Universal OCR for Resource Sharing: Exploring Shared Software to Make PDFs More Accessible

Date: August 2026

Lead Authors

Melissa Eighmy Brown (University of Minnesota)
Scott Prater (University of Wisconsin-Madison)

Contributors

See Project Team Members

BTAA Staff Liaison

Emily Shaw

Report Review & Approval History

Affiliated groups:

- Resource Sharing Coordinators Peer Group
- Digital Strategies and IT Services Peer Group
- Accessibility Peer Group
- Digitization Managers
- BTAA Libraries Executive Committee of the Deans
- BTAA Library Deans and Directors (Sept 2026)

Executive Summary

PDF files are inherently inaccessible to web users who have print disabilities[1]. In many cases, a PDF contains only an image of text and is consequently not readable by machines such as screen readers. Articles and book chapters scanned from print need to have Optical Character Recognition (OCR) applied either during scanning or before the document is posted for a patron, so that a screen reader may "read" the PDF.

While OCR is a necessary first step for creating accessible PDFs, it alone does not satisfy ADA Title II and WCAG 2.1 A/AA requirements, which also require structural tagging, accurate reading order, and alt-text— steps that continue to require manual, labor-intensive remediation.

The BTAA Universal OCR for Digitized Lending Pilot examined whether a shared, BTAA-wide OCR solution could support PDF accessibility remediation at scale. Early in the pilot, the focus centered on the feasibility of a centralized OCR service. Over time, as member workflows were reviewed, local OCR practices evaluated, and emerging vendor-developed solutions tested – including OCLC's Tesseract-based service -- the scope of inquiry broadened to include workflow diversity, vendor readiness, and the practical trade-offs between accessibility, turnaround time, and sustainability.

The pilot team found that a single, universal BTAA-wide OCR solution is *not currently feasible*. Contributing factors include:

- differing institutional scanning and delivery workflows;
- multiple ILL platforms (OCLC, Atlas, Ex Libris) with limited cross-vendor coverage;
- only ~30% of scans exchanged within BTAA;
- persistent OCR accuracy limitations across complex layouts, languages, and variable scan quality; and
- no vendor-agnostic, automated workflow that fits all BTAA environments.

At the same time, the pilot identified important shifts in the landscape that change near-term and long-term decision-making. Most ILL documents—the "80% use case"—can achieve acceptable OCR accuracy when text is created from clean, standard-layout scans. In parallel, major vendors are actively developing integrated OCR features, several of which are planned for release before April 2026, when new ADA Title II regulations take effect. These developments suggest that a hybrid environment that combines improved local scanning practices, selective local OCR, vendor

[1] WebAIM, "PDF Accessibility: Defining PDF Accessibility": https://webaim.org/techniques/acrobat/

automation, and targeted remediation for complex cases offers the most practical and scalable path forward in the near term.

Based on these findings, the pilot highlights several decision points for BTAA and member institutions:

- Content and Context Decisions: Determine when remediation is appropriate based on material type, scan quality, layout complexity, language, and patron need, recognizing that not all content should follow the same workflow.
- Timing Decisions: Balance immediate accessibility needs against evolving vendor capabilities, particularly in advance of April 2026, when regulatory expectations will increase.
- Workflow Decisions: Decide whether remediation should occur locally, through vendor services, or via referral, while maintaining reasonable ILL turnaround times.

The pilot recommends coordinated BTAA action in four areas:

1. **Enable patron requests for remediated documents**, supported by clear communication and local or referred remediation workflows.
2. **Adopt and share BTAA-wide best practices**, including prioritizing born-digital resources and applying BTAA scanning standards.
3. **Pursue coordinated vendor advocacy**, ensuring OCR and accessibility are core requirements in system development and licensing.
4. **Document and share local workflows** to identify common patterns, inform scalable approaches, and shape future shared solutions.

While a universal BTAA-hosted OCR service is not viable today, the pilot demonstrates that coordinated improvements in standards, workflows, and vendor engagement can substantially advance accessibility while maintaining the rapid turnaround expected of high-quality ILL services.

Background

Optical Character Recognition (OCR) is a transformative technology that converts digital images of typed or handwritten text into machine-readable and searchable text as well as being a foundational element for creating accessible documents. A pilot team was formed to investigate the possibility of a universal OCR service for all print materials digitized across the network, with the leading use case of interlibrary loan (ILL) and document delivery. Specifically, the pilot group was charged with conducting an environmental review of OCR practices, evaluating OCR tools

and integrations with ILL workflows, and developing a set of recommendations for implementation.

While performing OCR on digitized content is a step toward reducing access barriers and ensuring equitable access to collections and materials for all patrons, OCR alone is not enough to make a PDF document accessible. The ADA Title II[2] regulations for conventional electronic documents (including PDF files) that will go into effect in April of 2026 require documents to meet the WCAG 2.1 A and AA success criteria. As interpreted by the American Library Association[3], the actual remediation of the document to make it accessible can happen anywhere in the workflow. In practice, OCR is applied more often by the lending library at the point of scanning or before delivering a document to the borrowing (Appendix IV), which is then made available to the borrowing library's patron. These documents are delivered directly to the patron as quickly as possible (usually within 1-2 days).

