

EMMA—Educational Materials Made Accessible

Initial Release—April 18, 2023

Background

EMMA—Educational Materials Made Accessible—is a collaboration between the libraries and disability services offices (DSOs) at its member colleges and universities to facilitate the interchange of resources that have been remediated for accessibility.

Based at the University of Virginia, it consists of federated repositories of content and the technical infrastructure that enables DSOs to search across those repositories to obtain resources they need for the students, faculty, and staff they serve: either resources to remediate or resources that have already been remediated in the ways needed by the particular individuals who require them.

EMMA was created to address a significant problem, the fact that a resource remediated by a DSO is rarely provided to individuals at other institutions who need that same resource remediated the same way. This results in tremendous inefficiency, redundancy, and waste. The same book or article, for example, might be remediated by scores of DSOs over time, when the one first remediated might have been suitable for many or all of them.

This work is often quite labor intensive, and most DSOs are stretched thin, particularly at the beginning of each semester, when they scramble to provide accessible resources to students with disabilities. This can put those students at a distinct disadvantage, especially when the work involved means that they obtain the books and articles they need much later than their classmates do.

The Development of EMMA

The origin of EMMA traces back to a research project in 2015–16. A planning grant from the Institute of Museum and Library Services (IMLS) funded an investigation of the potential value of a shared, inter-institutional repository of accessible instructional materials in higher education.

That project, “Repository Services for Accessible Course Content,” resulted both in an article, “[Toward Accessible Course Content: Challenges and opportunities for libraries and information systems](#),”¹ published in the *Proceedings of the Association for Information Science and Technology* in December 2016, and an extensive white paper, “[Libraries: Take AIM! Accessible Instructional Materials and Higher Education](#),”² published in March 2017. This work documented, among other things, that it is widely but incorrectly believed by DSOs that they are prevented from sharing the remediated files that they create.

That led to a four-year project funded by The Andrew W. Mellon Foundation known as “FRAME: Federating Repositories of Accessible Materials for Higher Education.” FRAME’s mission was to eliminate as much as possible of that wasteful, redundant work by enabling remediated resources to be discovered and shared between responsible parties.

FRAME was a collaboration among academic libraries, repositories, technologists, and DSOs. It involved the libraries and DSOs at eight universities: George Mason University, the University of Illinois Champaign-Urbana, Northern Arizona University, Ohio State University, Texas A&M University, Vanderbilt University, the University of Virginia, and two campuses of the University of Wisconsin.

It also involved the integration of four significant repositories of content useful to students at colleges and universities: Benetech’s Bookshare, the Internet Archive, the HathiTrust, and ACE, the Accessible Content ePortal from the Ontario Council of University Libraries (OCUL). In addition, a fifth repository, EMMA, was created at the University of Virginia for remediated materials not originating in one of those four repositories, along with the technical infrastructure that integrates them.

¹ See <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/pa2.2016.14505301027>.

² See <https://dl.tufts.edu/downloads/d504rx736?filename=fn1079946.pdf>.

The Legal Foundation of EMMA

The first activity of the FRAME project was to convene a group of legal experts at the Association of Research Libraries headquarters in Washington, DC, “The Law and Accessible Texts,” in January of 2019.

That meeting and the subsequent work by those experts resulted in another white paper, “[Reconciling Civil Rights and Copyrights: The Law and Accessible Texts](#),”³ which clearly established that it is not a violation of copyright to provide an accessible version of a resource to a person who has a disability that impairs their ability to fully consume the published version. This is based on both U.S. law (e.g., the [Chafee Amendment](#)⁴) and international law (e.g., the [Marrakesh Treaty](#),⁵ to which the U.S. became a signatory on February 8, 2019).

It couldn’t be clearer: DSOs are permitted to do whatever they need to do to provide properly accessible resources to qualified students and faculty; in fact, their colleges and universities are *required* to do so.

While sharing may be restricted for some resources that have been obtained for remediation under *contractual* relationships that prohibit sharing, there is no *legal* obstacle to sharing remediated resources with qualified recipients. One of the goals of EMMA is to provide a trusted system that makes a wealth of resources available without such contractual limitations.

³ See <https://www.arl.org/resources/the-law-and-accessible-texts-reconciling-civil-rights-and-copyrights/>.

⁴ See <https://www.loc.gov/nls/about/organization/laws-regulations/copyright-law-amendment-1996-pl-104-197/>

⁵ Formally, “Marrakesh Treaty to Facilitate Access to Published Works for Persons Who Are Blind, Visually Impaired or Otherwise Print Disabled”; see <https://www.wipo.int/treaties/en/ip/marrakesh/>.

