

# About Bookshelf

Last Updated: 11, 2.02



National Center for Biotechnology Information (US), Bethesda (MD)

NLM Citation: About Bookshelf [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); 2010-.

# Table of Contents

<b>Bookshelf Overview</b> .....	1
About the Content .....	1
About the Archive .....	2
<b>COVID-19 Collection of Guidelines and Reviews</b> .....	5
<b>How to Include Content in Bookshelf</b> .....	7
Pre-Application Requirements .....	7
Evaluation and Setup Process .....	7
Scientific Quality Review .....	9
Technical Evaluation .....	9
Preview .....	10
Live Release .....	10
<b>Bookshelf Copyright Notice</b> .....	11
Copyright-protected Publications .....	11
Publications in the Public Domain .....	11
Disclaimer .....	11
Restrictions on Systematic Downloading of Books or Chapters .....	12
Publisher-specified Copyright Statements .....	12
<b>Bookshelf Disclaimer</b> .....	13
<b>Information for Librarians</b> .....	15
MARC Records .....	15
NLM Catalog .....	15
<b>Information for Authors and Publishers</b> .....	17
Selection of Titles for Bookshelf .....	17
The Application Process .....	18
Participation Agreement .....	19
NLM LitArch Open Access Subset .....	19
Books and PubMed Central International (PMCI) .....	19
Bookshelf Citations in PubMed .....	19
<b>Bookshelf and Research Funder Policies</b> .....	21
Partner Organizations .....	21
Funder-Sponsored Content .....	22

<b>Bookshelf Conversion Projects</b> .....	23
<b>NLM LitArch Open Access Subset</b> .....	25
<b>FTP Service</b> .....	27
Source Files from the NLM LitArch Open Access Subset .....	27
Locating OA Source Files .....	27
Display a Title in Bookshelf Using the Accession Number .....	28
Suggested FTP Client Configuration .....	28
<b>File Submission Specifications</b> .....	29
1 XML Coding .....	29
2 Images .....	29
3 PDFs .....	30
4 Supplementary Data .....	30
5 File Naming and Delivery .....	30
6 Additional XML Tagging Guidance .....	31
7 Tools .....	33
Submission sample .....	33
<b>OAI-PMH Service</b> .....	35
Copyright .....	35
Using the Books-OAI Service .....	35
High-Volume Retrievals .....	35
Access to Full Text .....	35
Supported Data Formats .....	35
OAI Books identifiers .....	36
Automatic Segmentation of Large Result Sets .....	36
Examples .....	36
<b>Content Selection for Bookshelf</b> .....	39
<b>Scientific, Editorial, and Technical standards: Reevaluation</b> .....	41
Scientific and Editorial Reevaluation .....	41
Technical Reevaluation .....	41

## Bookshelf Overview

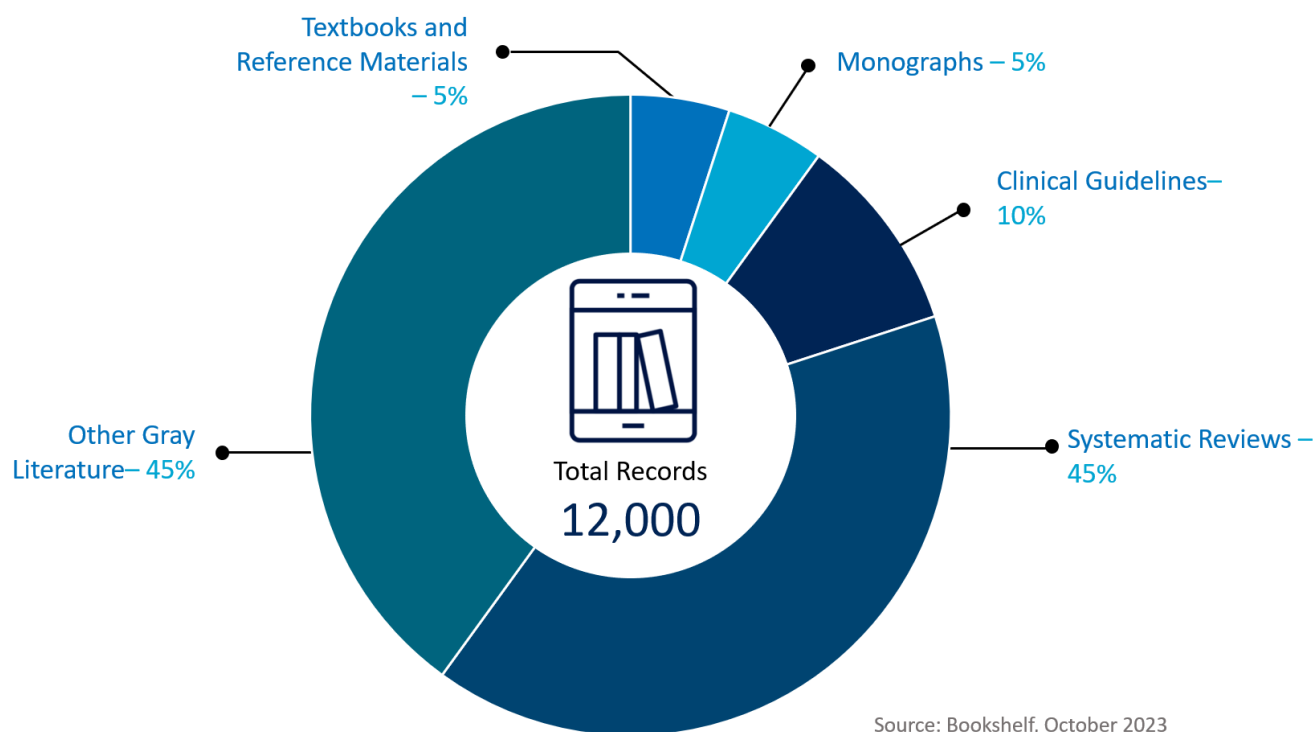
Updated: October 27, 2023.

Bookshelf provides free access to the full text of books and documents in the biomedical and life sciences as well as health care and medical humanities and social sciences at the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM). Through integration with other NCBI databases, such as PubMed, Gene, Genetic Testing Registry, and PubChem, Bookshelf also provides reference information for biological, chemical and other biomedical data and facilitates its discovery.

Bookshelf was developed and is managed by NLM's National Center for Biotechnology Information (NCBI).

## About the Content

Bookshelf started in 1999 with the third edition of Alberts' *Molecular Biology of the Cell*. Since then Bookshelf has grown to 9,000 titles beyond classic textbooks, to include monographs, reference books, systematic reviews, technical reports, clinical guidelines, web materials, grey literature, and other peer-reviewed documents.