BTAA library patrons collectively received 273,580 interlibrary loan and document delivery documents in FY 2024. For the 16 member libraries that reported numbers, this equates to approximately 68 documents per day that would need to be reviewed/remediated before delivery to meet ADA Title II regulations. Remediation to make a document fully accessible (e.g. adding structural tags, reading order, and alt text for images and tables) is a substantially manual process that takes anywhere from 20 minutes to several hours per document, depending on the software used, the complexity of the content, and the skill and speed of the human beings who are doing the remediation. Machine learning applications such as the PDF Remediation Tool developed[4] at Ohio State University Libraries show great potential for automating PDF remediation, and there are numerous vendors who provide document remediation services for a fee. However, at current resourcing levels, remediating every document before delivery to the requester is not feasible and would radically transform the ILL service into one where the turnaround time – a key metric used for evaluating ILL service – would be substantially delayed. Currently, remediation to meet accessibility requirements is mostly done on-demand and usually through disability accommodation services outside of the library.

Exemptions for meeting the ADA Title II requirements will vary from library to library according to their local legal counsel. While investigation of these requirements is out of scope for this pilot project, discussions around further remediation of documents will be necessary at many libraries. Processes and tools for remediation will evolve, as will our collective knowledge and

[2] American with Disabilities Act Title II Regulations: https://www.ada.gov/law-and-regs/regulations/title-ii-2010-regulations/
[3] https://www.ala.org/accessibility/applying-rule-library-content#:~:text=Q.%20As%20interlibrary%20loan%20is%20a%20shared%2Dlibrary%20activity
[4] https://smartchallenges.asu.edu/challenges/pdf-accessibility-ohio-state-university

understanding of needs. We will work together in making deliberate and demonstrable progress towards universal accessibility.

As a result of the investigation, the pilot team concluded that universal OCR is a necessary component, but not a sufficient one, for addressing the larger task of meeting accessibility requirements and lowering barriers to access and delivering remediated PDFs through interlibrary loan. This report will focus on the challenges and opportunities of a universal OCR service to complete a necessary step towards making PDFs shared through interlibrary loan more accessible. While we recognize there are other use cases for a universal OCR service, as well as other kinds of digital resources that would benefit from OCR processing, this report focuses on the leading use case.

OCR products and services

OCR of PDF documents can be accomplished locally with proprietary licensed or open-source tools, or it can be outsourced to third parties who perform the OCR on submitted PDF documents. This then returns a PDF with OCR'd text embedded in the document.

OCR tools

See Appendix V for a list of tools used locally by BTAA libraries to create PDF documents with embedded OCR text. Many use Adobe Acrobat, a licensed vendor product; some use OCR tools integrated into their scanning hardware. Two institutions use ABBYY FineReader, a licensed OCR tool that can also embed OCR'd text in a PDF. Only one institution uses Tesseract, a free open-source OCR tool that generates a plain text document from a PDF.

Benefits of local OCR tools

The main benefits of a locally managed OCR tool are lower costs and greater control. OCR tools embedded in scanning software are either included in the price of the scanning system or charged as an additional feature. Licensed products, such as ABBYY FineReader or Adobe Acrobat, do have significant licensing costs, but are generally cheaper than a third-party service, and depending on the licensing, may not charge or restrict OCR processing by the number of pages processed. More importantly, local tools allow the libraries to control at what point OCR happens in their workflow and give them an opportunity to examine OCR text, and if necessary, correct it, at the point in the processing that makes the most sense. Local tools can also be updated, supplemented, or replaced by other tools to improve OCR, and accessibility overall, as needed. Local OCR can be applied to all incoming requests (borrowing institution) or outgoing responses (lending institution), regardless of the ILL system (e.g. ILLiad, RapidILL) used to route the requests.

Limitations of local OCR tools

Local tools require local maintenance and support and additional ILL workflow steps. They are also mostly used at the lending institution during scanning or shortly thereafter, which means that many documents received by the borrowing institution may not have OCR applied and could require local OCR processing and remediation upon receipt, with the result that ILL workflows become more complicated. Licensed tools dedicated to OCR can also be expensive; in many cases it may be easier to contract a third-party ILL service that includes OCR as part of their service package.

Benefits of OCR services

OCR services remove the burden of OCR software and hardware maintenance from the local institution. When bundled with ILL services that are already contracted, they can also be cost-effective and reduce the steps in local ILL document delivery workflows. The three primary ILL service providers used by most BTAA libraries -- OCLC, Atlas, and Ex Libris -- are all actively exploring building OCR into their document delivery services.

