Florida Digital Archive (FDA) 
Policy and Procedures Guide

Version 3.0, May, 2011
Last reviewed May, 2010 without updates
Superseded versions:
  version 2.5, April 2009
  version 2.4, August 2007
  version 2.3, December 2006
  version 2.2, August 2006
  version 2.1, February 2006
  version 2.0, January 2006

This document covers mission, governance, division of responsibilities, archivable materials, rights, services, preservation strategies and other topics of interest to FDA users.

© Copyright 2006 – 2011 by the Florida Center for Library Automation.
<table>
<thead>
<tr>
<th>CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mission..............................................................</td>
</tr>
<tr>
<td>Background.......................................................</td>
</tr>
<tr>
<td>Governance.........................................................</td>
</tr>
<tr>
<td>FDA Affiliates....................................................</td>
</tr>
<tr>
<td>Designated Community.....................</td>
</tr>
<tr>
<td>Terminology: SIP, AIP and DIP...........</td>
</tr>
<tr>
<td>Model of operation and division of responsibilities</td>
</tr>
<tr>
<td>Archivable materials............................</td>
</tr>
<tr>
<td>Rights in archived materials...............</td>
</tr>
<tr>
<td>Billing.................................................................</td>
</tr>
<tr>
<td>Services..............................................................</td>
</tr>
<tr>
<td>Preservation Strategies.....................</td>
</tr>
<tr>
<td>Storage...............................................................</td>
</tr>
<tr>
<td>Security..............................................................</td>
</tr>
<tr>
<td>Continuity of Operations Plan............</td>
</tr>
<tr>
<td>Succession plan.................................................</td>
</tr>
<tr>
<td>Related documentation........................</td>
</tr>
</tbody>
</table>
Mission

The mission of the Florida Digital Archive (FDA) is to provide a cost-effective, long-term preservation repository for digital materials in support of teaching and learning, scholarship, and research in the state of Florida.

In support of this mission, the FDA guarantees that all files deposited by agreement with its Affiliates remain available, unaltered, and readable from media. For supported formats, the FDA will maintain a usable version using the best format migration tools available.

Background

Planning for the FDA began in 2001 in response to the perceived need of the directors of the libraries of the public universities of Florida to ensure the permanent availability of digital library materials such as electronic dissertations. Development was expedited by the award in 2002 of a three-year grant from the Institute of Museum and Library Services (IMLS) which concluded in September, 2005.

In order to implement the FDA, FCLA staff designed and developed the DAITSS application (Dark Archive in the Sunshine State). In November, 2005, an early version of the DAITSS application that lacked dissemination and withdrawal functions went into production for the Florida Digital Archive. In December 2006 the first version of DAITSS with all major planned functionality was completed. The software was released as open source under the GPL license in 2007. In April 2011 a wholly rearchitected and recoded version of DAITSS was installed as DAITSS 2.

The technical design, procedures and policies of the FDA are based on OAIS - Open Archival Information System Reference Model (ISO 14721:2003) and on ongoing work to define and certify trusted digital repositories, including Trusted Digital Repositories: Attributes and Responsibilities (RLG May 2002), the RLG/NARA Audit Checklist for Certifying Digital Repositories (RLG August 2005), and Trustworthy Repositories Audit & Certification: Criteria and Checklist (NARA, et al., February 2007).

Governance

The FDA is run by the Florida Center for Library Automation, a system-wide center of the state universities attached to the University of Florida for administrative purposes. It is under the administrative control of the Director of FCLA.

The FCLA Advisory Board acts as the Advisory Board of the FDA, recommending policy decisions to the FCLA Director. The FCLA Advisory Board
consists of the directors of the libraries of the eleven public universities, a representative of the state university system, a representative from the Division of Community Colleges, and the State Librarian of Florida.

FDA Affiliates

FDA Affiliates are institutions and organizations who have signed Agreements with the FDA allowing them to use the services of the FDA. Institutions eligible to be Affiliates are public university libraries in the state university system, PALMM partners, and others as approved on a case-by-case basis by the Board. A PALMM partner is an institution that has signed a formal partnership agreement with a library in the Florida state university system to participate in one or more PALMM projects. Non-library units within the public university system may submit materials for archiving only indirectly, by arranging for a Library to include these materials in its Agreement.

