardized rights statement	IN THE UNITED STATES No Copyright - United States

Figure 7. *Geography of Ohio*, page 96, as seen in Wikimedia Commons, with a focus on the record metadata.

This file has multiple **extracted images**:

Geography of Ohio - DPLA - aaba7b3295ff6973b6fd1e23e33cde14 (page 96) (cropped).jpg

Geography of Ohio - DPLA - aaba7b3295ff6973b6fd1e23e33cde14 (page 96) (cropped2).jpg

File history

Click on a date/time to view the file as it appeared at that time.

	Date/Time	Thumbnail	Dimensions	User	Comment
current	19:25, 16 May 2020		2,059 × 3,379 (1.22 MB)	DPLA bot (talk contribs)	Uploading DPLA ID aaba7b3295ff6973b6fd1e23e33cde14

Figure 8. *Geography of Ohio*, page 96, as seen in Wikimedia Commons, with a focus on the derivative images created from the original record and administrative metadata.

The filename is programmatically generated and embeds a great deal of information; the following example illustrates the various components of the filename.

Example Filename:

File:Geography of Ohio - DPLA - aaba7b3295ff6973b6fd1e23e33cde14 (page 96) (cropped).jpg

1. The prefix for all items in Wikimedia Commons, “File:”
2. The title of the work, in this case “Geography of Ohio” followed by a hyphen
3. The source of the digital object, universally “DPLA” for this project, followed by a hyphen
4. The unique identifier assigned by DPLA, in this case “aaba7b3295ff6973b6fd1e23e33cde14”
5. In the case of complex objects, the page number, in this case “(page 96)”
6. If the file was cropped using Wikimedia Common’s built-in image editing tool, “(cropped)” will be included between the page number and file extension to indicate the image is a derivative of an original
7. The file format extension, in this case “.jpg”

Even if the complex object being imported is not actually a book, the individual item records in Wikimedia Commons still uses the “(page X)” nomenclature to differentiate the individual objects.

The Summary section of the Wikimedia Commons record displays how the metadata is cross-walked into this environment. The Dublin Core Creator, Title, Description and Date fields are copied verbatim from the local Metadata Application Profile (MAP). To identify the contributing institution, and to differentiate between similarly named institutions, DPLA maintains a JSON file mapping all DPLA institutions with their Wikidata identifiers.⁷ This document also indicates which hubs/institutions are participating in the project at any given time through a true/false field that is toggled when an institution authorizes upload. This enables distinct category pages for each contributing institution and analytics to be tracked and provided to DPLA, relevant hubs, and contributing institutions.

The Source/Photographer field is one of the most important as it ensures that attribution for the contributing institution is clear. The field contains a narrative description of how DPLA facilitated this resource to be available in Wikimedia Commons. It also makes available information on the original contributing institution with links to the record as it is originally displayed (in Ohio Memory in this case) as well as in DPLA. Proper attribution of items was a topic that came up continuously when discussing this project with other organizations, so it should be reassuring to know that credit and direct links back to resources is enabled in this workflow.

The Permission and Standardized rights statement fields leverage the aforementioned URIs to be able to provide information to the user on the copyright status of the work as well as concrete information on how exactly they are able to use it responsibly for their own purposes. Finally, an interesting aspect of this record is the links provided to derivative images. In this case we can see the map displayed on this book page has two cropped derivative images.

USE IN WIKIPEDIA ARTICLES

All of the work described above is in service of one goal: to enable higher usage and exposure of digitized resources in Wikipedia articles. While it is possible to do this work manually, inserting images into articles without being a DPLA contributor or even having a digital repository to speak

of, the automated process is a clear advantage, especially when talking about large collections. For the map on page 96 in *Geography in Ohio*, we can see that the map of limestone distribution in Ohio has been included in an image gallery on the [Limestone Wikipedia article](#) (see figures 9 and 10).