Limitations of OCR services

OCR services do have a number of limitations that must be considered. The major factor to consider is ongoing costs, which are charged under a number of different models depending on the provider. Some services charge a flat subscription rate, and others bundle OCR into other services, not all of which may be needed. Other providers may charge by volume (number of pages scanned and OCR'd).

As is to be expected, OCR services have limited flexibility for integration in local workflows and few or no customizable options for the OCR process. Vendors decide on which tools they will use and how they are configured. They generally do not offer post-processing or quality control of the results as part of their service offering. In situations where OCR is part of resource sharing services, the source document provided by the lending institution may be modified and delivered directly to the borrowing patron, without providing the lending or borrowing institutions the opportunity to review and correct the results.

Challenges surfaced related to OCR and ILL documents

Digital Accessibility of PDF documents

ILL and document delivery services typically provide articles and book chapters in PDF format, which is the format currently supported by resource sharing systems. PDF files, especially those scanned from print, are inherently inaccessible. OCR alone is not enough to ensure that a PDF meets accessibility requirements; other modifications to the PDF also need to take place to make the document fully accessible (e.g. structural tagging, reading order, alternative text, and color contrast). Microsoft Word and other document formats lend themselves more easily to formatting for accessibility, but they are an unconventional delivery format.

BTAA Libraries OCR, ILL, and Document Delivery

The following challenges arise in the context of the complex borrowing and lending institution resource sharing workflows, which are described in greater detail in Appendix III.

Where OCR is applied in the workflow: 69% of BTAA libraries apply OCR to some or most documents as a lending library, 94% apply OCR to document delivery (scans provided to local patrons from local collections), and only 2 BTAA borrowing libraries apply OCR to most documents posted for their patrons (Appendix IV). Consequently, borrowing libraries receive many documents from BTAA and other sources without OCR that are delivered to BTAA patrons.

Ideally, all documents received by borrowing institutions would include OCR. However, 75% of BTAA libraries rely on hosted platforms (Appendix VI) which prevents them from installing and using proprietary, licensed software such as ABBYY FineReader on the hosted server where the documents are posted. To enable OCR processing within these hosted environments, integration with an OCR engine would involve collaboration with platform vendors to develop and implement the necessary functionality.

Timing of delivery: Documents are delivered directly to the borrowing library patrons through automated systems (e.g. “Trusted Sender” in Atlas’s Odyssey). Patrons may download documents before the borrowing library has a chance to apply OCR, even with an automated system in place.

How local OCR is integrated with vendor integrations: Resource sharing vendors have recently started evaluating and building integrations for OCR into their systems. Once these services become available, libraries will need to evaluate how best to use these offerings, either in conjunction with or as a replacement for current local OCR workflows.

Multiple resource sharing systems, with some libraries using more than one at the same institution: BTAA institutions use resource sharing services provided by the vendors OCLC, Atlas (ILLiad), or Ex Libris (Rapido and RapidILL), or some combination of the three. While some libraries work exclusively with ILLiad and OCLC systems, others process requests in Rapido as well. In addition to documents residing on self-hosted, Atlas, and OCLC servers, the University of

Wisconsin, the University of Minnesota, and the University of Oregon have a majority of documents residing on RapidILL and AWS servers through their use of Ex Libris's Rapido.

BTAA network sharing: Within the BTAA, only about 30% of article and chapter scan requests are filled by other BTAA members; the remaining 70% are provided by other entities (Appendix VII). As a consequence, a universal OCR service that is implemented on the lender side or by a single vendor will only apply to a small portion of all the ILL requests received by a BTAA library.

As the above description of the variety of workflows and vendors shows, a hosted BTAA OCR service would require integration in a number of different workflows, some unique to a single institution. Consequently, a BTAA-hosted service is beyond the current scope and capabilities of the BTAA Center for Library Programs.

Limitations of OCR

While OCR technology and accuracy has improved tremendously in the past few years, OCR tools are still far from approaching human abilities to read and transcribe text. The following factors can negatively impact OCR accuracy:

- Degraded source image quality. Poor contrast, low resolution, warped text blocks due to tight binding, or skewed pages all make it more difficult for tools to identify and accurately extract text.
- Extraneous markings. These include writing in margins, underlining, and annotations.
- Format and layout of text. Text included in tables or as a caption to a figure or image, or text that itself is part of an embedded image, in a variety of typefaces and locations on a page, are difficult for tools to identify and render accurately in the appropriate location. Multicolumn layouts, such as found in newspapers or many journal articles, and text that wraps around a figure or does not follow a straight line, also pose problems for OCR tools.
- Handwritten text. OCR tools are optimized to work on predictable, standard images of letters or ideograms. Texts that are primarily or entirely handwritten (as opposed to printed texts marked up with handwritten annotations) have idiosyncratic, unique features that deviate significantly from recognized letters or ideograms. Such texts are usually unparseable by an OCR tool, without significant training of the tool (see below).
- Less common languages and scripts. Most tools are designed to render text in the more common languages and scripts found online: the tools are trained on large amounts of materials found on the internet. The smaller the set of sample documents that can be used for training, the less accurate the results of OCR of a document written in that language or script will be. This is also true of historical texts whose orthographic practices differ substantially from current standards.