Designated Community

An Open Archival Information System (OAIS) is defined as an archive that has accepted the responsibility to preserve information and make it available to a Designated Community. The Designated Community is defined as an identified group of potential Consumers who should be able to understand a particular set of information. The designation of one or more Designated Communities is important for a preservation repository, because the repository must assure that archived digital information is independently understandable to that group.

For the FDA, the Designated Community consists of the professional staff of the FDA Affiliates. Staff members interact with the FDA and serve as proxies for the constituencies they serve in the academic and research communities. They must be able to render materials disseminated to them by the FDA and present these materials to users in understandable form. This may require them to write or acquire rendering software, for example METS-based page turners or media players, but it will not require extraordinary efforts, such as digital archeology or the acquisition of obsolete software or hardware.

Terminology: SIP, AIP and DIP

The terms SIP, AIP and DIP appear frequently below.

A SIP, or submission information package, is defined in OAIS as an information package delivered to the repository for archiving. In the OAIS specification, the SIP may contain one digital file or many files. In the FDA implementation, a SIP must follow certain rules documented in the Florida Digital Archive (FDA) SIP Specification.
An AIP, or archival information package, is the contents of the SIP as ultimately represented in archival storage, after additions, deletions, and transformations are performed by the repository.

A DIP, or dissemination information package, is a set of content disseminated from the repository. In the OAIS specification, the DIP may contain content from one or more AIPs. In the FDA implementation, a DIP disseminates a single AIP.

**Model of operation and division of responsibilities**

The FDA model of operation is based on the notion that responsibility for digital preservation is shared between the FDA and its Affiliates. The FDA attempts to ensure that archived materials can be disseminated back to the depositor in renderable form; that is, that the materials can be displayed, played or used as appropriate for the format on software currently available at the time of dissemination. However, the FDA is not itself a system for discovery, public access or rendering. As a result, Affiliates should note these limitations of the FDA:

- Information packages are disseminated from the FDA on request and may not be immediately usable. Disseminated materials may require the provision of appropriate rendering software and/or some transformation, such as the creation of derivative service copies.

- The FDA is responsible only for the materials it archives. If an Affiliate chooses to archive both the master and service formats of a particular item, the FDA will attempt to ensure the future renderability of both files via forward migration and/or other techniques. If the Affiliate archives only the master format, the FDA is not responsible for making service copies.

The following are the responsibility of an FDA Affiliate institution:

- Negotiate an Agreement for the use of the FDA. A template for the FCLA – *Library Agreement* is available on the Florida Digital Archive website. Agreements must be countersigned by representatives of the Affiliate and of FCLA.

- Maintain current Appendices to the Agreement, specifying which materials will be archived and which staff members are authorized to perform or request certain functions.

- Select the content to submit for archiving, and make sure that adequate descriptive metadata is maintained locally.
• Ensure the Affiliate has the right to deposit the content, and to give the FDA all permissions it needs to store, back up, and carry out preservation strategies on the material. Affiliate institutions assume liability for any breach of intellectual property rights occasioned by the deposit, copying and/or preservation treatment of deposited objects.

• Transmit the content to the FDA in the format required by the Florida Digital Archive (FDA)

• SIP Specification. The SIP specification requires a descriptor conforming to the DAITSS METS SIP Profile. If the Affiliate-owned content is transmitted on the Affiliate’s behalf by another institution, the Affiliate is responsible for coordinating such transmittals with the institution making the submissions.

• Retain a copy of all packages transmitted to the FDA until they are successfully archived in the FDA repository. The FDA does not retain copies of packages rejected by the archive.

• Maintain records of what is archived with the FDA. Upon submission to the archive, an Intellectual Entity Identifier (IEID) is assigned by the archive. The Affiliate’s records must associate this IEID with its corresponding local identifier for the content and its locally maintained metadata. Recording the filenames and checksums of any submitted content files and the local identifier and/or title included in the METS descriptor is also recommended. Administrative information may also be useful, such as the date the package was sent to the FDA, the date confirmation of ingest was received via an Ingest report, and additional checksums contained in the Ingest report.

• Use the information in Ingest Reports and Reject Reports to verify the disposition and condition of materials submitted to the FDA. See Archive Services Reports on the FDA website for more information about these reports.