The EMMA Technical Infrastructure

Most of the four-year Mellon grant period was devoted to developing a new repository for remediated resources that have been obtained from any source other than the individual participating repositories; a unified search for resources across all participating repositories; a user interface that streamlines both the search process and the process for depositing remediated resources; and the metadata that describes remediation processes and relevant aspects so that DSOs can document how they've remediated a resource and so that they can locate resources that have been remediated in the manner needed by an individual end user.

The Repositories

There are distinct differences between the participating repositories. EMMA has been designed to take advantage of what each repository offers while enabling DSOs and librarians at member institutions to search and access all of them in a uniform way.

Benetech's Bookshare

Benetech's [Bookshare](https://www.bookshare.org/cms/)⁶ is the most extensive collection of accessible books in the world, currently providing over a million books in any of five different accessible formats—Word, DAISY, EPUB, audio, and Braille—and also on its online platform. These books are provided free to qualified students in U.S. colleges, universities, and schools, and are available by low-cost subscription (less than \$1.00 per week) to adults who join as members.

Because of Benetech's highly developed technical infrastructure and its expertise in the provision of accessible content, its technical staff has been a significant partner with the technical staff at the University of Virginia in the development of the EMMA infrastructure.

⁶ See <https://www.bookshare.org/cms/>.

The Internet Archive

Best known for the Wayback Machine that preserves the history of over 800 billion web pages on the internet, the [Internet Archive](https://archive.org/)⁷ also provides access to millions of free books, movies, software, music, and other publications and media. Of particular relevance to EMMA is that the Internet Archive provides access to over eight million books and journal articles to [people with print disabilities](https://archive.org/details/printdisabled).⁸

Access to the Internet Archive is provided by a free membership, and access is also integrated with members of two other EMMA repositories, Bookshare and ACE.

HathiTrust

As stated on [its website](https://www.hathitrust.org/),⁹ “HathiTrust is a not-for-profit collaborative of academic and research libraries preserving 17+ million digitized items. HathiTrust offers reading access to the fullest extent allowable by U.S. copyright law, computational access to the entire corpus for scholarly research, and other emerging services based on the combined collection. HathiTrust members steward the collection—the largest set of digitized books managed by academic and research libraries—under the aims of scholarly, not corporate, interests.”

Because HathiTrust is a membership organization, only institutions that are members have access to its content. When resources from HathiTrust are deposited via EMMA, they are returned to HathiTrust and are thus sharable only among HathiTrust member institutions.

ACE

As stated on [its website](https://www.aceportal.org/),¹⁰ “The Accessible Content ePortal or ACE is a growing repository of accessible format texts available to

⁷ See <https://archive.org/>.

⁸ See <https://archive.org/details/printdisabled>.

⁹ See <https://www.hathitrust.org/about>.

¹⁰ See <https://www.ocul.on.ca/node/2192>.

users with print disabilities at participating Ontario Council of University libraries (OCUL) institutions. Aimed at making Ontario's university library collections more accessible, the repository has over 15,000 texts in five accessible formats (2 types of PDFs, Text, DAISY and EPUB).

“ACE operates as part of the [Scholars Portal Books](#) platform. Users with verified print disabilities are given access to browse, search and download texts from the secure repository through their home institutions. Users can also request additional texts to be added to the repository through their accessibility offices.”

EMMA

At its formal launch, EMMA as a membership organization encompasses the four repositories described above as well as a fifth repository, also called EMMA, at the University of Virginia. Other repositories are expected to be added from time to time.

As a repository, EMMA stores all the remediated files from member DSOs that have not been sourced from one of the other four repositories. Remediated files from those repositories are returned to them, but are available to members via the EMMA infrastructure.

EMMA also contains both the bibliographic metadata and, most importantly to its mission, the remediation metadata describing how deposited resources have been remediated. Searching with the EMMA infrastructure searches across all of the repositories, not just the EMMA repository. This is discussed in detail below.

Although searching EMMA is open to anyone, only staff at EMMA member institution libraries and DSOs have access to retrieving or depositing files. The end users—the students, faculty, and staff needing remediated resources—do not have direct access to EMMA other than to search. The DSOs are responsible for ensuring that files are only provided to qualified users with print disabilities who are legally entitled to them.

The EMMA Metadata

The EMMA metadata model was developed primarily in collaboration with the DSOs at the pilot universities. It was important for the system to be designed to fit their needs: what bibliographic metadata do they need in order to find the resources their end users require, and what accessibility and remediation metadata do they need to provide so that other DSOs will be able to identify just the right version of a resource for an individual student or other end user with a particular type of print disability?