- Multiple languages in a single document. Some tools are designed to detect a primary language in a document and OCR the document using the dictionary and conventions of that language. Consequently, documents containing more than one language may result in worse OCR results.

OCR tools can be *trained*, which refers to the practice of manually creating a sample set of correct transcriptions of texts from a given corpus, then configuring the tools to use that sample set to correctly transcribe other texts in that corpus. This technique can work well when used to OCR a large homogeneous set of materials, such as the collected handwritten letters of a person, or a large multivolume work set in a single standard typeface and layout. However, it does not scale to the interlibrary loan document delivery use case, where every PDF comes from a different source, with different characteristics.

Opportunities

80/20 Rule

100% accurate OCR on every scanned document is the goal to aim towards, but is impractical at this early stage. Defining what constitutes "good enough" OCR solutions to meet 80% of the use cases will put BTAA libraries in a strong position to improve OCR practices specifically, and PDF accessibility in general, into the future as technologies evolve.

A clean, flat scan of an article in a major language with standard layout and typeface, few images, tables or diagrams, few markings, will likely have OCR that is acceptable for meeting accessibility criteria for text that can be read by a screen reader. Most of the items requested through BTAA ILL partners will meet this criteria with relatively little effort, especially if scanning best practices are followed.

Just as important is defining communication and processing workflows to identify and remediate the other 20% of documents that require more hands-on intervention. This is required not only to meet patron accessibility needs on a case-by-case basis, but also to gather examples and analyze common problems that occur as a prerequisite for improving tools and workflows.

OCLC's implementation of Tesseract in Article Exchange

In response to concerns voiced within BTAA's Resource Sharing Development Partnership and elsewhere, OCLC is working on the integration of Tesseract, an open-source OCR engine, into their ILL electronic delivery system called Article Exchange. OCR can be applied to documents sent from the lending library to the borrowing library. In the spirit of the 80/20 rule described above, the

Universal OCR pilot team began meeting with OCLC to beta test their OCR service, providing test scans on either side of the 80/20 divide, evaluating the results, and providing feedback to the development team.

Limitations:

- Since not all of the BTAA libraries use OCLC exclusively for article and chapter requesting, this is not a solution for all of the documents shared across the BTAA libraries.
- Article Exchange is exclusively an OCLC ILL product and is not suitable for high-volume document delivery (scanned local content for our local patrons).
- Article Exchange is a manual, one document at-a-time process, so it can't be implemented at scale by BTAA institutions. Many BTAA libraries use a network drive to drop PDF files for batch-automated delivery (e.g. Odyssey Electronic Delivery Utility).
- The OCR functionality that OCLC is currently testing does not accept multiple languages for the same document. The only exception at this time is the ability for it to read and OCR both Simplified and Traditional Chinese.

OCLC's Article Exchange/Tesseract integration can be helpful for OCLC customers who are not consistently applying OCR to scanned documents for lending purposes; however, it is not a global solution for interlibrary loan and document delivery.

OCLC and Atlas have both engaged with this BTAA pilot workgroup to help gather requirements and evaluate the prototype OCR services they are developing. Continuing collaboration with them beyond the pilot could help address the limitations listed above.

Other vendor services

As stated in the Background, documents are usually delivered directly to patrons at each of our campuses: all BTAA institutions receive some or most of their documents through Odyssey/ILLiad, except for Oregon, which receives them through Rapido/Alma. Some institutions (Wisconsin, Minnesota) use both Rapido/Alma and Odyssey/ILLiad. In order to realize an OCR solution for all/most documents delivered to BTAA library patrons, it is necessary to work with all of our current resource sharing vendors to make this happen.

Atlas plans to develop an automated batch OCR solution for posted documents delivered via Odyssey from lenders using the OCLC Article Exchange API. They plan to also use this API for the ILLiad document delivery module. The batch API integration into Odyssey is planned for an upcoming version of ILLiad and will default to English language only in the initial release.

Ex Libris has also confirmed that they are developing an OCR service to incorporate into RapidILL behind the scenes. As of the time of writing, no date has been set on when this service will be

available, but they plan for the functionality to be in place by April 2026. With 88% (14 of 16) of BTAA institutions using ILLiad's Odyssey and 75% using RapidILL (3 of 16 currently use RapidILL in Alma outside of Odyssey), this may be enough of a critical mass to ensure the majority of documents delivered to BTAA patrons have OCR, with some limitations with languages.

