

**HATHITRUST DIGITAL LIBRARY  
 REVIEW OF COMPLIANCE WITH TRUSTWORTHY REPOSITORIES AUDIT &  
 CERTIFICATION: CRITERIA AND CHECKLIST (TRAC)  
 MINIMUM REQUIRED ELEMENTS**

The HathiTrust Digital Library (HathiTrust) is a collaborative effort of Indiana University, the University of Michigan, and the University of California (“Repository Administrators”) with support from the charter participating libraries of the Committee on Institutional Cooperation (CIC) and participation by other libraries. HathiTrust is funded in large part by base funding (i.e., not grant or other one-time funding sources) from the Repository Administrators, with contributions from the charter participating libraries of the CIC, and with additional funding made by libraries and library consortia wishing to archive digital content (“Participating Libraries”). The status of this document is a second public release and is currently still incomplete. Although work on this document is ongoing, unless otherwise noted, we consider work on these criteria to be in compliance with the guidelines as set with in the Checklist.

**A: Organizational Infrastructure** ..... 3

**A1.1 Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information.** .....3

**A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.** ..... 3

**A3.1 Repository has defined its designated community/communities and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation requirements will be met.**.....3

**A3.3 Repository maintains written policies that specify the nature of any legal permissions required to preserve digital content over time, and repository can demonstrate that these permissions have been acquired when needed.** .....4

**A3.5 Repository has policies and procedures to ensure that feedback from producers and users is sought and addressed over time.**.....4

**A4.3 Repository’s financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in accordance with territorial legal requirements.** .....4

**A5.5 If repository ingests digital content with unclear ownership/rights, policies are in place to address liability and challenges to those rights.** .....5

**B: Digital Object Management**..... 5

**B1.1 Repository identifies properties it will preserve for each class of digital object.** ....5

**B1.2 Repository clearly specifies the information that needs to be associated with digital material at the time of its deposit (i.e., SIP).**..... 5

**B1.3 Repository has mechanisms to authenticate the source of all materials.** ..... 6

**B1.4 Repository’s ingest process verifies each submitted object for completeness and correctness as specified in B1.2** .....6

<b>B1.5</b>	<b>Repository obtains sufficient physical control over the digital objects to preserve them</b>	<b>6</b>
<b>B1.6</b>	<b>Repository provides producer/depositor with appropriate responses at predefined points during the ingest processes</b>	<b>7</b>
<b>B1.7</b>	<b>Repository can demonstrate when preservation responsibility is formally accepted for the contents of the submitted data objects (i.e., SIPS)</b>	<b>7</b>
<b>B1.8</b>	<b>Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Ingest: content acquisition)</b>	<b>8</b>
<b>B2.10</b>	<b>Repository has a documented process for testing understandability of the information content and bringing the information content up to the agreed level of understandability</b>	<b>8</b>
<b>B4.1</b>	<b>Repository employs documented preservation strategies</b>	<b>8</b>
<b>B4.2</b>	<b>Repository implements/responds to strategies for archival object (i.e., AIP) storage and migration</b>	<b>9</b>
<b>B6.2</b>	<b>Repository has implemented a policy for recording all access actions (includes requests, orders etc.) that meet the requirements of the repository and information producers/depositors</b>	<b>9</b>
<b>B6.4</b>	<b>Repository has documented and implemented access policies (authorization rules, authentication requirements consistent with deposit agreements for stored objects)</b>	<b>9</b>
<b>C:</b>	<b><i>Technologies, Technical Infrastructure and Security</i></b>	<b>10</b>
<b>C1.7</b>	<b>Repository has defined processes for storage media and/or hardware change (e.g., refreshing, migration)</b>	<b>10</b>
<b>C1.8</b>	<b>Repository has a documented change management process that identifies changes to critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities</b>	<b>10</b>
<b>C1.9</b>	<b>Repository has a process for testing the effect of critical changes to the system</b>	<b>10</b>
<b>C1.10</b>	<b>Repository has a process to react to the availability of new software security updates based on a risk-benefit assessment</b>	<b>11</b>
<b>C2.1</b>	<b>Repository has hardware technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when hardware technology changes are needed</b>	<b>11</b>
<b>C2.2</b>	<b>Repository has software technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when software technology changes are needed</b>	<b>11</b>
<b>C3.4</b>	<b>Repository has suitable written disaster preparedness and recovery plan(s) including at least one off-site backup of all preserved information together with an off-site copy of the recovery plan(s)</b>	<b>11</b>
	<b><i>Appendix A: Shared Digital Repository Digital Assets Agreement</i></b>	<b>13</b>
	<b><i>Appendix B: Take-Down Policy - Addressing Copyright Concerns</i></b>	<b>19</b>
	<b><i>Appendix C: Copyright Holder Permission Agreement</i></b>	<b>20</b>