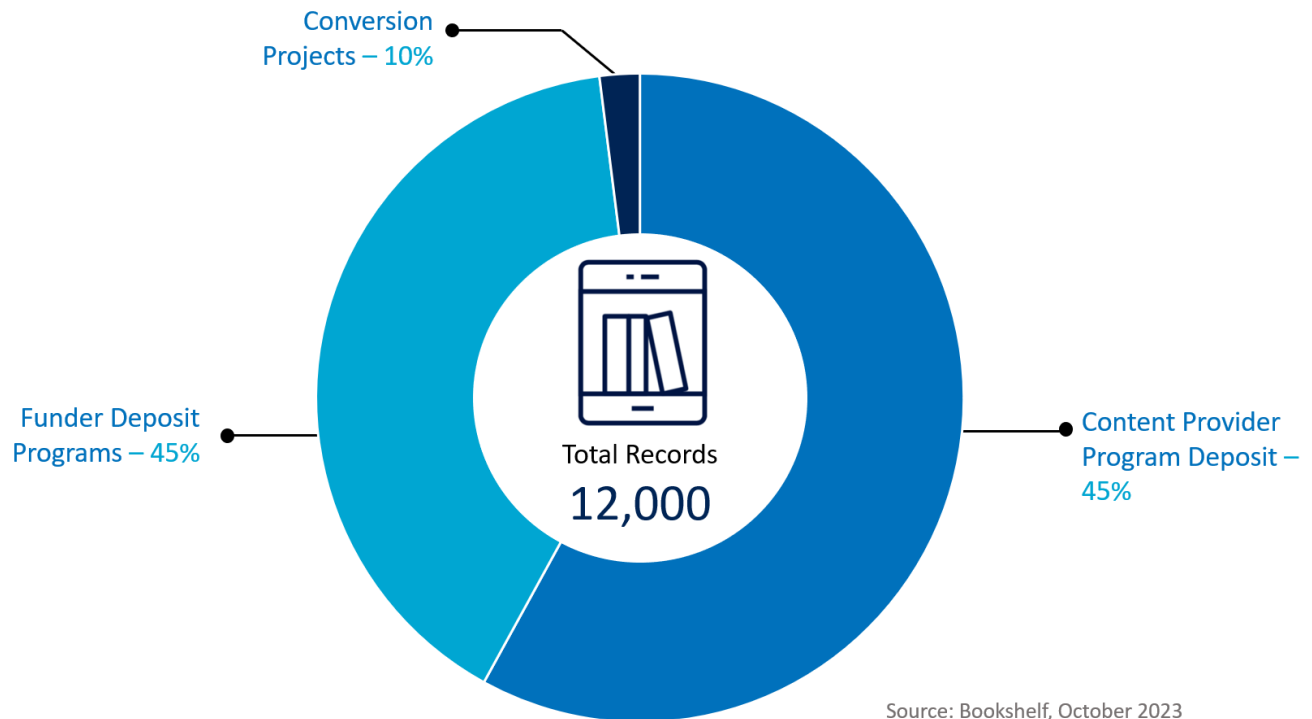
Content is added to the archive through

Content Provider and Publisher Program Deposit (45%)	Funder Program Deposit (45%)	Conversion Projects (10%)
NLM has agreements with publishers, scholarly societies, and other content owners to deposit books and documents directly to Bookshelf.	NLM partners with federal agencies, international organizations, and other funding or sponsoring organizations to accept the deposit of health reports, systematic reviews, clinical guidelines, monographs, and related reference works in compliance with funder policies and goals.	NLM collaborates with NIH or other institutionally-sponsored authors and editors to make available biological, chemical, and other highly regarded biomedical reference materials through conversion of Word documents.

Table continued from previous page.

Learn More	Learn More	Learn More
------------	------------	------------

The resulting collaborations with publishers, societies, funders, sponsors, authors and editors form the foundation of Bookshelf. Bookshelf is not a publisher and does not publish books and documents itself.



## About the Archive

NLM's fundamental responsibility is to permanently preserve periodicals, among other materials, pertinent to medicine. (See the [NLM Preservation Policy](#) for more information.) As such, Bookshelf is designed to provide permanent access to all its content, even as technology evolves. To do so,

	<p>Bookshelf immediately makes all content free to read as NLM believes that the best way to ensure the accessibility and viability of digital material over time is through consistent and active use of the archive. However, free access does not mean that there is no copyright protection (see the <a href="#">Bookshelf Copyright Notice</a>).</p>
	<p>Bookshelf stores content in eXtensible Markup Language (XML), which represents the structure and meaning of a document in a relatively simple and human-readable form. All Bookshelf content is converted to and stored in the <a href="#">Book Interchange Tag Suite (BITS)</a> XML format.</p>
	<p>Bookshelf leverages the existing <a href="#">PubMed Central® (PMC)</a> software architecture to store and make its content accessible to users.</p>
	<p>NLM operates the <a href="#">PMC International</a> network, which provides a framework to maintain copies of the Bookshelf corpus in other reliable international archives that share the goals of PMC and Bookshelf and operate on the same principles.</p>

Because Bookshelf has the capacity to store and cross reference data from diverse sources in a common format within a central repository, a Bookshelf user can quickly search either the entire collection of full-text books and documents or within a specific book or series to locate all relevant materials. The structure of the Bookshelf archive also supports the integration of the literature with other resources.

Bulk retrieval of files for text mining and other purposes is permitted within a subset of the Bookshelf archive.





## COVID-19 Collection of Guidelines and Reviews

Updated: May 3, 2021.

In collaboration with participating U.S. health agencies, international health organizations, and other sponsors and publishers of evidence-based clinical or public health information, the NCBI Bookshelf, a free online archive of full-text books and documents collected by the U.S. National Institutes of Health's National Library of Medicine (NIH/NLM), has developed a collection of COVID-19 guidelines and reviews. All COVID-19 Guidelines and Reviews on Bookshelf may be accessed by this regularly updated [search query](#).






# How to Include Content in Bookshelf

Updated: October 7, 2019.

This page describes the process for including content in Bookshelf as well as the pre-application requirements. Any content provider that submits an application that meets all the pre-application requirements undergoes an evaluation process. As part of the evaluation process, the National Library of Medicine (NLM) considers the scope of the content as well as its scientific and editorial quality. Content selected for inclusion in Bookshelf is also evaluated on technical quality. Content providers are encouraged to read through all steps and review Bookshelf's information for authors and publishers.

## Pre-Application Requirements

### For Content

	The content must be peer reviewed
	The content provider must be able to provide NLM with a PDF or Word document of the content or immediate access to it at a publisher or third-part site
	The full text of the content must be available in English

### For Content Providers

NLM looks for ongoing content provider conformance with published guidelines and best practices, including OAPEN Guides for OA monographs, IOM Committee on Standards for Developing Trustworthy Clinical Practice Guidelines, OER Authoring Resources, PRISMA guidelines, CARE case report guidelines, CONSORT, and ARRIVE.

## Evaluation and Setup Process

Submit Application → Screening → Scientific Quality Review → Technical Evaluation → Preview → Live Release

### Submit Application

Information you'll need to submit when applying








	Title of proposed content
	URL, PDF, or Word document of full text of proposed content
	Publisher name and information about its management, qualifications, and publishing policies





Table continued from previous page.

	Abstract or description of content
	Copyright holder
	Information about the authors and editors, including links to bios or CVs
	For more information and a link to the Bookshelf application, see our information for authors and publishers or contact <a href="mailto:bookshelf@ncbi.nlm.nih.gov">bookshelf@ncbi.nlm.nih.gov</a>

## Initial Application Screening

Once received, NLM will check that the application is complete and meets the pre-application requirements. NLM will communicate any inconsistencies about the application back to the applicant. If the application is complete, NLM will determine if the content is appropriate for Bookshelf. Content providers who do not pass the initial application screening are eligible to reapply in 24 months.

Factors that determine if content is in scope for Bookshelf

	Peer-reviewed with a clearly stated peer-review policy
	Generally, falls within the biomedical and life sciences
	Web content and clinical resources are regularly reviewed and updated
	The affiliations of the editors and authors should reflect the scope of the content and demonstrate editorial independence and diversity

The following content types are in scope:

- [Monographs](#);
- [Biographical Works](#);
- [Collected Works](#);
- [Government Publications](#);
- [Handbooks and Manuals](#);
- Reference works, e.g., [Dictionaries and Encyclopedias](#);
- [Standards and Guidelines](#);
- [Statistical Works](#);

- [Technical Reports](#);
- [Textbooks](#);
- [Web Content](#);
- and other types of [Grey Literature](#).

In very exceptional cases, Bookshelf will consider [self-published or independently sponsored works](#).



For more information,  
see our [Scope Details](#)

## Scientific Quality Review

After a thorough review of the content provider and its content, Bookshelf will inform the content provider whether the content meets Bookshelf's scientific quality standard.

Content that does not meet this standard is eligible for reapplication in 24 months.



For more information,  
see [Content Selection for Bookshelf](#)

## Technical Evaluation

First, Bookshelf asks the content provider to fill out a Technical Evaluation Checklist, which must be returned to Bookshelf before starting any work on the evaluation.