The screenshot shows the Wikipedia article for "Limestone". At the top, there is a navigation bar with the user name "Johndeweese" and a notification "You have a new Talk page message". Below this is the article title "Limestone" and a sub-header "From Wikipedia, the free encyclopedia". The main text begins with a disambiguation note: "For other uses, see Limestone (disambiguation)." The introduction paragraph states: "Limestone is a common type of carbonate sedimentary rock. It is composed mostly of the minerals calcite and aragonite, which are different crystal forms of calcium carbonate (CaCO₃). Limestone forms when these minerals precipitate out of water containing dissolved calcium. This can take place through both biological and nonbiological processes, though biological processes, such as the accumulation of corals and shells in the sea, have likely been more important for the last 540 million years.^{[1][2]} Limestone often contains fossils, and these provide scientists with information on ancient environments and on the evolution of life.^[3]"

Below the introduction, there is a section titled "Limestone" with a sub-header "Sedimentary rock". It features a photograph of a limestone outcrop in the Torcal de Antequera nature reserve of Málaga, Spain. Below the image is a caption: "Limestone outcrop in the Torcal de Antequera nature reserve of Málaga, Spain". Underneath the image is a "Composition" section with the text: "Calcium carbonate: inorganic crystalline calcite or organic calcareous material".

The left sidebar contains various navigation options such as "Main page", "Contents", "Current events", "Random article", "About Wikipedia", "Contact us", "Donate", "Contribute", "Help", "Learn to edit", "Community portal", "Recent changes", "Upload file", "Tools", "What links here", "Related changes", "Special pages", "Permanent link", "Page information", "Cite this page", "Wikidata item", "Print/export", "Download as PDF", "Printable version", and "In other projects".

Figure 9. The Wikipedia article on Limestone displaying the introduction and one image (but not the worked-example image).

Gallery [\[edit source \]](#)

A stratigraphic section of Ordovician limestone exposed in central Tennessee, U.S. The less-resistant and thinner beds are composed of shale. The vertical lines are drill holes for explosives used during road construction.	Photo and etched section of a sample of fossiliferous limestone from the Kope Formation (Upper Ordovician) near Cincinnati, Ohio, U.S.	Biosparite limestone of the Brassfield Formation (Lower Silurian) near Fairborn, Ohio, U.S., showing grains mainly composed of crinoid fragments	A concretionary nodular (septarian) limestone at Jinshitan Coastal National Geopark, Dalian, China	Limestone from Lake Tai, used in gongshi, a Chinese stone art	Folded limestone layers on Cascade Mountain in Provo Canyon, Utah	Fossils in limestone from the northern Black Sea region	Limestone distribution in Ohio, from "Geography of Ohio," 1923

Figure 10. The image gallery in the Limestone Wikipedia article, with the map from *Geography of Ohio* included and seen at the bottom right.

Editing Limestone

Content that violates any copyrights will be deleted. Encyclopedic content must be verifiable through citations to reliable sources.

B I > [Advanced](#) > [Special characters](#) > [Help](#) > [Cite](#)

```

==Gallery==
<gallery>
File:OrdOutcropIN.JPG|A stratigraphic section of [[Ordovician]] limestone exposed in central [[Tennessee]], U.S. The less-resistant and thinner beds are composed of [[shale]]. The vertical lines are drill holes for explosives used during road construction.
File:Limestone etched section Kopefem new.jpg|Photo and etched section of a sample of fossiliferous limestone from the [[Kope Formation]] (Upper Ordovician) near [[Cincinnati]], [[Ohio]], U.S.
File:BrassfieldEncrinite042112.jpg|Biosparite limestone of the [[Brassfield Formation]] (Lower [[Silurian]]) near [[Fairborn]], Ohio, U.S., showing grains mainly composed of [[crinoid]] fragments
File:大连国家地质公园12-龟背石.jpg|A [[concretionary]] [[nodule (geology)|nodular]] (septarian) limestone at [[Jinshitan Coastal National Geopark]], [[Dalian]], China
File:太湖赏石-Rock in the form of a fantastic mountain MET DT208239.jpg|Limestone from [[Lake Tai]], used in [[gongshi]], a Chinese stone art
File:Folded Rock Provo Canyon.JPG|Folded limestone layers on [[Cascade Mountain (Utah)|Cascade Mountain]] in [[Provo Canyon]], [[Utah]]
File:Calcinced fossils.jpg|Fossils in limestone from the northern [[Black Sea]] region
File:Geography of Ohio - DPLA - aaba7b3295ff6973b6fd1e23e33cde14 (page 96) (cropped).jpg|Limestone distribution in Ohio, from "Geography of Ohio," 1923
File:Chalk ("Upper Chalk" Formation, Upper Cretaceous; White Cliffs of Dover, England, southern Britain).jpg|Chalk is a variety of limestone. It is a softer, and more powdery material.
</gallery>

== See also ==
* {{annotated link|Coral sand}}
* {{annotated link|In Praise of Limestone|'In Praise of Limestone'}}
* {{annotated link|Kurkar}}
* {{annotated link|Limepit}}
* {{annotated link|Sandstone}}
    
```