## **A: Organizational Infrastructure**

### ***A1.1 Repository has a mission statement that reflects a commitment to the long-term retention of, management of, and access to digital information.***

The mission of the Shared Digital Repository is to contribute to the common good by collecting, organizing, preserving, communicating, and sharing the record of human knowledge.

### ***A1.2 Repository has an appropriate, formal succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.***

The Repository Administrators have funded HathiTrust Digital Library for an initial five-year period beginning January 2008, with a planned process of review and renewal. A review of the model used for funding and management of HathiTrust is scheduled for the third year of the first five-year period. This and subsequent processes of review provide HathiTrust with an opportunity to develop appropriate plans (e.g., succession) if necessary.

In cooperation with the Participating Libraries and in conjunction with that three-year review, HathiTrust has planned a constitutional convention for early 2011. In that process, HathiTrust will, in collaboration with the Participating Institutions, shape the next stage of governance for operating the repository through this partnership.

Currently, long-term curation of content in HathiTrust is part of the base-funded responsibilities of the University of Michigan Library, and ongoing funding for the Library is provided to that end. The Dean of Libraries reviews this funding and purpose with the Provost annually. Should the funding or organizational imperatives of the University of Michigan Library change, the Library will develop a succession plan and will devote multi-year funding to support of the transition to another host institution.

### ***A3.1 Repository has defined its designated community/communities and associated knowledge base(s) and has publicly accessible definitions and policies in place to dictate how its preservation requirements will be met.***

The primary designated community of HathiTrust is comprised of the faculty and students or other immediate constituencies (e.g., in the case of public research libraries) of the institutions represented by the Participating Libraries. By extension, meeting the needs of this designated community will ensure that HathiTrust meets the needs of higher education more generally. In “collecting, organizing, preserving, communicating and sharing [these records] of human knowledge” to our primary designated community, we will also be meeting many needs of the larger public.

Primary services that the archive provides are long-term preservation of the content held (both bit-level and preservation and format migration) and support for an array of basic uses of that content, including:

- Persistence of object address (OAIS “Reference”);
- Reading (dependent on the user and his/her rights);
- Searching;
- Assembling materials into private and public (i.e., shareable) collections.

HathiTrust relies on external service providers such as Google for many discovery and use functions of the content represented in the repository.

See elsewhere (particularly B1.4-1.5, B1.7, B4.1-4.2, C2.1-2.2) for documentation on mechanisms the HathiTrust archive employs to ensure long-term preservation of this content.

***A3.3 Repository maintains written policies that specify the nature of any legal permissions required to preserve digital content over time, and repository can demonstrate that these permissions have been acquired when needed.***

HathiTrust is devoted to an array of archive and access services in support of the record of human knowledge. As such, all objects in the archive are either in the public domain, have the necessary permissions to support the level of access afforded, or are simply archived in such a way as to ensure an enduring copy of the content. HathiTrust provides reading access only to those publications where permitted by law or by the rights holder. In cases where a rights holder has granted HathiTrust permission to provide reading access to a publication, the administrative office of the University of Michigan Library retains a record of those permissions. The conclusions of all such determinations are registered in a rights database that controls access. All other forms of access are conducted in light of US copyright law and with the guidance of the University of Michigan’s Office of the General Counsel.