### For Collections and Updated Web Content

Next, if the content is a collection of documents or regularly updated content, the content provider submits a set of at least 3-5 sample documents that is representative of the collection's content types and styles. The set of samples must meet Bookshelf's File Submission Specifications. In addition, Bookshelf will request a cover image for the collection or website, as well as a brief description of the collection to appear on its landing page in Bookshelf.

### For Books

If the content is a single book or other type of document, the content provider submits it to Bookshelf according to Bookshelf's File Submission Specifications.



## What does Bookshelf evaluate?

Bookshelf evaluates the content to ensure that:




The XML Coding meets Bookshelf's specifications, including explicit markup of all structural and textual elements of a source and enough metadata to build an accurate citation

Table continued from previous page.

	The deposit adheres to Bookshelf's File Submission Specifications
	The presentation of the content, as generated from the XML and associated files, is an accurate, error-free representation of the version of record




If necessary, the content provider must correct reported errors and submit a revised package for a subsequent round of evaluation.

*If a deposit fails to meet Bookshelf's file submission specifications at any stage of the application process, the application may be rejected.*

	For more information, see File Submission Specifications
-----------------------------------------------------------------------------------	----------------------------------------------------------

## Preview

Once the content passes Bookshelf's technical evaluation phase and is set up in Bookshelf's systems, Bookshelf will request the following from the content provider:

	Approval of a preview of the content as generated from the XML and associated files
	A formal NLM LitArch Participant Agreement
	A complete set of document files belonging to an accepted collection or website according to the terms of the participant agreement

## Live Release

NLM countersigns the content provider's LitArch Participation Agreement and releases the content to the Bookshelf public site with the content provider's approval.

Ongoing participation in Bookshelf requires a content provider to continue to meet NLM's quality standards for Bookshelf and the NLM collection. The process for reevaluating accepted content is described in Bookshelf's Reevaluation section.

## Bookshelf Copyright Notice

Created: March 28, 2011; Updated: April 18, 2016.

This page provides information about copyright restrictions that apply to content in Bookshelf.

### Copyright-protected Publications

For content in Bookshelf, read the license and copyright statements on the bottom of each page. The license and copyright statements define what uses of the document are permitted, and apply to all associated files, including images and supplementary material, unless otherwise indicated. The terms and conditions of use are not identical for all documents. In the absence of a license or copyright statement, the document is in the public domain. Contact the publisher if you have any questions about the permissible uses of the documents.

### Publications in the Public Domain

Some of the content found in Bookshelf is authored and published by the National Center for Biotechnology Information (NCBI) or other institution of the U.S. government. No permission is needed to reproduce or distribute this type of content, but the authoring institute or agency must be given appropriate attribution.

Occasionally, copyrighted material (such as a figure, table, or supplementary material), may appear in a U.S. government publication in Bookshelf. In these cases, a special copyright statement such as “Reprinted with permission from...” will accompany the material. Permission must still be sought from the original author/publisher of such material before it can be reproduced or distributed.

### Disclaimer

This disclaimer relates to PubMed, PubMed Central (PMC), and Bookshelf. These three resources are scientific literature databases offered to the public by the U.S. National Library of Medicine (NLM). NLM is not a publisher, but rather collects, indexes, and archives scientific literature published by other organizations. The presence of any article, book, or document in these databases does not imply an endorsement of, or concurrence with, the contents by NLM, the National Institutes of Health (NIH), or the U.S. Federal Government.

Content in NLM literature databases may be published by academic publishers or institutions, scholarly societies, or government and non-governmental organizations. To be added to a database, a publication must apply and be selected by NLM for inclusion in MEDLINE, PMC, or Bookshelf. PubMed indexes and makes searchable the contents of these databases; MEDLINE is the primary component of PubMed. Details on the content selection processes for each database can be found at:

- [MEDLINE](#)
- [PubMed Central](#)
- [Bookshelf](#)

Once publications are selected for inclusion in a database, NLM does not review, evaluate, or judge the quality of individual articles and relies on the scientific publishing process to identify and address problems through published comments, corrections, and retractions (or, as in the case of preprints, withdrawal notices). The publisher is responsible for maintaining the currency of the scientific record and depositing all relevant updates to the appropriate NLM database.

For more information about our content and how it relates to you, please see: <https://pubmed.ncbi.nlm.nih.gov/disclaimer/>

## Restrictions on Systematic Downloading of Books or Chapters

Crawlers and other automated processes may NOT be used to systematically retrieve content from the Bookshelf web site.

Bookshelf has an auxiliary service that may be used for automated retrieval and downloading of a special subset of books, namely, books in the Open Access Subset. This auxiliary service, the NLM LitArch FTP service, is the *only* service that may be used for automated downloading of content in Bookshelf.

## Publisher-specified Copyright Statements

Use these links to the following publishers' sites for additional copyright statements that apply to their respective content in Bookshelf.

- [BC Decker](#)
- [CADTH](#) (Canadian Agency for Drugs and Technologies in Health)
- [Garland Science](#) (an imprint of Taylor & Francis Group)
- [Lippincott-Raven](#) (an imprint of [Lippincott Williams & Wilkins](#))
- [Remedica](#)
- [Sinauer Associates](#)
- [University of Washington](#)
- [W. H. Freeman](#)



## Bookshelf Disclaimer

Created: March 28, 2011; Updated: September 11, 2023.

This disclaimer relates to PubMed, PubMed Central (PMC), and Bookshelf. These three resources are scientific literature databases offered to the public by the U.S. National Library of Medicine (NLM). NLM is not a publisher, but rather collects, indexes, and archives scientific literature published by other organizations. The presence of any article, book, or document in these databases does not imply an endorsement of, or concurrence with, the contents by NLM, the National Institutes of Health (NIH), or the U.S. Federal Government.

Content in NLM literature databases may be published by academic publishers or institutions, scholarly societies, or government and non-governmental organizations. To be added to a database, a publication must apply and be selected by NLM for inclusion in MEDLINE, PMC, or Bookshelf. PubMed indexes and makes searchable the contents of these databases; MEDLINE is the primary component of PubMed. Details on the content selection processes for each database can be found at:

- [MEDLINE](#)
- [PubMed Central](#)
- [Bookshelf](#)

Once publications are selected for inclusion in a database, NLM does not review, evaluate, or judge the quality of individual articles and relies on the scientific publishing process to identify and address problems through published comments, corrections, and retractions (or, as in the case of preprints, withdrawal notices). The publisher is responsible for maintaining the currency of the scientific record and depositing all relevant updates to the appropriate NLM database.

For more information about our content and how it relates to you, please see: <https://pubmed.ncbi.nlm.nih.gov/disclaimer/>



## Information for Librarians

### MARC Records

#### Recent Releases

In April, 2011, the National Library of Medicine (NLM) Cataloging Section began maintaining MARC 21 records for Bookshelf titles. Batches of records are released as they become available or are updated, and may be downloaded by release date from Bookshelf's public ftp site at: <ftp://ftp.ncbi.nlm.nih.gov/pub/bookshelf/>.

### NLM Catalog

It is also possible to view all available catalog records for books and resources in Bookshelf by using the NLM Catalog. Within the Search field on the NLM Catalog home page at <http://www.ncbi.nlm.nih.gov/nlmcatalog>, enter the following search query:

```
nlmcatalog_books[sb]
```



## Information for Authors and Publishers

Updated: September 29, 2022.

Bookshelf provides free access to the full text of books and documents in the biomedical and life sciences as well as health care at the National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM). The collection is referred to as the [NLM Literature Archive](#). Content in Bookshelf is made available by participating authors, editors, funders, sponsors, and publishers. Through integration with other NCBI databases, such as PubMed, Gene, Genetic Testing Registry, and PubChem, Bookshelf provides reference information for biological data and facilitates its discovery.

This page describes how content is selected for Bookshelf and how to apply.