• Work with Florida Digital Archive staff to resolve problems in a timely manner.

• If for any reason the archiving of any materials is no longer required, make a request to the FDA to withdraw that content.

• Request dissemination of content when needed back for any reason, for example, in order to make new derivatives from a master.

The following are the responsibility of the FDA:
• Ingest and store materials submitted by Affiliates in accordance with the Affiliate’s Agreement.

• Restrict authorization to submit materials, withdraw materials, disseminate materials, and request reports to the individuals specified in the Agreement with the Affiliate.

• Provide detailed Ingest or Reject information for every SIP received.

• Preserve original files exactly as submitted, with demonstrated integrity, viability and authenticity.

• For files in supported formats, use appropriate preservation strategies to ensure a renderable version of the file can be disseminated at any time.

• Provide dissemination information packages (DIPs) on request. The DIP will contain the original version of ingested files. It may also contain modified renderable versions (normalized and/or migrated) when appropriate.

• Provide appropriate reports to Affiliates for management purposes.

• Manage the archive transparently, as required for Trustworthy Digital Repositories.

### Archivable materials

An FDA Affiliate may archive any content in any format, as long as the Affiliate has appropriate rights (see Intellectual Property Rights below). Content to be archived is specified in Appendix A of the Agreement between FCLA and the Affiliate. The Affiliate is considered the "owner" of the content. For more information about specifying content to be archived, see Setup Procedures for New FDA Affiliates on the FDA website.

Appendix A can be modified by the Affiliate at any time by providing FCLA with a dated replacement in writing via email. The revised Appendix must be sent from the email account of the primary administrative contact for the Affiliate, as listed in the Agreement. Modifications will be implemented by FCLA within 10 working days.

### Rights in archived materials

The Affiliate retains all ownership and management rights in materials owned by the Affiliate.
The Affiliate is responsible for compliance with all applicable copyright laws and other laws applicable to deposited materials, and must have the authority to grant FCLA non-exclusive rights to copy, render and create derivative versions of deposited files.

If FCLA receives notification in writing that any digital file is held in the FDA in violation of applicable law, the FDA will disseminate a DIP to the owning Affiliate and withdraw the entire AIP containing the file. If this is challenged and resolved in favor of the Affiliate, FCLA will re-ingest the material at no charge.

Billing

FCLA does not charge for the use of the FDA. It is possible that at some time in the future, billing will be instituted with the advice and consent of the FCLA Advisory Board.

If billing is instituted, written notice of no less than 180 days will be provided to Affiliates. No charges will be incurred by Affiliates for FDA services provided up to the time that billing is implemented. If an Affiliate wishes to withdraw any of its materials from the FDA before billing is implemented it may do so at no charge by following standard procedures for withdrawal requests.

If billing is implemented, the billing algorithm and rates will be posted on the FDA website along with terms and conditions of payment established by FCLA. Affiliates who allow their content to remain in the archive commit to meeting their financial obligations to the FDA.

Services

Standard services of the FDA are listed below. Special services, such as the automatic creation of METS descriptors for a set of content files, may be available upon request. Contact FDA staff for information.

1. Ingest of SIPs.

SIPs transmitted according to FDA instructions will be submitted into the archive. For every SIP submitted, the Affiliate will receive an Ingest Report for successful ingest or an Reject Report if the SIP is rejected. Ingested SIPs are transformed into Archival Information Packages (AIPs) for storage.

As part of the submission process, files not described in the SIP descriptor are discarded. Content files are checked for viruses and for conformance to format specifications (validity). Finding a file validation error will not normally cause the SIP to be rejected, but it will be recorded as an anomaly, or warning, in the metadata stored for the file and reported to the Affiliate in the Ingest Report.
Note that if changes are required to a package already submitted to and archived by the FDA, the package cannot be changed by resubmitting the SIP. The FDA has no ability to update archived content. To correct a package without adding a duplicate, the existing AIP must first be withdrawn before submitting a SIP containing the corrected content. (See item 5. Withdrawal, below).

2. Secure storage and management of AIPs.

This includes maintaining onsite and offsite copies, and periodic fixity checking. It also includes periodic copying to new storage media (media refreshment).