***A3.5 Repository has policies and procedures to ensure that feedback from producers and users is sought and addressed over time.***

HathiTrust will respond to any request sent to a publicly-listed administrative email address support account within one business day. These responses may consist of requests for clarification or more information; at the very least, they will acknowledge receipt of the original request. Email requests are tracked in a helpdesk application, and assigned to appropriate staff members on the same business day of receipt.

HathiTrust has an active program of user testing, with the results of these tests influencing interface design and software functionality. The Repository Administrators are active participants in the wider digital library community, and stay current with developments in technology, standards and best practices.

***A4.3 Repository’s financial practices and procedures are transparent, compliant with relevant accounting standards and practices, and audited by third parties in***

***accordance with territorial legal requirements.***

The budget of HathiTrust is a separately maintained “agency” budget, managed by the University of Michigan Library, and other financial components of the operation are represented in the budgets of several University of Michigan Library Information Technology (LIT) departments, including:

1. Core Services: staff (e.g., system administration), hardware and maintenance.
2. Digital Library Production Service: staff (e.g., most publicly available services).
3. LIT administration: staff (e.g., coordination and review).

Other financial components, such as central accounting, exist elsewhere in the University of Michigan Library’s budget structure. Most budget line items associated with support of the archive are identified as part of that activity, wherever practicable.

The University Library’s financial procedures are subject to audits by the University of Michigan Office of University Audits. No cost elements of the archive currently rely on grants or charged-for services. All documented activities are subject to FOIA and may be reviewed with appropriate requests made to the University of Michigan’s Freedom of Information and Policy Administration Coordinator.

***A5.5 If repository ingests digital content with unclear ownership/rights, policies are in place to address liability and challenges to those rights.***

The Repository Administrators make every effort to ensure that they have appropriate rights to ingest and provide access to content. In those cases where the rights are unclear (e.g., when we encounter copyright information that, relative to the work in hand, is ambiguous or contradictory), the Repository Administrators archive but do not provide access to the work. Where possible, the Repository Administrators secure rights to use works that are in copyright. HathiTrust’s policy governing procedures for responding to complaints is represented in the University of Michigan Library’s “take-down policy” on addressing challenges to access rights, available in Appendix B.

**B: Digital Object Management**

***B1.1 Repository identifies properties it will preserve for each class of digital object.***

HathiTrust is committed to preserving the intellectual content and in many cases the exact appearance and layout of materials digitized for deposit. HathiTrust stores and preserves metadata detailing the sequence of files for the digital object. HathiTrust has extensive specifications on file formats, preservation metadata, and quality control methods, included in the University of Michigan digitization specifications, dated May 1, 2007 (<http://www.lib.umich.edu/lit/dlps/dcs/UMichDigitizationSpecifications20070501.pdf>). HathiTrust is committed to migrating the formats of materials created according to these specifications as technology, standards, and best practices in the digital library community change.

***B1.2 Repository clearly specifies the information that needs to be associated with***

***digital material at the time of its deposit (i.e., SIP).***

HathiTrust has a set of requirements for digital objects deposited in the repository that clearly identify the digital object, define the files associated with any given digital object (e.g., the page images that comprise a book), define the relationships between files (e.g., the association of text files with page images for any given page), and identify any technical or administrative metadata. HathiTrust has defined a METS profile for these metadata. Digital objects must be accompanied by METS metadata that conforms to this profile; in cases where no METS metadata are provided, these metadata are created by HathiTrust in the process of ingest and validation.

Status: incomplete - pending additional documentation of specific files and identifier syntax required for valid submission.

***B1.3 Repository has mechanisms to authenticate the source of all materials.***

HathiTrust consists of content from the work by the Participating Libraries to digitize their collections, relying significantly on digitization services provided by Google. In all cases, a bibliographic record with associated digital object identifier is available in HathiTrust's administrative systems for each object ingested into HathiTrust. Additionally, for materials received from other institutions, HathiTrust receives and retains a HathiTrust Digital Assets Agreement and Digital Assets Submission Inventory (see Appendix A).