### Selection of Titles for Bookshelf

Participation in Bookshelf is open to any English-language work that is in scope for and meets NLM's standards for the archive. A work must qualify on two levels, both the Scientific quality of the title, and the Technical quality of its digital files.

Proposed full-text book and document titles are reviewed by the National Library of Medicine for inclusion in the NLM collection. To be considered for Bookshelf, a work must fall within the scope of core biomedical subjects as defined in the [Collection Development Guidelines of the National Library of Medicine](#).

Consideration for participation in Bookshelf is given to varied types of peer-reviewed, full-text content. The following format and literature types are within scope for Bookshelf:

- [Monographs](#);
- [Collected Works](#);
- [Government Publications](#);
- [Handbooks and Manuals](#);
- Reference works, e.g., [Dictionaries and Encyclopedias](#);
- [Standards and Guidelines](#);
- [Statistical Works](#);
- [Technical Reports](#);
- [Textbooks](#);
- [Web Content](#);
- and selected other types of [Grey Literature](#).

Works consisting entirely of original research articles, theses, and dissertations, are not acceptable. It is important that the Bookshelf collection reflects the kinds of content available in the wider book and reference literature, as far as is possible. Proposed content may be written for scientists, health professionals, educators and students, or consumers.

Supplementary materials and data sets are only accepted in association with the complete full-text.

Except in rare cases when proposed content meets NLM's [self-published or independently sponsored works](#) criteria, proposed content must be accepted for publication by a credible publisher or content provider. All new content providers and sponsors are reviewed by the National Library of Medicine to ensure they meet scientific and scholarly communications standards.

In cases where a content provider is regularly producing related materials of the same type, NLM may choose to evaluate this content as a collection, rather than independently evaluating each item as a monograph or other type of content. This arrangement will allow future materials in the collection to be included in Bookshelf

without further applications or scientific quality review. Bookshelf decides to evaluate content as a collection when:

- Enough content in the collection has been published to evaluate the consistency of quality.
- The peer review process and other editorial policies and processes are the same across the entire collection.
- The material is published consistently and at regular intervals.
- All materials in the collection have the same intended purpose and audience.
- The structure, formatting, and layout remains consistent across editions/issues, with only minor changes.

In such cases, NLM expects publishers to have at least a two-year history of quality scholarly publishing in the life sciences. NLM may consider an application from a publisher that has been publishing scholarly content for less than two years if there is evidence that the management and individuals responsible for editorial quality and operations have adequate experience in comparable positions at other organizations.

## The Application Process

Authors, editors, funders, sponsors or publishers may put forward titles for review. Authors and editors who would like their books to be included on the Bookshelf should first agree on the application with the book publisher or with the copyright holder.

To begin the application process, download and complete the [application](#) for single titles or the [application spreadsheet](#) for multiple titles. Send an email to [bookshelf@ncbi.nlm.nih.gov](mailto:bookshelf@ncbi.nlm.nih.gov) and attach the completed application. If you have not archived content in Bookshelf previously, you will also need to complete the [publisher information](#) sheet.

In addition to the application, Bookshelf requests the applicant to provide the following:

1. A PDF version of the content or URL where the content may be accessed.
2. An abstract or blurb about the content
3. Brief biographies or CVs of authors or editors.  
Note: In the case of content in which chapters are contributed by different authors, the biography of the editors is sufficient

If a proposed title is deemed to be within scope for the NLM Collection, then its scientific merit is the primary consideration in selection for publication and indexing in Bookshelf. A scientific evaluator will review the content.

When the content evaluation is completed, Bookshelf will inform the applicant whether or not the book meets the scientific quality standard. For books which qualify on this basis, the process will continue as follows:

- Bookshelf will request the applicant to submit full-text XML files (including abstract), images (including cover image), and all supplementary materials such as video, audio, or data files, and a PDF version of the book if available.
- Bookshelf will evaluate the digital files to assess the technical quality, including that all required files have been supplied and are correctly named, the XML validates according to a mutually-agreed DTD, and the online organization of the content, as generated from the XML and associated files, is an accurate and complete rendering of the content.
- If errors or anomalies are found in the data, Bookshelf will request the applicant to correct the data and resubmit. If the data fails to meet the technical quality standards, the book can be rejected.
- Bookshelf reevaluates corrected data. If it is found to be satisfactory, the book can be included in Bookshelf.

If a book is rejected, there is no appeal process; applicants may resubmit the book after a 24 month period.

Ongoing participation in Bookshelf requires a content provider to continue to meet NLM's scientific and technical quality standards for Bookshelf and the NLM collection over time. Bookshelf may choose to reevaluate content if these standards are not met.

## Participation Agreement

A [participation agreement](#) must be signed by NLM and the copyright holder of the book before Bookshelf begins the production process and releases the book on the Bookshelf website. Once the full text of the book has been made freely available on the Bookshelf website, per the terms of the agreement, Bookshelf does not support removal of the content nor any post-publication embargoes. Bookshelf provides free access to the full text of all content in its archive regardless of participation in the NLM LitArch Open Access Subset.

## NLM LitArch Open Access Subset

The content that you make available in NLM LitArch may also be made available in the Open Access subset, if you so choose. As part of this subset, NLM may make all of the content freely available to users for automated downloading (XML, images, PDF) via the NLM LitArch FTP site. Content in the subset is either in the public domain or made available under a [Creative Commons](#) or similar license that generally allows for less restricted redistribution and reuse than traditional copyrighted work. For more information, see the NLM LitArch Open Access Subset.

## Books and PubMed Central International (PMCI)

Bookshelf shares content with Europe PMC. For more information, see [PMC International](#).

## Bookshelf Citations in PubMed

Bookshelf will submit citation(s) to [PubMed](#) for a subset of books and resources included in its collection, as determined by NLM. The display of book citations in PubMed is based on [Citing Medicine, 2<sup>nd</sup> edition](#). More information on the display and search of book citations and abstracts in PubMed is available in the [NLM Technical Bulletin](#).





## Bookshelf and Research Funder Policies

Updated: October 7, 2019.

Bookshelf was established in 1999 as the National Library of Medicine's (NLM) full-text repository of life sciences books and documents. Since 2004, with the addition of [HSTAT](#), Bookshelf has been the home for full text health technology assessments, systematic reviews, clinical guidelines, and related health documents sponsored by US federal agencies and other private and international organizations partnering with NLM. Since 2009, Bookshelf has also been the home for [National Academies reports funded by the National Institutes of Health](#). Since 2014, Bookshelf has been the designated repository for monographs and chapters submitted in accordance with the [Wellcome Trust's open access policy on monographs and book chapters](#).

These partnerships support NLM's mission of enabling biomedical research; supporting evidence-based health care and public health; and promoting healthy behavior through the acquisition, organization, and preservation of the world's scholarly biomedical literature. Additionally, these efforts support and catalyze NLM's commitment to open science policies and practices. In bringing together the research outputs of many funding organizations and scientific disciplines in a central archive, NLM seeks to improve discovery of, provide consistent access to, and ensure long-term preservation of the scientific record.

NLM will consider partnering with funding and sponsoring organizations to leverage the Bookshelf repository infrastructure when the following conditions are met:

The funder or sponsor has an established process, e.g., peer review, for selecting its projects.

The funder has a public or open access policy that requires the deposit of funded research monographs and book chapters in Bookshelf.

The funder is able to demonstrate how partnering with Bookshelf will further the NLM and NIH missions as well as their own.

There is evidence that the funder's policy and implementation will contribute a sufficient volume of content to Bookshelf on a continuing basis.

To become a partner, a funder must sign a partnership agreement with NLM outlining the responsibilities of the NLM and the partner organization or make arrangements with an existing NLM partner (i.e., [Europe PMC](#)). In general, partners may leverage the Bookshelf infrastructure but remain responsible for the maintenance and enforcement of their individual policies.

### Partner Organizations

The following US government agencies, international organizations, and private funders have active partnerships with NLM to leverage the Bookshelf infrastructure.