Insert **Cite your sources:** <ref></ref>

Edit summary (Briefly describe your changes)

This is a minor edit Watch this page Permanent

By publishing changes, you agree to the [Terms of Use](#), and you irrevocably agree to release your contribution under the [CC BY-SA 3.0 License](#) and the [GFDL](#). You agree that a hyperlink or URL is sufficient attribution under the Creative Commons license.

Publish changes **Show preview** **Show changes** **Cancel**

Figure 11. The Source View editing option for the Limestone article in Wikipedia, allowing direct editing of the Wikitext.

Once images are in Wikimedia Commons, embedding the images in Wikipedia articles is a simple process. One option for Wikipedia editors is to use a What You See Is What You Get (WYSIWYG)

HTML editor that should be familiar to most users. Alternately, there is also a Source View editing option which uses the custom markdown called [Wikitext](#) to format pages in Wikipedia (see figure 11).

Source View editing allows more precision when inserting images into Wikipedia articles and makes it easier to understand how they will ultimately be displayed in the article. The way in which different page elements flow around one another in articles can be surprising when using the WYSIWYG editing option as images assumed to show up where you placed the cursor can ultimately be placed in very different locations than expected.⁸

USAGE ANALYTICS

Analytics tools are available that allow organizations to track the articles containing their assets, showing what image was embedded in an article and how many views the article received. TLCPL’s initial ingest added a total of 129,725 discrete image assets to Wikimedia Commons. From that pool, images were added to a total of 227 Wikipedia articles between May 2020 and February 2021 (see figure 12). In that time period the articles had a total of 11.7 million page views (see figure 13).⁹ In February 2021 alone, the 227 enriched articles received 1.87 million page views. By comparison, the total number of records TLCPL had available in Ohio Memory was 129,395 in February 2021, and those records received 26,602 unique page views. The major strength of this project is to display locally created digitized resources where researchers would already be on the open web and take advantage of that much wider level of exposure.¹⁰