***B1.4 Repository's ingest process verifies each submitted object for completeness and correctness as specified in B1.2***

The HathiTrust ingest process conducts the following tests on each submitted item:

1. Metadata: internal tests to ensure MARC21 conformance and completeness
2. OCR text: tested for well-formedness using JHOVE;
3. Image files: tested for well-formedness using JHOVE;
4. Metadata in image files: internal tests for consistency with conventions;
5. Digital signatures (MD5 checksums) for all OCR text and image files: checksum verification (see B1.1);
6. Additionally: a one-to-one correspondence is ensured between OCR text and image files.

Note: these processes do not detect completeness problems originating in capture (e.g., missing pages), nor do they detect readability or other subjective problems.

***B1.5 Repository obtains sufficient physical control over the digital objects to preserve them***

HathiTrust collects the following:

1. Submission/deposit agreements: HathiTrust assumes archival responsibility for materials deposited by Participating Libraries or external organizations when a submission or deposit agreement has been signed by an appropriate authority

from that organization or institution. Appendix A is the boilerplate for these agreements. Signed agreements are deposited with University of Michigan Library administration and filed centrally.

2. Workflow documents: The MDP workflow diagram (<http://www.lib.umich.edu/mdp/MBooksFlowchart.pdf>) documents ingest procedures for content from internal digitization processes, vendor-supplied data, and data from both the Google digitization effort and local digitization activities. In the coming year, the diagram in will be replaced with a diagram representing the general mechanisms to be developed by HathiTrust, and with text to further explicate workflow processes.
3. Records of preservation events: As items are ingested, transformed, or other preservation operations are conducted, a cumulative record of these events and the date and time of their occurrence is maintained. The records use the PREMIS markup conventions and are embedded in the METS document corresponding to the item.

Status: incomplete - pending replacement workflow diagram

#### ***B1.6 Repository provides producer/depositor with appropriate responses at predefined points during the ingest processes***

At each of the milestones listed below, logs with the following information (including error reports) are created and provided to the depositing organization:

1. Metadata: upon loading of descriptive metadata, bibliographic records and item records are created with associated identifiers; report includes records that load successfully and, along with error reports, those that failed to load;
2. Digital objects: upon ingest, inventory and validation report for content that is loaded; additionally, at the point that an item record is updated (see B1.7), a report noting that HathiTrust has taken formal responsibility for preservation of a specific digital object;
3. Rights metadata: after metadata and digital objects are loaded, logs reporting outcome of automatic determination of the copyright status of the object, drawn from bibliographic information in the catalog. No access is provided without this step.

Status: incomplete - pending revisions for sharing with an external depositor.

#### ***B1.7 Repository can demonstrate when preservation responsibility is formally accepted for the contents of the submitted data objects (i.e., SIPS)***

After the ingest process is complete for a given item (step 2, above), the administrative database is updated with a special permanent marker indicating that the item has been digitized, is stored in the repository, and is being preserved. The presence of this marker is an indication of formal acceptance.

As part of the regular communications of HathiTrust, the Repository

Administrators provide reports on the activities and status of the archive. Additionally, the Repository Administrators will work with Participating Libraries to develop specifications for a service by which a Participating Library may request an audit of that institution's content with regard to data integrity.

***B1.8 Repository has contemporaneous records of actions and administration processes that are relevant to preservation (Ingest: content acquisition)***

The METS metadata for each document is updated to include information regarding the ingest processes, including preservation actions (documenting the mandatory semantic units in conformance with the PREMIS Data Dictionary 1.0), with associated dates.

***B2.10 Repository has a documented process for testing understandability of the information content and bringing the information content up to the agreed level of understandability***

In addition to basic validation or bit-level integrity of items (see, for example, C1.7), the basic level of understandability for digital objects in HathiTrust is constituted minimally by the completeness and correct sequence of page images for any given volume. For content digitized by Google, HathiTrust relies on an automated mechanism that tests the volume's completeness, legibility and accuracy as a digital surrogate. Where a manual review of completeness, legibility, order and general accuracy has taken place (e.g., in the case of locally-digitized content), HathiTrust registers the fact that this review was performed in the object's metadata. Additionally, in all cases, HathiTrust acts on reports from users about the content that we have online by responding to the user and making an attempt to correct errors.