3. Full preservation for supported formats.

Files in all formats will receive secure bit-level preservation treatment. Full preservation treatment will be performed for files in supported formats. Supported formats are listed in the list of Supported and Recognized formats list on the FDA Format Information website's format information page.

Full preservation includes bit-level preservation of the originally submitted files, as well as services intended to ensure that the information content of the files will remain usable into the indefinite future. These services vary according to the file type but may include the creation of normalized forms of the file and/or the reformatting of obsolete formats to reasonably comparable successor formats. See Preservation Strategies below for more information.

The FDA does not guarantee that normalized or migrated versions of any file will be identical in appearance or behavior to the original file. Note also that if an AIP contains content files in both supported and unsupported formats, there is no guarantee that the intellectual entity will remain usable as intended.

4. Dissemination.

Archived content will be disseminated by FTP upon the request of an authorized agent of the Affiliate. (Note that dissemination does not remove content -- the AIP remains in the FDA unless it is explicitly withdrawn.)

The DIP always includes the original version of all files deposited in the original SIP. If a preferred version of any file has been created, the DIP will also include the "last, best" version of the file. The DIP Descriptor is a METS file with a structural map section (<structMap>) describing the original set of files and, if applicable, a second structural map section describing the “last, best” set of files.

As part of the dissemination process, the AIP is refreshed, at which time file formats are re-identified using the most current identification routines, and files are migrated or normalized using the most current file processing routines. This
updated AIP is what is returned to the Affiliate as a DIP. This process ensures that the disseminated package is always as complete and up-to-date as possible.

5. Withdrawal.

(Withdrawal is temporarily unavailable in DAITSS 2.0.) An AIP will be withdrawn from the FDA upon the request of an authorized agent of the Affiliate. All withdrawal requests must be sent to the FDA Help Desk at FDA@prb.fcla.edu. Files belonging to the withdrawn AIP are deleted entirely from storage, but the FDA retains a permanent record that the intellectual entity was ingested and withdrawn. The Affiliate will receive a Withdrawal Report by email.

A common reason for withdrawal is to update information in a previously submitted package. The existing AIP must be withdrawn, and a new SIP must be submitted for Ingest.

An AIP may be withdrawn on the initiative of the FDA, if the FDA receives information that the AIP has been archived in violation of copyright.

6. Reporting.

The FDA provides periodic statistical reports to Affiliates about their own use of the FDA. General statistical reports are posted to the FDA web site. Ad hoc reporting is also available upon request.

Preservation Strategies

Preservation strategies supported by the FDA are based on format transformation, that is, changing file formats to delay or accommodate format obsolescence. The FDA performs two kinds of transformations:

- **Normalization.** If a file is in a format considered to be less than optimal for digital preservation a version of the file may be created in a more preservation-worthy format. In general, preferred formats are non-proprietary, well documented, and well understood by FDA staff. Normalized versions may not be equivalent to originals in appearance or functionality. For example, a PDF file (WAV example) might be normalized into a set of page-image TIFFs. In this case the appearance of the content is retained but functionality such as actionable hyperlinks is lost. If normalization is part of the Action Plan for a particular file format, files in that format will be normalized on ingest. The normalized version will not be stored in the AIP, but the entire SIP will be rejected if normalization fails. This ensures that normalization can be done if necessary, but spares the cost of storing normalized versions.
Migration. If a file is in a format considered at risk of obsolescence, a version may be created in a format considered to be a reasonable successor to the original format. All effort will be made to retain the appearance and behaviors of the original version, although this can not always be guaranteed. The successor format may be a higher version of the original format (for example, PDF 1.4 might be migrated to PDF 1.6) or it may be another format. If migration is part of the Action Plan for a particular file format, files in that format will be migrated on ingest.

The preservation strategies that will be implemented for any file format are documented in the Action Plan for the file format, available on the FDA website. Action Plans are reviewed periodically and revised when appropriate.

All preservation strategies are applied at the time a SIP is ingested, as part of ingest processing. This includes packages that are disseminated and then re-ingested, either as part of the Archive’s planned preservation processes or as part of an Affiliate-requested dissemination (see next paragraph). Normalized and migrated versions of files contained in the SIP become part of the AIP.