#### US Federal Agency Partners

US Department of Health and Human Services

National Institutes of Health (NIH)

Agency for Healthcare Research and Quality (AHRQ)

Centers for Disease Control and Prevention (CDC)

Office of the Surgeon General (OSG)

Substance Abuse and Mental Health Services Administration (SAMHSA)

US Department of Veterans Affairs (VA)

## **Private and International Partners**

Canadian Agency for Drugs and Technologies in Health (CADTH)

Institute for Quality and Efficiency in Healthcare (IQWiG)

Research Triangle Institute (RTI)

UK National Health Service (NHS)

National Institute for Health and Care Excellence (NICE)

National Institute for Health Research (NIHR)

Wellcome Trust

World Health Organization (WHO)

European Observatory on Health Systems and Policies

Guidelines Review Committee

Health Evidence Network (HEN)

International Agency for Research on Cancer (IARC)

If you represent a funding organization that should be included on the list above, please let us know.

General inquiries about Bookshelf and funder policies can be sent to [bookshelf@ncbi.nlm.nih.gov](mailto:bookshelf@ncbi.nlm.nih.gov)

## **Funder-Sponsored Content**

If a partner funding organization sponsors the publication of content, NLM will recognize the funding organization as the content owner and expects the funder to take responsibility of the scientific and editorial quality of the content and ensure that it meets NLM's standards for Bookshelf.

## Bookshelf Conversion Projects

Updated: October 7, 2019.

Bookshelf was established in 1999 as the National Library of Medicine's (NLM) full-text repository of life sciences books and documents. One of the ongoing goals of Bookshelf is to facilitate the discovery of reference information for biological, chemical, and other biomedical data, both through retrieval in Bookshelf, but also through its integration with other NCBI databases, such as PubMed, Gene, Genetic Testing Registry, and PubChem. To support this goal, Bookshelf collaborates with NIH or other institutionally-sponsored authors and editors to make this content available through the conversion of Word documents meeting Bookshelf style guidelines to XML meeting Bookshelf's technical requirements.

NLM will consider such collaborations when the following conditions are met for the content to be converted:

The content meets NLM's guidelines for [independently sponsored works](#)

The content is sponsored by NIH or a credible institutional sponsor, such as a university department

The content is peer reviewed and adheres to a review and update cycle if continuously published

The content owner is able to demonstrate how collaborating with Bookshelf will further the NLM and NIH missions as well as their own, such as providing reference material to data archived in NCBI and other NLM databases

NLM will recognize the institutional sponsor as the content owner and expects the content owner to take responsibility of the scientific and editorial quality of the content and ensure that it meets NLM's standards for Bookshelf. The content owner must sign a participant or other partnership agreement with NLM outlining the responsibilities of NLM and the collaborating sponsor.

Bookshelf collaborates with participating sponsors, authors, and editors on the following active conversion projects:

[Assay Guidance Manual](#)

[EndoText](#)

[GeneReviews](#)

[LactMed](#)

Among the conversion projects that are no longer active include:

[Baculovirus Molecular Biology](#)

[Enterococci](#)

[Molecular Imaging and Contrast Agent Database \(MICAD\)](#)

[Streptococcus pyogenes](#)

In addition, Bookshelf conversion projects include support for documentation about NCBI and NLM resources.

General inquiries about Bookshelf and funder policies can be sent to [bookshelf@ncbi.nlm.nih.gov](mailto:bookshelf@ncbi.nlm.nih.gov)



## NLM LitArch Open Access Subset

A subset of the total collection of books and documents in the NLM Literature Archive ([NLM LitArch](#)), accessible through the Bookshelf website, are available through the NLM LitArch Open Access subset. Contents in the NLM LitArch Open Access subset generally include works which are in the public domain, works which are available under a [Creative Commons](#) or similar license, and works whose authors or publishers have explicitly agreed to the terms of the NLM LitArch Open Access subset. Except for public domain works, works in the NLM LitArch Open Access subset are still protected by copyright, but are made available under a Creative Commons or similar license that generally allows more liberal redistribution and reuse than a traditional copyrighted work. The license terms are not the same for each work. Read the license text which is available with each downloadable file to determine terms of use.

Please note the following regarding the NLM LitArch Open Access subset:

- Most books and documents available in NLM LitArch and accessible via the NLM Bookshelf site are subject to traditional copyright restrictions, and are not part of this subset. Read the Bookshelf Copyright Statement for more information.
- The NLM LitArch FTP service is the *only* service that may be used for automated downloading of content from this open access subset.
- Systematic retrieval (bulk downloading) of content through any other automated process is prohibited, even if you are only retrieving documents from this subset.
- Some publishers and content providers use the term “open access” for books and reports that are available for free at the time of publication, but they are still subject to the same copyright restrictions. Such documents are not part of the Bookshelf Open Access subset.

Use the FTP service to download the complete sets of files for books and documents in this subset. Each downloadable file typically consists of the following data file types: XML, images, and PDF.



## FTP Service

The FTP service for the [NLM Literature Archive](#) (NLM LitArch) may be used to download the source files for any book in the NLM LitArch Open Access (OA) Subset and can be used as a source for data mining. The NLM LitArch FTP site is normally updated once a day.

If you have questions or comments about the NLM LitArch FTP site, please write to the Support Center, using the “Support Center” link at the bottom right corner of any Bookshelf web page.

## Source Files from the NLM LitArch Open Access Subset

This FTP service may be used to download the source files for any book in the NLM LitArch Open Access subset. The source files for a book may include:

- Files with the .xml suffix, which are XML files encoded in the NLM/BITS DTD
- PDF file(s) if available
- Image files or graphics for display versions of mathematical equations

The URL to access the FTP site is <ftp://ftp.ncbi.nlm.nih.gov/pub/litarch/>

All the source files for a book are packaged in a single .tar.gz file. The FTP site has a 2-level-deep folder (directory) structure. Folder names are randomly generated and the .tar.gz file for a book is randomly assigned to a second-level folder.

## Locating OA Source Files

To locate a specific title or set of titles in the OA subset, refer to this file list which is available in 2 formats, text (.txt) and comma-separated values (.csv).

Text: [ftp://ftp.ncbi.nlm.nih.gov/pub/litarch/file\\_list.txt](ftp://ftp.ncbi.nlm.nih.gov/pub/litarch/file_list.txt)

CSV: [ftp://ftp.ncbi.nlm.nih.gov/pub/litarch/file\\_list.csv](ftp://ftp.ncbi.nlm.nih.gov/pub/litarch/file_list.csv)

Each list entry includes the following information:

- The fully qualified name of the .tar.gz file for a book, with directory structure
- Citation details:
  - Book title
  - Publisher
  - Publication date
- The accession number for the title which is a unique, persistent ID for the title that can be used to display it in Bookshelf
- Last updated (YY-MM-DD HH:MM:SS)

Example:

```
cc/8d/healthus08_NBK19617.tar.gz    Health, United States, 2008: With Special Feature on
the Health of Young Adults    National Center for Health Statistics (US)    2009
NBK19617    2012-03-12 15:42:52
```

where

```
File:    cc/8d/healthus08_NBK19617.tar.gz
Title:   Health, United States, 2008: With Special Feature on the Health of Young Adults
Publisher:    National Center for Health Statistics (US)
Publication date:    2009
Bookshelf Accession ID:    NBK19617
Last updated:    2012-03-12 15:42:52
```

## Display a Title in Bookshelf Using the Accession Number

Insert the accession number (accid) to the Bookshelf URL as follows:

`http://www.ncbi.nlm.nih.gov/books/<accid>/`

For example, to display

Health, United States, 2008: With Special Feature on the Health of Young Adults

National Center for Health Statistics (US) 2009

Accession number: NBK19617 (from the example above)

Use: <https://www.ncbi.nlm.nih.gov/books/NBK19617/>

## Suggested FTP Client Configuration

After a series of experiments using ftp clients with NCBI's FTP server, we've found that the configuration of ftp clients can seriously affect performance. NCBI recommends setting the TCP buffer size to 32Mb. For more information, please see <ftp://ftp.ncbi.nlm.nih.gov/README.ftp>.