HathiTrust provides several methods (e.g., e-mail links and report forms in the user interface) for people to report problems. Contact via any of these methods results in a response within one business day. Additionally, the University of Michigan Library Digital Library Production Service (DLPS) maintains an active program of user testing in order to continuously improve access to materials in the repository.

***B4.1 Repository employs documented preservation strategies***

HathiTrust currently ingests only documented acceptable preservation formats, including TIFF ITU G4 files stored at 600dpi, JPEG or JPEG2000 files stored at several resolutions ranging from 200dpi to 400dpi, and XML files with an accompanying DTD (typically TEI or METS). HathiTrust supports these formats because of their broad acceptance as preservation formats and because the formats are documented, open and standards-based, thus giving HathiTrust a means by which it can effectively migrate its contents to successive preservation formats over time, as necessary. The Repository Administrators have undertaken such transformations in the past; moreover, HathiTrust offers end-user services that routinely transform digital objects stored in HathiTrust to "presentation" formats using many of the widely available software tools associated with HathiTrust's

preservation formats. HathiTrust gives attention to data integrity (e.g., through checksum validation) as part of format choice and migration (see also C1.7).

***B4.2 Repository implements/responds to strategies for archival object (i.e., AIP) storage and migration.***

HathiTrust stores AIPs in the repository file system. Each AIP is stored in a separate directory that also contains all files and metadata associated with the digital object. Each AIP contains technical and administrative metadata (e.g., md5 checksums). Each object is described in an associated METS document, also stored in that directory.

Objects stored in the repository are in a restricted set of formats (see B4.1). Each format conforms to a well-documented and registered standard (e.g., ITU TIFF and JPEG2000) and, where possible, is also non-proprietary (e.g., XML). The SDR has migrated large SGML-encoded collections to XML, and Latin-1 character encodings to UTF-8 Unicode. Our success in migrating from older formats to newer formats demonstrates our commitment to our collections and our ability to keep materials in our repository viable. All migrations are documented in change logs.

***B6.2 Repository has implemented a policy for recording all access actions (includes requests, orders etc.) that meet the requirements of the repository and information producers/depositors***

Access to digital objects is currently made possible through two mechanisms: a web-based system and a data API. The web-based system allows users to view page images and search text of items in the repository or just search text, depending on the copyright status of the item. The data API allows objects in the repository, with their associated metadata, to be retrieved by third-party systems. Actions using both mechanisms are logged, and statistics are generated from the logs and made viewable. As currently specified in charter document(s), a formal request for withdrawal will consist of (a) deletion of the specified digital object(s), (b) the creation of a tombstone record, and (c) a time-and-materials basis for providing a copy of contents to the requesting institution.

***B6.4 Repository has documented and implemented access policies (authorization rules, authentication requirements consistent with deposit agreements for stored objects).***

Access to content by users, part of HathiTrust's preservation strategy, is governed by copyright law. HathiTrust provides reading access only to those publications where permitted by law or the rights holder. In cases where a rights holder has granted the University of Michigan Library permission to provide reading access to a publication, the administrative office of the University of Michigan Library retains a record of those permissions. A permissions agreement is attached as Appendix C. Similarly, when partner institutions or organizations deposit materials in the archive, a signed agreement is filed with the University of Michigan Library administration (Appendix A).

Access policies are exercised using IP address detection, user authentication, and geography detection in conjunction with the determined copyright status of each item stored in the rights database. All other forms of access (e.g., computational research and access for users with print disabilities) are conducted in light of US copyright law and with the guidance of the University of Michigan's Office of the General Counsel.

A description of the access policies, an overview of implementation, and a full description of the rights database are available upon request as separate documentation.

## **C: Technologies, Technical Infrastructure and Security**

### ***C1.7 Repository has defined processes for storage media and/or hardware change (e.g., refreshing, migration).***

HathiTrust replaces storage regularly, approximately every 3-4 years or as the usable life of storage equipment dictates. HathiTrust staff members conduct migrations from one storage medium to another using tools that validate checksums internally. (Digital objects are stored both online and on tape, and the online storage system conducts regular scans to detect and correct data integrity problems.) A total file count is done following a large data transfer, and regularly scheduled integrity checks follow.