The librarians and representatives from the participating repositories also had an important role, particularly in the development of the bibliographic metadata. An important concern was being able to sufficiently identify a remediated resource that uses one of their holdings as a source so that they can properly integrate it back into their systems. And Benetech had a major role in the way the metadata is implemented under the hood: much of EMMA's functionality is designed to align with Bookshare's.

To avoid unduly burdening the DSOs, the required bibliographic metadata for deposit consists of only seven properties: the file name of the remediated resource, its title, its author or other creator, its primary language, the repository it is to be deposited in, the type of resource it is (e.g., text, sound, etc.), and its format (e.g., PDF, Word, EPUB, etc.).

The required remediation metadata is even simpler, consisting of just four necessary properties: whether the whole resource has been remediated or not (often only select chapters of a book may have been, for example); which portions were remediated if the whole resource wasn't; whether the resource being deposited has been remediated (some resources are used "as-is," and some are "born accessible"); and a free text property for any remediation comments the DSO wants to provide.

The user interface for deposit makes it clear which fields are required. The system won't allow a resource to be deposited

without all of those required properties having been provided. But it also enables much richer optional metadata to be provided.

There are twelve additional bibliographic properties that are optional, including such obviously useful ones as identifiers, the publisher name, publication date, rights information, version information, subject, and others. In the interest of not adding work to the over-burdened DSOs, those are encouraged but not required. Part of the design of the EMMA workflow is to enable librarians at member institutions to augment, review, and refine the metadata; they are the experts at bibliographic metadata, not the DSOs. It is expected that in most cases the libraries will flesh out much of the optional bibliographic metadata.

On the other hand, the DSOs are the experts in accessibility and remediation. There are thirteen additional remediation metadata properties that are optional, and it is likely that DSOs will want to provide many of them. In fact, it is likely that they already keep track of them in their own internal workflows. These include properties that record what types of features a resource includes that may present remediation challenges, like tables, math, and images; and properties that record what a DSO has done, such as adding image descriptions or MathML, tagging table headers, and so forth.

The EMMA metadata is described in three companion documents in the suite of EMMA documentation. It is included in summary form in both “Searching for Resources in the EMMA Repositories” and “Depositing Remediated Resources in EMMA,” and it is documented in detail in “A Guide to the EMMA Metadata.”

Searching via EMMA

EMMA’s user interface provides a streamlined, user-friendly process for searching across the many millions of available resources in the five repositories. Plentiful drop-down lists, prompts, and help features were developed and refined based on the piloting done by the staff at the universities participating in the

four-year Mellon grant-funded period and executed by technical staff at the University of Virginia.

The Search dashboard consists of an expandable set of search fields that enables the precision needed when searching across such a vast corpus of content.

- The top-level field provides a drop-down list consisting of “ISBN/DOI etc.,” “Title,” “Creator,” “Publisher,” and “Keyword.” This field can be duplicated any number of times, to search by any or all of these choices, and by more than one entry of a given type.
- The “Sort By” field offers choices of “Relevance,” “Title,” “Publication Date,” “Upload Date,” and “Remediation Date.”
- The “Repository” field provides “(any),” “EMMA,” “ACE,” “Bookshare,” “HathiTrust,” and “Internet Archive.”
- The “Feature” field provides a list of a number of different Braille formats, Speech formats, and PDF formats.
- The “Format” field provides a drop-down consisting of “BRF,” “DAISY,” “DAISY Audio,” “EPUB,” “Braille,” “PDF,” “Microsoft Word,” “Tactile,” “Kurzweil,” and “RTF.”
- The “Accessibility” field provides a list of 39 potential accessibility features that might be relevant for a particular end user, such as “Alternative Text,” “Captions,” “MathML,” “Print Page Numbers,” “Tagged PDF,” “Transcript,” and many others.
- A “Page Size” field enables specification of the number of results to display per page.

These features not only make it easy to enter specifications, they also ensure that typographical errors don’t occur.

Search results are displayed in a hierarchical fashion by title and repository; these can then be expanded to show all of the different

formats available for a title in a given repository. The title-level results show the available bibliographic metadata; the format-level results show available remediation and administrative metadata.

When a DSO has located the resource they want, a click on the “Retrieval Link” in that record will download the file to a location they specify.

This search process is described in detail in the accompanying document, “Searching for Resources in the EMMA Repositories.”