# File Submission Specifications

Updated: November 19, 2024.

## 1. XML Coding

NLM LitArch requires full text content in XML format. The XML must be valid against a mutually agreed upon Document Type Definition (DTD). Submissions are preferred in the [Book Interchange Tag Suite \(BITS\) DTD](#). For detailed information on using the BITS DTD for submissions to NLM LitArch, please read the [Bookshelf Tagging Guidelines](#).

Alternative DTDs must provide explicit and detailed mark-up of all structural and textual elements of a source, including, but not limited to:

- Bibliographic metadata, such as titles, contributor names, or publication dates
- Organizational units, such as chapters, appendices, sections, or subsections
- Typical textual elements, such as paragraphs, figures, tables, footnotes, or reference citations
- Links, for example from the text body to figures, tables, or cited references

Regardless of the XML DTD used, the following metadata information must be present and tagged with correct values in every XML file:

- Author/Editor (may only be omitted if same as publisher)
- Book Title
- Edition (may only be omitted if it is a first edition)
- Publisher Name
- Place of Publication
- Date of Publication
- Language (if other than English)

NOTE: If you are using the Book Interchange Tag Suite (BITS), LitArch strongly suggests that you tag your documents so that they conform to the PMC Style Checker.

See also additional XML tagging guidance.

## 2. Images

### 2.1. Content Images

Generally, authors submit raw image data files to a publishing house in various formats (ppt, pdf, tif, jpg, xml, etc.). The files are then normalized to produce print or electronic output. NLM LitArch requires the normalized output, which is high-resolution, and of sufficient width and quality to be considered archival. Images generated at low resolution for display purposes are not acceptable.

- Uncompressed high-resolution TIFF or EPS files are required for all images, or, where these are unavailable, PNG, JPEG, or GIF format, at a resolution equivalent to the original source images published in the document
- All files must be cross-platform compatible
- Do not include thumbnail versions of any full-size images that are deposited in NLM LitArch
- Graphics must be legible throughout all submissions
- Graphics must be submitted in the expected orientation for display

NLM LitArch image specifications correspond to the PMC specification for journal submissions. More specific details on figure, equation, and table image quality can be found on the [PMC Image Quality Specifications page](#).

## 2.2. Cover Image

NLM LitArch displays a cover thumbnail for each book. A high-resolution source image of the cover or an alternative image representing the book should be submitted at time of delivery.

The aspect ratio for the image should be 4:5.

## 3. PDFs

One or more separate **non-password protected** PDFs which directly correspond to either (a) the complete book or (b) to each of the individual chapters in the book should be provided.

If print-quality PDFs are available, please submit them. If the book is not printed, the resolution of the images in the PDF should be no less than: Line Art 800 dpi, Halftones 300 dpi, Color 600 dpi. All fonts used in the file need to be fully embedded. Compression for images should be lossless (zip) or highest-quality JPEG. Illustrations should be encoded as vector data with no erroneous conversion to bitmaps.

## 4. Supplementary Data

NLM LitArch requires all available supplementary material to be submitted in a portable format, such as PDF, .doc, .csv, etc. Supplementary material should not be externally linked from the document text to an online location as a substitute for submission. Supplementary material has been defined to include all of the following:

- Voluminous material that was used to support the conclusions of the narrative, such as a genomic database or any data set which can never accompany a document based on sheer mass
- Material added to the work for enhancement purposes, such as a quiz, a PowerPoint presentation, or videos
- Textual or semi-textual content for which elements cannot be adequately described by the XML schema employed, for example fillable forms, questionnaires, or flowcharts

### 4.1. Video

NLM LitArch expects good-quality video and will downsample for web streaming if necessary. If the meaning of the video is not clear due to low quality, it must be improved prior to submission.

- Preferred Settings:
  - Audio codec: AAC
  - Sample audio bit rate: 128 kbit/s
  - Video codec: H.264
  - Video resolution: 480 vertical lines or better
  - Format: MPEG-4 (mp4) container
- Accepted file formats: mov, avi, mpg, mpeg, mp4, mkv, flv, wmv

Video files larger than 1GB should be split to several episodes, each less than 1GB.

## 5. File Naming and Delivery

NLM LitArch requires that data is named and packaged in a compressed archive, such as a zip file, and transferred to the NLM LitArch FTP site. Please write to [bookshelf@ncbi.nlm.nih.gov](mailto:bookshelf@ncbi.nlm.nih.gov) for an FTP account.

Data files:

- Names of image files and supplemental data files must match the names referenced in the XML file

- The name of the cover image should correspond to a book identifier used in the XML, for example an ISBN or a publisher ID
- Files should be grouped by type in appropriately named subfolders:
  - XML files should be delivered in subfolder named xml
  - Image files should be delivered in subfolder named images
  - PDF files should be delivered in subfolder named pdf
  - Supplementary data files should be in subfolder named suppl
- In case of multiple PDFs and XML files, the PDF and corresponding XML base file names must match
- All file names must be unique within a package
- All files names must contain a valid file type extension
- File names should not contain spaces, and should only contain letters, numbers, dashes, periods, and underscores
- XML file names should not exceed 20 characters in length

ZIP file package:

- NLM LitArch accepts .zip, .tar, .gz, and .tgz
- bzip2 compressed .zip files are not accepted
- The name of the package should, if possible, correspond to a book identifier used in the XML, for example an ISBN or a publisher ID
- Send an email notification to [bookshelf@ncbi.nlm.nih.gov](mailto:bookshelf@ncbi.nlm.nih.gov) for each submission including the compressed file name and the source title
- Do not send files as attachments to an email message
- For corrections please explicitly include what has been changed in the email notification. The name of replacement files (.xml, .pdf, .tif, etc.) must be identical to the original (replaced) file names

## 6. Additional XML Tagging Guidance

In addition to the minimum requirements above, below is guidance on XML tagging issues that arise commonly in content in LitArch.

### 6.1. Copyright Statement

If the content submitted to LitArch is under copyright, then a copyright notice, or statement must be included in the XML. See the [copyright tagging guidelines](#) for detailed tagging information and examples.

### 6.2. Open Access and License Statements

#### Terms.

A **license** is a statement which specifies the copyright permission granted for a document.

An **open access** (OA) document, in the context of LitArch, is a document that is published with a [Creative Commons](#) license or a similar license that allows any user to redistribute the document without requesting permission from the copyright holder. Additionally, this license may allow or prohibit the creation and distribution of derivative works or modified versions of the document, and/or may limit reuse and redistribution to non-commercial purposes only.

An **open access** document, in the context of LitArch, is NOT simply a document that is freely available at the time of publication.

The following conditions apply to all OA documents in LitArch:

- The applicable license will be clearly indicated in both the XML and PDF versions of each OA document. The license data in the XML must conform to LitArch’s guidelines for recording the terms of a license. See License Tagging Information below
- LitArch will include the content (full-text XML and other files) for all OA documents in PMC's Open Access subset (“OA subset”)
- Content for documents in the OA subset will be freely available to users for downloading via the LitArch FTP service, and similar services that allow automated downloading of documents
- Reuse or redistribution by LitArch users of content that is in the OA subset will be subject only to proper attribution of the original source and authorship, and to the license included in a particular document

### 6.2.1. License Tagging Information

Any document that is to be open access must indicate that explicitly in the source XML file. For content that is supplied in the Book Interchange Tag Suite (BITS), see the [license tagging guidelines](#) for detailed tagging information and examples.

For content supplied to LitArch in a form other than BITS, the guidelines are:

1. Include a precise summary license statement in the source XML file.
2. If you are referring to any standard license – such as a Creative Commons license – there should be a URI included for an unambiguous reference.

### 6.3. Book and Chapter Excerpts

Book or chapter records may be submitted to PubMed and the citation in PubMed may include an excerpt. See the [excerpt tagging guidelines](#) for detailed tagging information and examples.