Status: incomplete - add periodic checksum validation as additional check beyond what the storage system does internally.

### ***C1.8 Repository has a documented change management process that identifies changes to critical processes that potentially affect the repository's ability to comply with its mandatory responsibilities.***

Changes in software releases of all components of the system (from ingest to access) are developed and tested in an isolated "development" environment to prepare for release to production. When ready for release, developers record the changes made and increment version numbers of system components as appropriate using a version control system. New versions of software are released using automated mechanisms (in order to prevent manual errors). Major changes and upgrades in hardware architecture are recorded in monthly reports of unit activity, and thus are traceable to that level of detail.

### ***C1.9 Repository has a process for testing the effect of critical changes to the system.***

See C1.8. Additionally, subsets of production data are available in the development environment to allow developers to ensure proper system behavior before releasing changes to production.

***C1.10 Repository has a process to react to the availability of new software security updates based on a risk-benefit assessment.***

HathiTrust staff apply security updates to the operating system and to networking devices as soon as they become available in order to minimize system vulnerability. As with new software releases, security updates are tested in a development environment before being released to production. Software packages that present a lower security risk and that have a greater potential to affect application behavior (web servers, language interpreters, etc.) are generally installed, configured and tested manually to allow for greater control in managing updates. Software updates are not applied automatically; moreover, updates that present a potential for having an impact on system behavior are applied and tested first in the development environment. If no impacts are seen, HathiTrust staff apply these updates in production after a testing period of at least one week.

***C2.1 Repository has hardware technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when hardware technology changes are needed.***

The primary designated user community of HathiTrust is comprised of the faculty and students or other immediate constituencies (e.g., in the case of public research libraries) of the institutions represented by the Participating Libraries of the HathiTrust effort, as well as the archiving needs of those libraries. The hardware, consequently, is selected to ensure a minimum of outages and sufficient robustness to support a large number of simultaneous users. Hardware is also selected to ensure easy expansion of storage and adaptability of other hardware needed for ingest of content.

Both with regard to end-users and depositing institutions, HathiTrust staff upgrade hardware on a regular basis (i.e., every three or four years), and to help detect more rapid growth in demands, the web server and storage infrastructures have their own performance monitoring that indicate overload conditions.

***C2.2. Repository has software technologies appropriate to the services it provides to its designated community(ies) and has procedures in place to receive and monitor notifications, and evaluate when software technology changes are needed.***

In order to design, build and modify software for the designated end-user community, HathiTrust conducts an active usability program and seeks input from the Strategic Advisory Board of HathiTrust. Similarly, with regard to software development in support of the archiving needs of the Participating Libraries, HathiTrust focuses on the development of highly functional ingest and validation mechanisms. HathiTrust also seeks and responds to guidance from the Strategic Advisory Board with regard to archiving services.

***C3.4 Repository has suitable written disaster preparedness and recovery plan(s) including at least one off-site backup of all preserved information together with an off-site copy of the recovery plan(s).***

HathiTrust consists of highly redundant storage located in two locations separated

by several hundred miles (Ann Arbor, MI and Indianapolis, IN), and is backed up to tape at a third location several miles from the Ann Arbor data center. HathiTrust will develop a disaster recovery plan in the second year of operation (i.e., 2009).

Status: incomplete - pending development of disaster recovery plan

## Appendix A: HathiTrust Digital Assets Agreement

Complete the Agreement form below for submission of digital assets to the HathiTrust Digital Library (“HathiTrust”). Print two copies of the completed form and have them signed by your representative who is authorized to enter into legal agreements. Send these copies, with a Digital Assets Submission Inventory for each discrete submission, to HathiTrust. Upon acceptance, HathiTrust will return one signed original.

NAME OF INSTITUTION: \_\_\_\_\_

ADDRESS: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

NAME OF CONTACT: \_\_\_\_\_

TITLE OF CONTACT: \_\_\_\_\_

EMAIL: \_\_\_\_\_

TELEPHONE: \_\_\_\_\_

### 1. Introduction

The digital assets described in this Agreement will be deposited with HathiTrust for archiving and/or distribution for non-commercial purposes. This Agreement does not request or require transfer of copyright to HathiTrust.