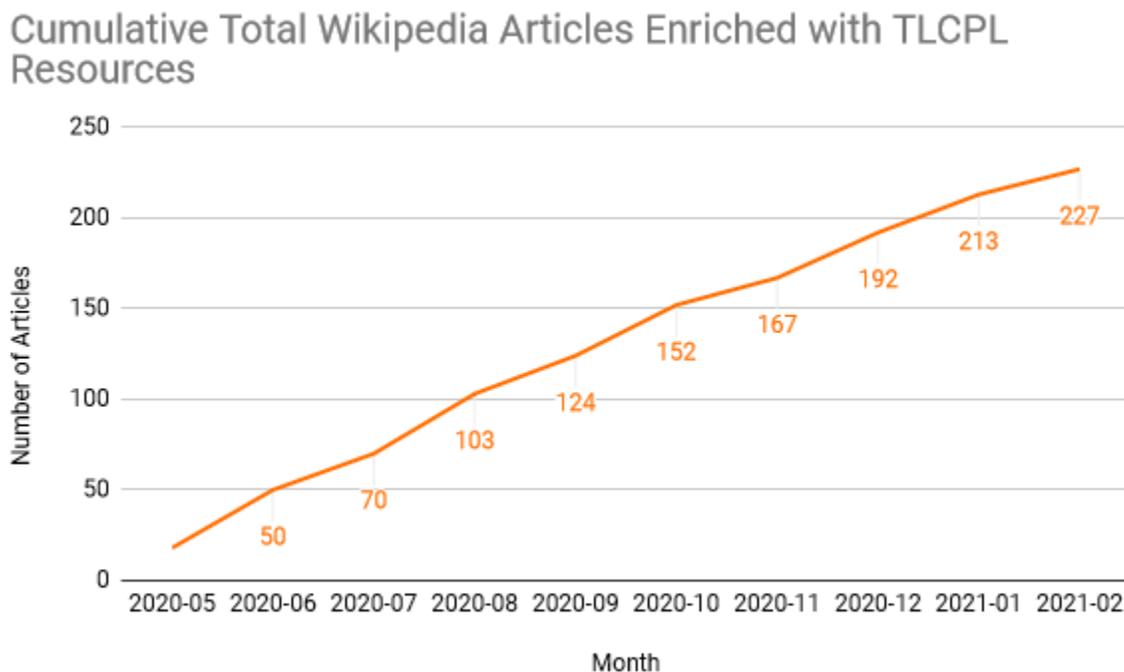


Figure 12. A graph displaying the cumulative total number of articles with inserted images from TLCPL resources from May 2020 to February 2021.

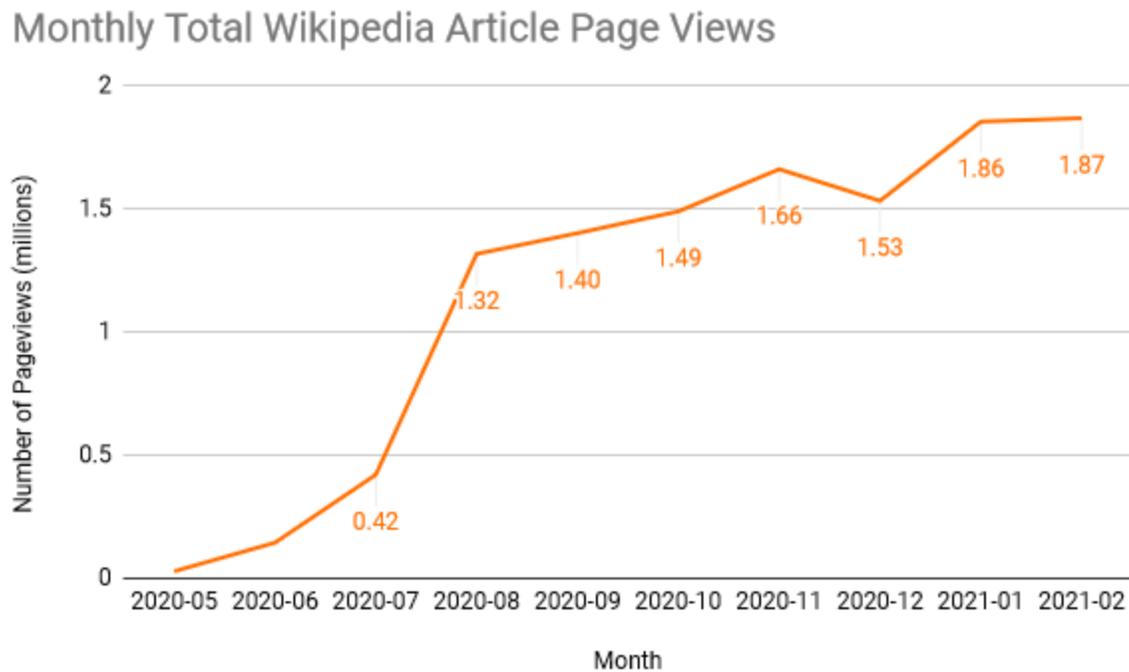


Figure 13. A graph displaying the monthly total number of page views of Wikipedia articles with inserted images from TLCPL resources from May 2020 to February 2021.

There is a valid discussion to be had about the comparison of these metrics, as comparing page views to unique page views is not a one-to-one match, but no matter the measurement it is fairly clear this audience is an order of magnitude beyond what might conventionally be available.