Perhaps the most critical consideration for vendor services is at what point the vendor applies OCR in a local institution's workflow. Ideally, OCR would be applied *before* the email or text notification is sent to the patron that the item is ready for retrieval. Workflows are also complicated by the fact that some institutions use vendor-host servers to make items available, while other institutions receive and make accessible items on their own server. In certain hybrid vendor/local workflows, documents could be taken off of the local or vendor server, OCR would be applied, and then the documents would be put back on the server for the patron to download. Any vendor solution should be optimized to eliminate the need for local institutions to intervene in the processing and should minimize the end-to-end delivery time as much as possible.

HathiTrust Accessible Text Request and Resource Sharing Service

All BTAA libraries are members of the HathiTrust digital repository. This partnership provides an opportunity to coordinate digitization and accessibility of whole works digitized by member libraries. Most documents in HathiTrust have OCR applied after the image scan. A small number may not have OCR; including some handwritten texts, visual materials, and works in languages that cannot be OCR'd.[5] The quality of the OCR may vary depending on the source material (e.g. tight binding or low contrast).

The HathiTrust Accessible Text Request (ATR) service[6] allows member institutions to fulfill requests for whole works that are under copyright for the purpose of providing accessible materials to patrons. However, requests are handled on a case-by-case basis and usually are submitted after an accommodation need has been declared/documented by the user with local disability services, so this is not a universal solution for accessibility.

The HathiTrust Resource Sharing program, which is slated for go-live in early 2026, allows HathiTrust member libraries to download and share a portion, article, or chapter from HathiTrust to fulfill document delivery and interlibrary loan requests under Section 108 (d)(g) and (i) of U.S. copyright law, as long as the supplying library holds a print copy of the title in their library collection that is currently not on loan.

[5] https://www.hathitrust.org/accessibility/
[6] https://www.hathitrust.org/member-libraries/services-programs/atrs/

Currently, six BTAA interlibrary loan operations—Chicago, Iowa, Maryland, Michigan, Minnesota, and Wisconsin—are engaged in high-quality whole book scanning for users with print disabilities. Since BTAA libraries are already contributing full-book digitizations to HathiTrust, it could make sense to include resource sharing departments engaged in ongoing full-book scanning as contributors to HathiTrust, in conformance with HathiTrust digitization guidelines. However, given the individualized nature of ILL and course material requests across the BTAA network, a data-driven analysis of overlapping needs could help determine the value of such an effort, that is, whether the books digitized by one BTAA library would be likely to be later used by other BTAA libraries.

Local OCR solutions

Most local OCR services use either licensed Adobe products to generate OCR and embed it directly in scanned PDFs, or they use the open-source Tesseract software to generate plain text OCR and embed the text in the scanned PDF as a post-processing step. Some institutions also take advantage of the OCR capabilities of their local scanning hardware.

In addition to local software, many AI providers now provide cloud-hosted OCR generation services that offer significantly high-quality OCR, especially on documents with simple formatting and in more commonly written languages. Services such as Google Cloud Document AI, Microsoft Azure AI Vision, Amazon Textract, ABBYY Vantage, and others incorporate OCR into their service offerings. Costs for these services vary greatly and can be difficult to determine up front, as they generally are scaled to amount of usage. They also require that institutions upload documents to the service for OCR, which may violate licensing agreements or privacy policies in certain cases if the agreement with the OCR service provider does not explicitly guarantee that the documents will not be consumed for any purposes other than generating OCR'd text.

A combination of local OCR, quality control (QC), remediation workflows, and vendor solutions may offer a strong hybrid environment for delivering OCR'd documents. Documents could be OCR'd locally, then sent to patrons with or without passing first through an internal QC step. Documents that fail QC, or are returned by patrons due to insufficient quality OCR, could then be sent to selected, approved AI OCR service(s) to obtain higher quality OCR. This hybrid approach would keep costs lower and reduce delivery times while maximizing OCR quality.

Summary of Evaluation

We cannot recommend at this time any single global OCR solution that would make digitized and electronic resources handled in the ILL context universally accessible. The reasons for this include:

- OCR quality is highly variable; dependent on the quality of original print, quality of scan, and complexity of the document (i.e. multiple languages, images, charts); and currently requires human quality control to meet accessibility requirements;
- Workflows, practices, and tools vary across libraries. Many institutions may have already invested in OCR tools locally, so an external solution will not be one-size-fits-all;
- There is currently no local or vendor automated solution that encompasses all possible avenues of PDF delivery used by interlibrary and document delivery. However, the vendor landscape is rapidly changing and the three main resource sharing vendors (OCLC, Atlas, and Ex Libris) used by all BTAA member libraries are developing OCR offerings in the foreseeable future, which should result in all BTAA library patrons receiving their requested PDFs with OCR already applied.
- PDF accessibility, as defined by WCAG 2.1 A and AA requirements, is more than just OCR, and remediation takes significant time to complete.