### 2. Definitions

“Agreement” – this document, including all of its terms and conditions and any completed Digital Assets Submission Inventory (DASI)

“Digital Assets” – a collection of computer files that contain digital content (images, texts, sounds, video) and/or data descriptive of the content and its digital format

“Digital Assets Submission Inventory (DASI)” – the document that specifies submission content and which, along with this document and its terms and conditions, constitutes the

## Agreement

“Federate” – to virtually bring together digital assets for access and/or preservation purposes in such a way as to allow transparent searching as if they were a single database for presentation purposes

“Non-commercial Purposes” – use for purposes that do not generate a profit, either directly or indirectly

“HathiTrust Administrators” – The University of Michigan Library, Indiana University and the University of California

“Submission” – Either (a) a continuous stream of data from a particular source (e.g., Google); or (b) a specific body of content characterized by uniform technical and legal characteristics.

### 3. License

The Depositor grants HathiTrust and its agents, operating on behalf of the Participating Libraries, the non-exclusive right to use the Digital Assets for non-commercial Purposes for a term of:

\_\_\_\_\_ Perpetuity

\_\_\_\_\_ 5 years commencing upon execution of this Agreement

Such right shall include, but not be limited to, the right to:

- 3.1 electronically view, present and display to others the Digital Assets, including providing access via the web and, display as permitted by the rights indicated in the Digital Assets Submission Inventory (DASI);
- 3.2 Federate and incorporate the Digital Assets into databases containing other digital assets;
- 3.3 augment or create metadata to enhance service capacity;
- 3.4 electronically store, archive, copy and/or convert the Digital Assets for preservation purposes.

### 4. Depositor’s Obligations

4.1 The Depositor hereby warrants that:

4.1.1 the rights and permissions information as specified in the Digital

Assets Submission Inventory is reliable and up to date;

4.1.2 The information in the bibliographic record, including but not limited to date, place of publication, and whether the Digital Asset constitutes a U.S. government document is reliable and up to date.

## **5. HathiTrust Administrators' Obligations and rights**

5.1 The HathiTrust Administrators shall:

5.1.1 make reasonable efforts to manage the Digital Assets during the term as set forth in section 3 above, developing them as appropriate to make them available for Non-commercial Purposes;

5.1.2 make reasonable efforts to comply with and inform end users of known copyright and user restrictions pertinent to the Depositor's Digital Assets;

5.1.3 cooperate with Depositor to ensure that Digital Assets are replaced or removed as needed to comply with claims related to the Digital Asset's copyright and user restrictions;

5.1.4 obtain express written permission from the Depositor or copyright owner to use the Digital Assets for any Commercial Purpose;

5.1.5 provide basic services including storage, backup, management, fixity checks, and periodic refreshment by copying the data to new storage media;

5.1.6 The HathiTrust software systems and services used to support online access to the Digital Assets are the property of the University of Michigan, and they may be modified periodically as deemed necessary. Depositor will be notified 30 days prior to any modifications that may affect access to the archive.

5.2 HathiTrust may assume custodial responsibility for previously accepted Digital Assets "orphaned" by the dissolution of the Depositor and not formally assigned to the custody of another agency.

5.3 Although due care will be made to preserve the physical integrity of the Digital Assets, HathiTrust shall incur no liability for the loss of or damage to deposited Digital Assets.

**6. Dissolution of HathiTrust**

If HathiTrust dissolves without formal assignation of this Agreement and obligations herein, or otherwise discontinues its management of the Digital Assets, the Depositor may revoke the license effective immediately upon notice to HathiTrust, or in the event of HathiTrust’s dissolution, upon notice to its successor.