Ultimately what might be one of the most exciting metrics for an organization looking to implement this work is the amount of time it actually took to execute this project. Since TLCPL's records already satisfied the requirements to be copied over to Wikimedia Commons, the actual import process was able to begin during the April 30, 2020 Zoom call between TLCPL and DPLA staff that was set up to discuss the project; from the perspective of the contributing organization, this process takes essentially no time or effort. Once the process is started, staff at the contributor institution will be informed when the records have finished being copied.

The actual work of locating images for inclusion into articles and inserting them took roughly an hour of work a week, or roughly ten minutes per article, sometimes less. Approximately 38 hours of work was spent identifying images and inserting them into articles between May 2020 and February 2021.

While not of central concern to the project or its usage, the editorial work is also interesting and uses enough creativity and problem solving to be an enjoyable activity. Because the resources in Wikimedia Commons are available to be used by anyone (as in, anyone with a device and an internet connection), this makes it a wonderful opportunity for interns or volunteers to contribute. Volunteers could work on the editorial portion of this project remotely with no real barriers. While all the effort of getting TLCPL digitized images into Wikipedia described here has been using previously existing articles, this work could make an excellent partnership opportunity

with schools to write and create whole new articles for which there is an abundance of already digitized resources to support.

CONCLUSION

The work of remediating and writing metadata to participate in large consortial efforts such as DPLA is always going to be a major undertaking, but projects like this that can leverage automation and partnerships show just how powerful these relationships can be. Making locally digitized resources available through DPLA, copying them over to Wikimedia Commons, and then embedding those images into Wikipedia articles is an excellent opportunity to meet users where they already are—online. This work provides exceptionally high usage statistics and is fertile ground for outreach and programming opportunities to get partners, volunteers, and interns involved with making those digitized resources available in Wikipedia.

ACKNOWLEDGEMENTS

Special thanks to Jen Johnson, Library Consultant at the State Library of Ohio, and Virginia Dressler, Digital Projects Librarian at Kent State University, for their support in enabling this work and article.

ENDNOTES

- ¹ This presentation is available on YouTube at <https://youtu.be/OBS0KSYBcBI>. Information on all past DPLA webinars and programming can be found at <https://pro.dp.la/events/workshops>.
- ² The entire collection of all resources contributed to Wikimedia Commons via DPLA can be found at https://commons.wikimedia.org/wiki/Category:Media_contributed_by_the_Digital_Public_Library_of_America.
- ³ Statistics related to contributor totals were created from a Wikimedia database query published at <https://quarry.wmflabs.org/query/51256>.
- ⁴ *Geography of Ohio* was published as part of a series of bulletins by the Ohio State Geological Survey. The book was authored by Roderick Peattie, an assistant professor of geology at Ohio State University, in 1923. This item was digitized by the Toledo Lucas County Public Library and uploaded as part of Public Domain Day 2019. The digitized version of this book is available through Ohio Memory at <https://ohiomemory.org/digital/collection/p16007coll33/id/115214>.
- ⁵ More information, including a complete list of Content Hubs and Services Hubs and their geographic distribution, is available on the DPLA website at <https://pro.dp.la/hubs/our-hubs>.
- ⁶ As shared by Dominic Byrd-McDevitt in a webinar on March 18, 2021 entitled “DPLA + Wikimedia: One Year In + Ten Million Views,” available at <https://www.youtube.com/watch?v=jLoJ0gvvSNU>.
- ⁷ The JSON file is available for view on DPLA’s GitHub page at https://github.com/dpla/ingestion3/blob/develop/src/main/resources/wiki/institutions_v2.json.

- ⁸ For more information on step-by-step instructions for adding images into Wikipedia articles after harvest, see the blog post at <https://johndeweese.com/2021/03/18/adding-images-to-wikipedia-articles-via-dpla/>.
- ⁹ All statistics on Wikipedia page views are drawn from the BaGLAMa 2 utility available at <https://glamtools.toolforge.org/baglamma2/#gid=430>.
- ¹⁰ Up-to-date statistics and data are available at the Digitization Statistics Dashboard created to communicate about major projects in Digitization Services at the Toledo Lucas County Public Library and available at https://docs.google.com/spreadsheets/d/1jV0ZzT6H_Jl1tQ8v2zDXmf5YGn0dfBnhQbFFiFCEpZM/edit?usp=sharing.