Recommendations

These recommendations are scoped around PDF accessibility in the specific context of interlibrary loan and document delivery services with the BTAA; PDF accessibility in general is outside the scope of this pilot. However, any PDF future accessibility pilot should explicitly integrate workflows, tools, and organizational structures that could be used to include digital resource accessibility issues related to interlibrary loan as outlined in this report.

1. Establish and operationalize patron remediation pathways.

Borrowing libraries should provide clearly communicated avenues for individual patrons to request and receive remediated interlibrary loan and document delivery documents.

In accordance with ADA Title II, Section 35[7], BTAA libraries should clearly communicate available avenues for remediation after delivery on document cover sheets, within patron email notifications.

The specific approach to remediation may vary across institutions. In preparation for these requests, all libraries should ensure the following skills, processes, or referral mechanisms are in place:

[7] "...a public entity shall take any other action that would not result in such an alteration or such burdens but would nevertheless ensure that individuals with disabilities receive the benefits or services provided by the public entity to the maximum extent possible." https://www.ada.gov/law-and-regs/regulations/title-ii-2010-regulations/#-35204-duties

- Develop in-house or referral processes to handle requests for remediated documents. As mentioned above, a combination of local OCR, quality control (QC), remediation workflows, and vendor solutions should be leveraged to create a strong hybrid environment for delivering OCR'd documents. How that work is accomplished might be through different services available at each university, and this may or may not be housed within the libraries.
- In-house remediation may involve building relevant technical skills: checking for accessibility, adding OCR, tags, headers, reading order, and alternate text.
- Provide an estimate for the time it will take to complete the remediation.

Implementation across the network: This will be facilitated by the BTAA Resource Sharing Coordinators peer group to ensure support and harmonization for all member libraries.

Deliverable: Documented remediation pathway and patron-facing communication language at each member institution.

2. Establish and document best practices for digital resource lending and borrowing for accessibility.

BTAA Resource Sharing Coordinators, in partnership with other library stakeholder departments, can determine strategic goals and implementation roadmaps for digital resource lending that elevate accessibility as a collective priority of the BIG Collection.

Actions include:

- Prioritize lending of existing e-resources and digitized content over scanned print content.
    - E-resources and digitized content often already contain OCR'd text and are more likely to partially or completely meet accessibility requirements.
    - Participate in the HathiTrust Resource Sharing program once it becomes available to member libraries in 2026. Most documents in HathiTrust have OCR, but the quality varies.
- When feasible, the borrowing library may evaluate acquiring a licensed e-book (purchase-on-demand) or other digital resource to offer a more accessible alternative. However, it is important to note that not all born-digital materials are inherently accessible; some may require additional remediation, while others may be protected by DRM software that prevents resource sharing and/or may have licenses that prohibit resource sharing or modification.

- Accept requests for digital reproductions of whole in-copyright print works for eligible users with disabilities and provide OCR to the file created as permitted under Section 121 (the Chafee Amendment) of the US copyright law (Title 17, USC). This will increase accessibility for users with visual impairments or print disabilities.
- Apply the BTAA Libraries Scanning Standards that are part of the BTAA Resource Sharing agreement. OCR quality is greatly improved by creating scans based on these best practices.

Implementation across the network: This will be facilitated by the BTAA Resource Sharing Coordinators peer group to ensure support and harmonization for all member libraries.

Deliverable: Network-aligned accessibility guidance embedded within BTAA digital lending documentation and scanning standards.

3. Advocate for OCR and digital accessibility as a critical requirement for resource sharing.

Continue to leverage the BTAA's collective expertise, influence, and negotiating power to advance the pursuit of universal accessibility in partnership with resource sharing platform vendors.

BTAA Libraries will:

- Develop and document the BTAA libraries' strategy for advancing universal accessibility in resource sharing across the BTAA (and beyond) through engagement with resource sharing vendors, partnerships, and communities over the next 3-5 years. Forums for engagement and discussion include, but may not be limited to:
    - OCLC (RS4G, SHARES)
    - Atlas Systems (ILLiad, Odyssey)
    - Ex Libris (Rapido, RapidILL, Alma Resource Sharing)
    - Project ReShare
    - Rethinking Resource Sharing Initiative
- Engage with appropriate user groups to prioritize accessibility work and leverage enhancement voting, including through ELUNA RapidILL User Group (Ex Libris), OCLC community forums, and Atlas Systems User Voice.
- Build PDF accessibility into vendor agreements as a critical requirement of interlibrary loan and document delivery service to patrons, including incorporating standardized

license language that addresses compliance with WCAG 2.1[8] (for example, the Library E-resource Accessibility suggested license language[9]).

Implementation across the network: The Resource Sharing Partnerships Advisory Group is well positioned to develop and steward the BTAA's Resource Sharing Vendors Partnerships (RSVP) strategy and coordinate its ongoing implementation, drawing as needed upon the expertise of colleagues in other BTAA Libraries stakeholder groups, such as the Library Accessibility Peer Group and the Digital Strategies and IT Services Peer Group.
The RSVP Advisory Group has been nascent since the spring of 2025. An experienced and knowledgeable long-term Resource Sharing Coordinator who was also an active participant on the Universal OCR project team has volunteered to lead the Advisory Group's formation and begin convening in September 2026 to carry out the deliverable described below.