If there is not an implemented Action Plan for the file format, bit-level preservation will be carried out for the file until the time when full preservation becomes available. At that time, the AIP containing the file can be disseminated and re-ingested, causing the full preservation treatment to be applied.

Storage

For every file in the AIP, two master copies are written. One copy is stored at the UF Computing & Network Services facility in Gainesville (CNS) and one copy is stored at the Northwest Regional Data Center in Tallahassee (NWRDC).

The two master copies are treated as a single file by DAITSS, the repository software application underlying the FDA. This means that when any action is performed on a file, it must be successfully performed on both master copies to be considered complete. For example, a fixity check involves calculating a message digest over the bits of a file and comparing this to a previously stored message digest. For a fixity check to be complete, message digests must be calculated for both of the master copies of the file and verified to match the stored message digest.

In addition to the master copies, traditional backup copies on tape are maintained in Gainesville and Tallahassee.

Security

Data security is ensured by a combination of physical security and cybersecurity.
The UF Computer Network Services (CNS) in Gainesville and the North West Regional Data Center (NWRDC) in Tallahassee are responsible for the physical security of computer tapes containing archived data, computer servers running the archive application, and other computer hardware necessary for the operation of the FDA.

CNS is a secure data center at the University of Florida. The central machine room and all of the core network fiber huts are secured with an electronic lock system based on the Lenel OnGuard software and Lenel controllers. All doors that open on to public space are configured to fail to a secure state. Alarm conditions are investigated by operations staff or reported to the University Police Department at operations' discretion.

Proximity cards and fobs are used for access. Access rights are granted according to work requirements and staff interaction needs. Most technical staff have 24/7 access to their work areas (including the machine room and fiber huts in most cases) and work hours access to other areas where there are staff members. For access during non-work hours, PIN codes are required in addition to the card or fob.

Air handlers and leak detectors are continuously monitored by operations staff. Backup power is provided by a single UPS for all computer equipment and a diesel generator which can power the UPS and the air conditioning systems for 24 hours without refueling.

NWRDC was designed and engineered to be a state-of-the-art data center that could guarantee customers' security, accessibility and connectivity. NWRDC employs an advanced security system and is monitored 24 hours a day, 7 days a week. Exteriors, doorways and hallways are under continuous recorded video surveillance. Multiple security points are in place throughout the entire building. Air temperature and humidity levels are continuously monitored and controlled. Redundant cooling systems are implemented. To protect equipment and data, NWRDC is equipped with redundant power and two generators that can run at full capacity for weeks with emergency refueling as needed, a smoke detection and fire suppression system, and lightning protection.

Cybersecurity refers to computer and network security. FCLA has a designated security administrator who enforces security procedures and maintains security tools that exceed the best practice recommendations of the University of Florida. The FDA is a "dark archive" with no online public access to stored data. Network access to archived data and to the FDA server computers is limited to a small number of FCLA staff working from known IP addresses. Unauthorized attempts to gain access to data and server computers are logged and reviewed daily for unusual security threats. Security patches to the operating systems are checked for daily and installed within 24 hours of release.
Continuity of Operations Plan

A Continuity of Operations Plan, or COOP, is a tested plan to ensure that an essential service can continue to function across a wide range of potential emergencies, including localized acts of nature, accidents, technical failures, and physical or technological attacks.

The primary computer server for the FDA resides at CNS in Gainesville, FL. A backup server capable of running the DAITSS software is in place at the NWRDC in Tallahassee. According to the COOP for the FDA, a hardware failure at the primary location will result in a switch over to the backup server within hours of the failure. When the FDA is running in COOP mode on the backup server, Affiliates may or may not be able to submit new SIPs, depending on the nature of the disaster. No SIPs will be ingested. However, disseminations will be available on request.

Please contact the Florida Digital Archive staff if you have specific COOP questions.

Succession plan

In the event that the Florida Digital Archive ceases production operations for any reason, FDA Affiliates will have two options. In option 1, content archived for the Affiliate will be returned to the Affiliate as FDA Dissemination Information Packages (DIPs). In option 2, the content will be sent to another preservation repository of the Affiliate's choice in either the FDA DIP format, or the RXP (Repository eXchange Package) format. The RXP was developed to enable heterogeneous preservation repositories to transfer archived AIPs without loss of provenance.