Depositing Remediated Resources

Depositing remediated resources is in many ways complementary to searching for them. Fundamentally, the same metadata and a similar user interface are used for both. But where search is optimized to find the most possible useful results with the least comprehensive metadata—many searches can return accurate and useful results with nothing more than an ISBN—deposits are optimized by having rich, comprehensive metadata.

The deposit process has been designed with the understanding that DSOs are typically overloaded with work, so although it enables a very rich complement of metadata to be provided, it only requires the minimum necessary to make the system work. However, in practice, it has become clear that it is in the interest of the DSOs to provide rich metadata, particularly in recording what they’ve done to remediate resources, as mentioned above, in order to make it possible for other DSOs to locate the best file, having been remediated in the way the individual print-disabled end user they’re serving requires.

To make the process of contributing this metadata easier for a DSO—and to ensure that it’s accurate—the user interface for deposit provides a wealth of drop-down lists of controlled vocabularies, with prompts in most fields and with links to help documentation throughout.

There is even a bibliographical metadata lookup feature to help populate a record based on metadata found in Google Books, WorldCat, and Crossref.

There are four ways to deposit remediated resources into EMMA.

- **Single File Upload.** This is designed to enable a DSO to record their work and deposit files as they complete them.
- **Batch Upload.** This enables DSO's to build "manifests" of multiple resources to deposit. This is a popular option because many DSOs prefer to wait until work slows down later in a semester to deposit the resources remediated that semester, or to deposit batches prepared for each end user.
- **Using a DSO's Own Tracking Spreadsheet.** DSOs often use Excel, Google Sheets, or other similar tools to keep track of their work. These often include much of the information needed by EMMA, but also information that should not be conveyed to EMMA, particularly personally identifiable information (PII) about the students they serve and their disabilities. Moreover, those DSO-specific spreadsheets don't use exactly the same terminology that EMMA requires. EMMA has developed a methodology by which these spreadsheets can, with simple modification of some column headers, be exported to EMMA in a way that EMMA only receives the information it needs, in the form that it needs it. This produces a "starter manifest" that can then be completed in the Bulk Upload process.
- **Using the EMMA Manifest Template.** This is an Excel spreadsheet that has been created to streamline the accurate provision of metadata. It has the column headers that EMMA will recognize, and it provides drop-downs for controlled vocabularies that significantly streamline data entry by the DSOs. Best of all, the DSO is free to add columns to keep track of other information, like the PII about its users, that should not be conveyed to EMMA. It is

expected that some DSOs will choose to use this template in place of the spreadsheets they had previously used.

All of these deposit methods are documented in detail in “Depositing Remediated Resources in EMMA,” a companion document in the EMMA Documentation Suite.

Useful Resources

Here are resources mentioned in this document, as well as other resources relevant to accessibility that may be of interest.

Background

The report of the initial IMLS funded project, “Repository Services for Accessible Course Content”: “[Toward Accessible Course Content: Challenges and opportunities for libraries and information systems](#),” published in the *Proceedings of the Association for Information Science and Technology* in December 2016.

The white paper published by that same project, “[Libraries: Take AIM! Accessible Instructional Materials and Higher Education](#),” published in March 2017.

The white paper establishing the legal foundation for the FRAME project, “[Reconciling Civil Rights and Copyrights: The Law and Accessible Texts](#).”

Two laws on which that legal foundation is based: in U.S. law, the [Chafee Amendment](#), and in international law, the [Marrakesh Treaty](#).

The Repositories

Benetech’s [Bookshare](#).

The [Internet Archive](#).

[HathiTrust](#).

[ACE](#), the OCUL Accessible Content ePortal.

The EMMA Documentation

Searching for Resources in the EMMA Repositories. *(Link to come.)*

Depositing Remediated Resources in EMMA. *(Link to come.)*

A Guide to the EMMA Metadata. *(Link to come.)*

Standards, Guidelines, and Tools

DAISY's [Accessible Publishing KnowledgeBase](#).

The [W3C Web Accessibility Standards Overview](#).

The [W3C Web Accessibility Initiative](#) (WAI).

The [W3C Web Content Accessibility Guidelines](#) (WCAG).

[How to Meet WCAG \(Quick Reference\)](#).

The [W3C EPUB 3 Overview](#).

[EPUB Accessibility 1.1](#).

[EPUB Accessibility Techniques](#).

[The Schema.org Accessibility Properties for Discoverability Vocabulary](#).

[PDF/UA in a Nutshell](#) from The PDF Association.

[Create and Verify PDF Accessibility](#) (Acrobat Pro).

[Providing Accessibility Metadata in ONIX](#).

Benetech's [Image Description Guidelines](#).

DAISY's [WordToEPUB](#) Converter Tool.