Deliverable: A documented, cohesive high-level strategy for engaging with resource sharing vendors and partnerships as a consortium to advance universal accessibility and other shared goals over the next 3-5 years, to help guide member engagement and inform forthcoming shared service contract negotiations.

4. Formalize and share documented workflows and tools within the BTAA resource sharing network for remediating PDFs in order to inform scalable, common solutions.

There cannot be any one-size-fits-all solution for remediating PDFs for accessibility, including OCR, that will meet all possible accommodation needs. Most institutions have implemented, or will implement, hybrid solutions that combine in-house library scanning and tools, campus remediation services, and vendor services and workflows. BTAA Libraries can work together to share best practices and recommendations for these hybrid approaches, helping to ensure accessibility needs are met across BTAA campuses. Nonetheless, given the complexities of accessibility remediation, current systems' technical limitations, the idiosyncrasies of each BTAA library and its institutional context, and the individual needs for accommodation that will vary from patron to patron, no global solution can ever be provided that meets everyone's needs.

However, continuing to gather, discuss, and document current workflows as they evolve will shed light on common issues or tasks that may lend themselves to more centralized or shared approaches. For example, this workgroup has already found that all BTAA libraries rely on three vendors for their resource sharing management: OCLC, Atlas, and Ex Libris. Presenting

https://www.w3.org/TR/WCAG21/
[9] https://btaa.org/library/programs-and-services/reports/library-e-resource-accessibility---standardized-license-language

coordinated, common requirements to these vendors to push forward their development of OCR and PDF remediation services would go a long way towards standardizing and simplifying the resource sharing work of BTAA libraries and beyond. Other opportunities for coordinated effort may arise as we work jointly to expand our resource sharing services to better integrate accessibility remediation.

Implementation across the network: The Resource Sharing Coordinators peer group will facilitate structured knowledge sharing and maintain a shared repository of documented workflows.

Deliverable: Network-level inventory of accessibility-related workflows and tools, updated annually.

Project Team Members

Bryan Birchmeier (University of Michigan)
Emily Campbell (University of Michigan)
Rachael Cohen (Indiana University)
Kate Dohe (University of Maryland)
Melissa Eighmy Brown (University of Minnesota)
Brian Miller (The Ohio State University)
Christina Miranda-Izguerra (University of Chicago)
Scott Oostendorp (University of Michigan)
Scott Prater (University of Wisconsin - Madison)
Will Schlaack (University of Illinois at Urbana-Champaign)
Emily Shaw (BTAA)
Andy Spencer (University of Wisconsin - Madison)
Jeff Teichmann (Rutgers University)
Rachel Watters (University of Wisconsin - Madison)

Appendix

I. Acknowledgements

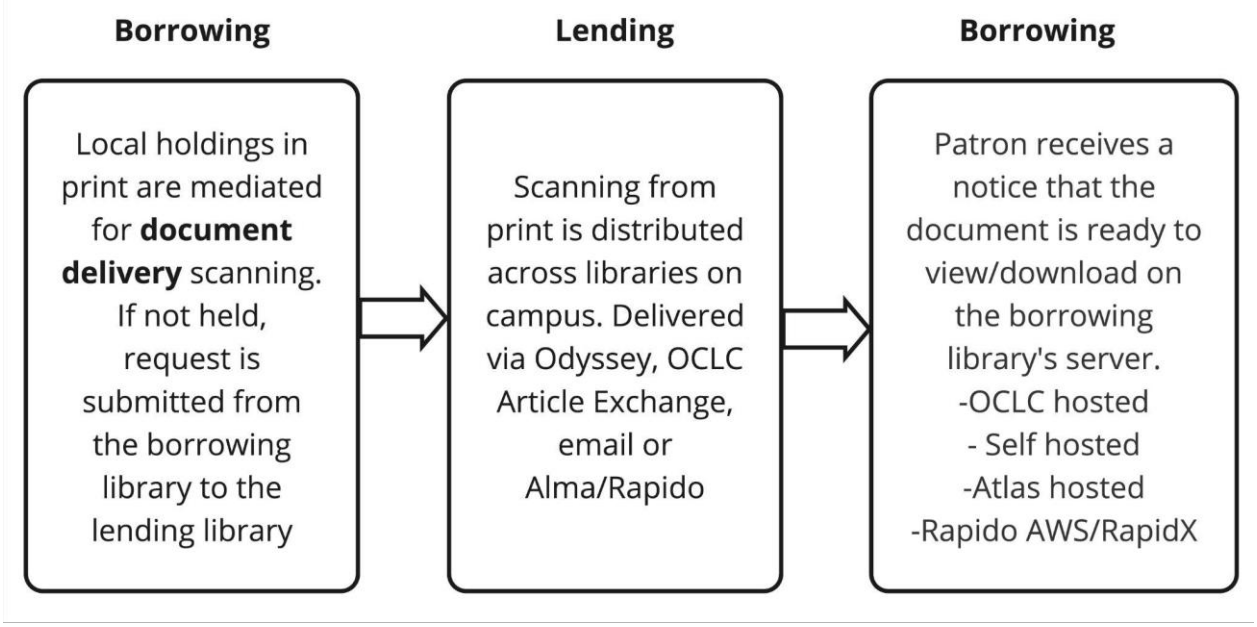
We gratefully acknowledge the support and collaboration of Tony Melvyn at OCLC and Dustin Stokes at Atlas and their teams for their investigation and reporting on OCLC's Article Exchange OCR service under development. We also thank Emily Shaw and Treshani Perera at BTAA for their critical support in making this pilot project possible.