### 6.4. Accessibility

NCBI Bookshelf supports XML tagging practices as defined by the [JATS4Reuse Recommendation on Accessibility](#), ensuring that content is accessible to all users, including those who rely on assistive technologies.

Key accessibility practices, as outlined in the [Bookshelf Tagging Guidelines](#), include:

- Headings: Use structured heading elements to create a clear document hierarchy that enhances navigation for screen readers.
- Image Descriptions: Provide meaningful alternative text for images, figures, and tables, ensuring essential information is accessible to users of assistive technologies.
- Audio and Video: Include descriptions or transcripts for multimedia content, making audio and video accessible to all users.
- Table Structure: Implement clear headers and logical structuring in tables to ensure that assistive devices can accurately interpret data.
- Lists: Use list elements solely for true list content, ensuring proper nesting for clarity, and avoid using list tags for non-list formatting purposes.
- External Links: Clearly tag external links to distinguish them for assistive technology users, aiding navigation.
- Mathematical Content: Represent mathematical expressions using MathML, following best practices for compatibility with screen readers.

## 7. Tools

For content submitted in the BITS Book DTD and marked up as per the Bookshelf Tagging Guidelines, a number of tools are available:

### 7.1. Online XML Validator

Use the [PMC XML Validator](#) to validate XML files against a DTD. The DTD must be identified with a properly formed DOCTYPE declaration. The book XML document must be uploaded as a single file of less than 1Mb.

### 7.2. PMC Style Checker

The [Online PMC Style Checker](#) is an interactive tool which provides a detailed report of all items in a document tagged using the NLM Book DTD that do not comply with the [Bookshelf Tagging Guidelines](#). The report will list items as either warnings or errors. Errors are required fixes, and warnings are suggested fixes.

### 7.3. Math Previewer

Use the [PMC Math Preview Tool](#) to create a GIF or PNG rendering of MathML or LaTeX code.

## Submission sample

Download a [sample submission file](#) corresponding to [this](#) report. See [Bookshelf Tagging Guidelines](#) for additional examples of fully-tagged samples.



## OAI-PMH Service

Updated: July 26, 2018.

The NCBI Bookshelf OAI-PMH service (Books-OAI) provides access to metadata of all items in the Bookshelf archive, as well as to the full text of a subset of these items.

Books-OAI is an implementation of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), a standard for retrieving metadata from digital document repositories. Visit the [Open Archives Initiative site](#) for more information about the protocol and other activities of the OAI group.

Books-OAI supports OAI-PMH version 2.0. It does not support earlier versions of the protocol.

If you have questions or comments about this, or any of the other services provided by the NCBI Bookshelf, please write to [bookshelf@ncbi.nlm.nih.gov](mailto:bookshelf@ncbi.nlm.nih.gov).

## Copyright

Most of the items in this archive are copyright protected, with copyright held by the author(s) or the depositing publisher. In general, the OAI service cannot be used to retrieve the full text of documents in the NCBI Bookshelf. The only exceptions to this policy are for documents that are in the public domain and those that are made available under an Open Access provision (as defined in the Open Access Subset). See the [Bookshelf Copyright Notice](#) for more information.

## Using the Books-OAI Service

The base URL for the service is <https://api.ncbi.nlm.nih.gov/lit/oai/books/>.

See the examples, below, for details about how to use this service to retrieve metadata for books and documents.

## High-Volume Retrievals

If you are using a script that makes more than 100 requests of any kind, please run it outside of the Bookshelf system's peak hours. Also, please make sure that your system does not make concurrent requests, even at off-peak times. Peak hours are Monday to Friday, 5:00 AM to 9:00 PM, U.S. Eastern time.

The Books-OAI Service supports HTTP compression, and the Accept-Encoding HTTP header must be set to `'gzip, deflate'` for the most efficient data transfer. See also [HTTP compression](#) for further information about this capability.

## Access to Full Text

Some Bookshelf participants allow harvesting of the full text of all items, others allow it for only some items, and many do not allow it at all. See the Open Access Subset for specifics.

In addition, the parameter `set=book-open` identifies the complete collection of items in Bookshelf for which the full text may be harvested.

## Supported Data Formats

Records may be retrieved from the Bookshelf archive in either the “NCBI Book Tag Set” or the “Book Interchange Tag Set” XML format for metadata or full-text book records. A DTD, W3C Schema, and complete documentation for these formats are available from <https://dtd.nlm.nih.gov> and from <https://jats.nlm.nih.gov/extensions/bits/>. The prefixes are:

- Metadata: metadataPrefix=nbk\_meta
- Full Text: metadataPrefix=nbk\_ftext

## OAI Books identifiers

The identifiers used by this system are of the form:

```
oai:books.ncbi.nlm.nih.gov:nbkid
```

where *nbkid* is the numerical portion of the document's accession ID. For example, the OAI identifier for the document NBK315783 is `oai:books.ncbi.nlm.nih.gov:315783`.

## Automatic Segmentation of Large Result Sets

If a `ListIdentifiers` request results in more than 500 hits, Books-OAI will return the first 500 with a `<resumptionToken>` that can be used to get the remaining items.

If a `ListRecords` request results is more than 25 hits for full text or 50 hits for metadata, Books-OAI will return the first 25, or 50 records, respectively, with a `<resumptionToken>`.

## Examples

### Identify

- Identify the Books-OAI interface.  
<https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=Identify>

### ListMetadataFormats

- List all metadata formats.  
<https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListMetadataFormats>
- List metadata formats available for a document with the identifier `oai:books.ncbi.nlm.nih.gov:315783`.  
<https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListMetadataFormats&identifier=oai:books.ncbi.nlm.nih.gov:315783>

### ListIdentifiers – return a fixed number of identifiers per request.

Typically, the response to this request will contain 500 identifiers, and will include a resumption token at `OAI-PMH/ListSets/resumptionToken`.

- List (the first set of) identifiers for all documents released in Bookshelf after 1/1/2014 in metadata format.  
[https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListIdentifiers&from=2014-01-01&metadataPrefix=nbk\\_meta](https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListIdentifiers&from=2014-01-01&metadataPrefix=nbk_meta)
- If the above request returned the resumption token `"oai%3Abooks.ncbi.nlm.nih.gov%3A179164!2014-01-01!!nbk_meta!"`, then the following request returns the next set of identifiers:



[https://api.ncbi.nlm.nih.gov/lit/oai/books/?  
verb=ListIdentifiers&resumptionToken=oai%3Abooks.ncbi.nlm.nih.gov%3A179164!2014-01-01!!  
nbk\\_meta!](https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListIdentifiers&resumptionToken=oai%3Abooks.ncbi.nlm.nih.gov%3A179164!2014-01-01!!nbk_meta!)

- List identifier for documents in the setSpec “micad”, released between March to July 2012, and available in full-text XML.

[https://api.ncbi.nlm.nih.gov/lit/oai/books/?  
verb=ListIdentifiers&from=2012-03-01&until=2012-07-31&metadataPrefix=nbk\\_ftext&set=micad](https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListIdentifiers&from=2012-03-01&until=2012-07-31&metadataPrefix=nbk_ftext&set=micad)

## ListRecords – return a fixed number of records per request.

Typically, the response will contain 25 full-text XML records.

- List (the first set of) full text XML for the setSpec “assayguide” released between 3/22/2014 to 6/12/2014

[https://api.ncbi.nlm.nih.gov/lit/oai/books/?  
verb=ListRecords&from=2014-03-22&until=2014-06-12&set=assayguide&metadataPrefix=nbk\\_ftext](https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListRecords&from=2014-03-22&until=2014-06-12&set=assayguide&metadataPrefix=nbk_ftext)

## ListSets – return a fixed number of sets per request

- This response will typically contain the first 500 sets. Use the “resumption token” to get the next set.