**Agreed and Executed on Behalf of Depositor:**

Depositor’s Authorized Representative

Name: \_\_\_\_\_

Title: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

**Agreed and Executed on Behalf of HathiTrust:**

HathiTrust’s Authorized Representative

Name: \_\_\_\_\_

Title: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

## HathiTrust Digital Assets Submission Inventory

### Information about Contributing Institution

Name of Institution	
Website	
Authorized Representative	
Title	
Phone Number	
Fax Number	
Email	
Technical Contact	
Phone Number	
Fax Number	
Email	

### Title of Collection

### Brief Description of Content and Content Subject Area

### Authority of Data/Provenance

Object Type / Material Type Description

Object Type	File Size (estimate in bytes)	# of Objects
Texts - ASCII		
Texts - XML		
Images - TIFFs		
Images – JPEG2000		

Rights (Please indicate copyright status and copyright owner. If permissions have been obtained from the copyright holder to display in copyright items that are part of this submission, please attach the permission documentation to this document.)

Access Restrictions (Please indicate if any access restrictions are applicable to these digital assets indicating to whom and dates of restriction)

Data Submission Method (Please indicate method of submission)

Return from Google	
FTP (zip or tar files)	
disk	
CD	
DVD	

Indicate if data is compressed or uncompressed (please specify file type)

## **Appendix B: Take-Down Policy - Addressing Copyright Concerns**

The University Library makes every effort to ensure that it has appropriate rights to ingest and provide access to content. In those cases where the rights are unclear (e.g., when we encounter copyright information that, relative to the work in hand, is ambiguous or contradictory), we archive but do not provide access to the work. Where possible, the University Library secures rights to use works that are in copyright. Parties who have questions or who wish to contest the use of specific works may contact University Library Information Technology Administration at [lit-info@umich.edu](mailto:lit-info@umich.edu) or

University of Michigan Library Administration  
818 Hatcher South  
University Library  
University of Michigan  
Ann Arbor, MI 48109-1205

With all such communications, please include:

1. A physical or electronic signature of the copyright owner. NOTE: If an agent is providing the notification, also include a statement that the agent is authorized to act on behalf of the owner
2. Identification of the material that is claimed to be infringing or to be the subject of infringing activity and that is to be removed or access to which is to be disabled, and information reasonably sufficient to permit the Library to locate the material. Providing URLs in your communication is the best way to help us locate content quickly.

The University Library will respond to all correspondence within one business day. If the Library is not able to determine, within one business day, that the Library is permitted to use the work(s) in question, the Library will cease providing access to the work until or unless it can determine that such uses are permitted. The Library may work with the University of Michigan's Office of the General Counsel to make determinations about appropriate use. Parties who wish to contest the Library's uses of specific works may, at their discretion, issue a DMCA take-down notice to the University's DMCA compliance officer. For further information on the University of Michigan's DMCA compliance process, please see: <http://www.itd.umich.edu/itua/copyright.php>.

**Appendix C: Copyright Holder Permission Agreement**

AGREEMENT BETWEEN \_\_\_\_\_  
AND The University of Michigan Library CONCERNING *(Title and full citation)*

\_\_\_\_\_

I, \_\_\_\_\_, hereby authorize the University of Michigan to produce digital copies of the above named publication (“the Work”) for its library collection and to make the full text of this publication available to the public in digital form without restrictions.

\_\_\_\_ I also authorize the University of Michigan to make and distribute reprints or other paper copies of the work for noncommercial scholarly purposes.

I represent and warrant to the University of Michigan that I am a copyright holder of the Work with the right to make this authorization because (please initial the appropriate reason):

- \_\_\_\_ I still possess the original copyright that I obtained as author or publisher;
- \_\_\_\_ The copyright was transferred back to me by the publisher or other legal owner;
- \_\_\_\_ I obtained the copyright by transfer, gift, divorce decree, or inheritance;
- \_\_\_\_ Other (please specify) \_\_\_\_\_.

I also represent that the Work does not, to the best of my knowledge, infringe or violate any rights of others. I further represent and warrant that I have obtained all necessary rights to permit the University of Michigan to reproduce and distribute the Work and that any third-party owned content is clearly identified and acknowledged within the Work."

It is understood that this authorization constitutes a non-exclusive, perpetual license, and that I retain all other rights to this work to which I as copyright holder am entitled.

*Signed and Dated*

Grantor  
  
\_\_\_\_\_  
  
\_\_\_\_\_

Accepted (U-M)  
  
\_\_\_\_\_  
  
\_\_\_\_\_