II. Charge

III. The ILL and document delivery workflow

Interlibrary loan and document delivery are time-sensitive services for articles and chapters, typically estimating a 1-2 day turnaround time for delivery. The work that happens behind the scenes is complex. The borrowing (requesting) library submits requests to the lending library in accordance with Section 108 of U.S. copyright law. However, if the material is found to be in the local collection, those requests are filled through local document delivery: electronic resources or scans from print are delivered to the patron. The lending library delivers electronic resources if licensing allows, and print materials are scanned in a distributed model across campus library locations. Documents from lending libraries are delivered directly to patrons and reside on the borrowing library's server, typically for 2-4 weeks.

The percentage of scans provided to other libraries from electronic content (which is usually already OCR'd) ranged from 27% to 84% across BTAA survey respondents, with an average of 55% being supplied from electronic content (i.e. e-journal articles or e-book chapters).



IV. Where OCR is applied

Different OCR tools and delivery systems are used across the BTAA network for interlibrary loan and document delivery. The majority (15 of the 16 libraries) surveyed apply OCR to some or all of their documents. Of those:

- **69% apply OCR to lending scans (scans sent from a providing library to the borrowing libraries' patron)**, and this is **most often accomplished through scanners or OCR scanning software at the point of scanning**. However, some libraries remediate files after scanning.

- **94% apply OCR to document delivery scans (local scans for local patrons**), but a few of those do so only when specifically requested by the patron. The high percentage of BTAA libraries applying OCR to document delivery is most likely due to the fact that this is a smaller set of documents to work with, and these documents are sent directly to local patrons.
- **12.5 % (2 of 16) apply OCR to borrowing documents** by running ABBYY FineReader on their self-hosted ILLiad server.

V. OCR tools used in BTAA Libraries

Libraries may use more than one OCR technology

Software/Tools	# of Libraries noted use
Adobe Acrobat	7
BSCAN	4
ScannX software (purchased with scanner)	3
ABBYY FineReader	2
Bookeye scanner (SCAN2OCR software integration with some models)	3
Book Pavilion (Plustek)	1
Tesseract	1

VI. Odyssey(Electronic Delivery): Server where documents reside

Server platforms for documents	# of BTAA libraries
OCLC hosted	9
Self hosted	4
Atlas hosted	2
Do not use ILLiad	1

Not listed: Rapido AWS and RapidX server for those libraries using Rapido for ILL and document delivery (Minnesota, Wisconsin, and Oregon). Oregon doesn't use ILLiad.

VII. Sharing between BTAA partners

In a one year period, users at the 16 responding BTAA institutions collectively received 192,659 article/chapter documents from external supplying libraries and other providers, averaging 12,041 borrowing requests per institution. Most libraries were able to provide numbers for articles and chapters received from BTAA libraries or supplied to BTAA libraries. Those percentages are in the table below. The average number of documents received from BTAA and supplied to BTAA libraries is 30% of their total requests; however, this varies depending on a libraries' participation in RapidILL. RapidILL members share less within the BTAA network.

Fiscal Year 2024 statistics for articles and chapters

University	Full RapidILL Member?	% Received from BTAA	% Supplied to BTAA
Indiana University	No	28%	41%
Michigan State	Yes	11%	30%
Northwestern	No	82%	45%
Ohio State	No	56%	32%
Penn State	Yes	8%	25%
Purdue	Yes	22%	16%
Rutgers	Yes	21%	20%
University of Chicago	Yes	14%	37%
University of Illinois	No	46%	42%
University of Iowa	Yes	42%	35%
University of Maryland	Yes	19%	26%
University of Michigan	Yes	35%	31%
University of Minnesota	Yes	18%	27%
University of Nebraska	Yes	26%	25%
University of Oregon*	Yes	Unknown	Unknown
University of Wisconsin	Yes	9%	11%
Averages		30%	30%

*These metrics were not available from the University of Oregon at the time of the survey.