<https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=ListSets>

## GetRecord – retrieve a record by identifier

- Get record in Bookshelf metadata XML format with an identifier.

[https://api.ncbi.nlm.nih.gov/lit/oai/books/?  
verb=GetRecord&identifier=oai:books.ncbi.nlm.nih.gov:315783&metadataPrefix=nbk\\_meta](https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=GetRecord&identifier=oai:books.ncbi.nlm.nih.gov:315783&metadataPrefix=nbk_meta)

- Get record in Bookshelf full-text XML format with an identifier.

[https://api.ncbi.nlm.nih.gov/lit/oai/books/?  
verb=GetRecord&identifier=oai:books.ncbi.nlm.nih.gov:315783&metadataPrefix=nbk\\_ftext](https://api.ncbi.nlm.nih.gov/lit/oai/books/?verb=GetRecord&identifier=oai:books.ncbi.nlm.nih.gov:315783&metadataPrefix=nbk_ftext)



## Content Selection for Bookshelf

Updated: January 15, 2021.

The National Library of Medicine (NLM) decides whether the scientific and editorial character and quality of a proposed title merit its inclusion in Bookshelf. In making this decision, NLM first considers the suitability of the title for the NLM collection, based on the criteria in the [Collection Development Guidelines of the National Library of Medicine](#). To apply to Bookshelf, a title must have a clearly stated peer-review process and demonstrate features that contribute to the objectivity, credibility, and quality of its contents.

Examples of the types of criteria that may be considered when evaluating content for Bookshelf include:

Category	Possible considerations
<b>Policies</b>	<ul style="list-style-type: none"> <li>Is the peer review process documented?</li> <li>Does the title clearly state its editorial and ethical policies, and does it adhere to these policies?</li> <li>Are commercial sponsorships clearly addressed (i.e., do not raise questions about objectivity of published content)?</li> </ul>
<b>Figure and table quality</b>	<ul style="list-style-type: none"> <li>Are figures and tables well-constructed and of sufficiently high resolution (i.e., not blurry)?</li> <li>Are figures and tables well-annotated and easy to read and interpret?</li> <li>If figures/images have been re-used from other sources, are permissions, license details, and attribution clearly documented?</li> </ul>
<b>Language quality</b>	<ul style="list-style-type: none"> <li>Is the writing clear, concise, and logical?</li> <li>Does the language impede scientific meaning or cause confusion?</li> </ul>
<b>Formatting and organization</b>	<ul style="list-style-type: none"> <li>Is content organized in a logical fashion and cohesive?</li> <li>Are there indicators of sufficient editorial attention, as evidenced by the elimination of editorial errors (e.g., incorrectly numbered sections, mislabeled tables/figures)?</li> </ul>
<b>Editors and authorship</b>	<ul style="list-style-type: none"> <li>Are full names and affiliations of the editors and authors/contributors provided?</li> <li>Do editors and authors/contributors have appropriate credentials and subject matter expertise?</li> </ul>
<b>Scientific quality</b>	<ul style="list-style-type: none"> <li>Is the content scientifically accurate and unbiased?</li> <li>Is the content substantive?</li> <li>Are statements supported appropriately by citations?</li> <li>Are references appropriate in number and up-to-date?</li> <li>Is the content appropriate to its intended purpose and audience?</li> </ul>
<b>Other</b>	<ul style="list-style-type: none"> <li>For web content and clinical guidelines/resources, is there an appropriate review cycle for communicating and maintaining the currency of the information?</li> <li>For web content, is the website well organized and easy to use/navigate?</li> </ul>

In addition to the above, as outlined in the [Collection Development Guidelines of the National Library of Medicine](#), reviews should generally contain substantive summaries and analysis of recent research in a field. Case reports should include thorough and detailed case presentations, as well as substantial discussions about the relevance to clinical practice and research.

For content that reports the results of original research, the research should be scientifically rigorous. NLM uses the definition of scientific rigor provided by the [NIH Office of Extramural Research](#), which defines scientific rigor as “the strict application of the scientific method to ensure robust and unbiased experimental design, methodology, analysis, interpretation and reporting of results. This includes full transparency in reporting experimental details so that others may reproduce and extend the findings.” NLM also expects original research to adhere to applicable ethical standards, including the protection of human and animal subjects.

There are numerous resources available that can aid authors and editors in improving the quality of their content. As examples, some resources NLM encourages content providers and authors to reference are:

- Monographs: [OAPEN Guides for OA monographs](#)

- Clinical Practice Guidelines: [IOM Committee on Standards for Developing Trustworthy Clinical Practice Guidelines](#)
- Textbooks: [OER Authoring Resources](#)
- Systematic reviews and meta-analyses: [PRISMA](#) guidelines
- Case reports: the [CARE case report guidelines](#)
- Clinical trials: [CONSORT](#)
- Animal studies: [ARRIVE](#)

In addition, publishers should follow current best practices. NLM supports the publishing practices outlined by:

- [Directory of Open Access Books](#)
- [Open Access Publishing in European Networks \(OAPEN\)](#); particularly the Jisc and OAPEN: Publisher information on open access monographs (2016).
- [Open Access Scholarly Publishers Association \(OASPA\)](#)
- [Committee on Publication Ethics \(COPE\)](#)
- [International Committee of Medical Journal Editors \(ICMJE\)](#)
- [Council of Scientific Editors \(CSE\)](#)

## Scientific, Editorial, and Technical standards: Reevaluation

Updated: October 7, 2019.

Ongoing participation in Bookshelf requires a content provider to continue to meet NLM's scientific, editorial, and technical quality standards for Bookshelf and the NLM collection over time.

NLM will reevaluate content for continued participation in Bookshelf if there are problems identified with the content that the provider fails to address or that appear to be systemic; if there are verifiable concerns about the scientific, editorial, or technical quality of the content; or if there are significant changes in the ownership, policies, or practices for producing the content.

The reevaluation process for a Bookshelf-participating journal is similar to the review process for new applications to Bookshelf.

- For details on how content is assessed for scientific and editorial quality, see [Content Selection for Bookshelf](#)
- For details on the technical requirements for participation in Bookshelf, see [File Submission Specifications](#)

Content providers are also expected to continue to meet the terms of specified in the LitArch participation agreements, including:

- Deposit of any revised or updated content soon after the time of publication
- Notification about any change in status of content (such if the conclusions, guidelines, or recommendations presented in a document are no longer valid or effective)

### Scientific and Editorial Reevaluation

Before a reevaluation based on scientific and editorial reasons begins, Bookshelf staff will notify the content provider of NLM's concerns and place a hold on processing any additional content until the review is complete. The reevaluation process will focus on recent content and may take up to 12 weeks.

#### Note.

Content may be reevaluated by NLM only if it has an active LitArch-participation agreement. Some content in Bookshelf is deposited in compliance with the public access policy of collaborating funders and are published by publishers that are not active participants in Bookshelf.

If content was accepted for archiving in Bookshelf by virtue of its inclusion in MEDLINE, the reevaluation process will involve the [Literature Selection Technical Review Committee](#).

If, as a result of the reevaluation, the content is found to currently meet NLM's scientific and editorial standards for Bookshelf, Bookshelf will continue to accept and archive content from the provider, including all content published during the holding period. If the content is found to no longer meet NLM's standards, the Literature Archive agreement will be terminated and the provider will be eligible to re-apply 24 months after the date of the termination. No additional content, including any content submitted or published during the holding period, will be accepted for archiving in Bookshelf.

### Technical Reevaluation

Bookshelf will communicate to a content provider any systemic technical or production issues, including the failure to provide updated content in a timely manner. If the content provider repeatedly fails to respond or successfully resolve these technical and production issues, NLM reserves the right to terminate its LitArch

agreement with the provider. The provider will be eligible to re-apply 24 months after the date of the termination.

NLM's decision to continue or cease archiving content in Bookshelf is final. NLM encourages content providers to use feedback resulting from the reevaluation process to improve the overall scientific, editorial, and technical quality of their content.