An Affiliate will be able to specify that option 1 or option 2 applies to all AIPs for that account. Alternatively, if the Affiliate can identify exclusive subsets of content by project coding, the Affiliate can request different options for different subsets.

Related documentation

Related documentation appears in the Digital Archive Information and Software and Documentation sections of the Florida Digital Archive website.
Florida Digital Archive Procedures

The following procedures and processes implement the above-stated FDA policies.

Note that in DAITSS 2 there are three separate phases of processing that result in an AIP: *Transmittal* (or deposit of SIPs to the FDA), *Submission* of SIPs to DAITSS software, *Ingesting/archiving* of the successfully submitted package into the FDA repository.

**Transmittal of packages (SIPs) to the FDA.**

*Transmittal* of packages to the FDA can be made by one of the following methods, or by special arrangement with the FDA:

1. **Direct online submission via the DAITSS 2 user interface.** (This process both *transmits* a SIP and *submits* it for processing.)
2. **Delivery to FCLA’s ftp server.** SIPs deposited on FCLA’s ftp server will be moved to the FDA production server daily at 9am for submission to the DAITSS repository software for processing.
3. **Delivery to FCLA via external hard drive.** SIPs delivered to FCLA via external hard drive must be accompanied by a completed transmittal form signed by an individual authorized to deposit materials per the institution’s Appendix A of their Agreement with FCLA. SIPs delivered via external hard drive will be submitted by FDA operators.
4. **Delivery to FCLA’s ETD Service.** Materials delivered to FCLA’s ETD service are also automatically delivered to the FDA production server daily at 9am for submission to the DAITSS repository software for processing.

*Transmittal* (deposit) of SIPs must be performed by an individual authorized to deposit materials in the FDA, per the institution’s Appendix A of their Agreement with FCLA, or by an individual authorized to submit materials to FCLA’s ETD Service.

Each transmitted SIP will be submitted to the FDA’s DAITSS repository software for processing as described below.

**Submission of packages (SIPs) to the FDA’s DAITSS software and SIP validation.**

After SIPs are transmitted to the FDA, they are submitted to the DAITSS software by one of two methods:

1. **Direct online submission via the DAITSS 2 user interface.** This process does the following:
   a. assigns the SIP a package ID (Internal Entity ID),
b. validates the SIP and,
c. if the SIP is valid, copies it to the DAITSS workspace for ingesting. 
   If the SIP is invalid it is rejected and deleted from the FDA server. A 
   reject report is displayed and delivered via email or ftp, and is 
   stored in the DAITSS database for future reference.

2. Submission via a DAITSS Operator. For SIPs transmitted to FCLA by ftp 
   or external hard drive, the SIP is copied to an FDA staging area and is 
   submitted in batches by a DAITSS Operator. Batch submission 
   processing is otherwise identical to that of direct submission via the 
   DAITSS 2 user interface.

Ingesting/archiving of packages into the FDA repository

If a SIP submission is successful (the SIP is valid and not rejected and deleted in 
the submission step), the package is processed and stored in the FDA repository 
as an AIP, and an ingest report is produced and sent to the submitting institution, 
either via email or it is sent to the institution’s directory on FCLA’s ftp server.

Requesting Dissemination of AIPs

Archived content will be disseminated by FTP upon the request of an authorized 
agent of the Affiliate. (Note that dissemination does not remove content -- the AIP 
remains in the FDA unless it is explicitly withdrawn.) All dissemination requests 
must be sent to the FDA Help Desk at FDA@prb.fcla.edu. The dissemination 
package (DIP) will be placed in the Affiliate’s FDA FTP directory, and the Affiliate 
will receive a Dissemination Report.

Requesting Withdrawal of AIPs from the FDA repository

(Withdrawal is temporarily unavailable in DAITSS 2.0.) An AIP will be withdrawn 
from the FDA upon the request of an authorized agent of the Affiliate. All 
withdrawal requests must be sent to the FDA Help Desk at FDA@prb.fcla.edu. 
Files belonging to the withdrawn AIP are deleted entirely from storage, but the 
FDA retains a permanent record that the intellectual entity was ingested and 
withdrawn. The Affiliate will receive a Withdrawal Report.

Changing AIP contents

Note that if changes are required to a package already submitted to and archived 
by the FDA, the package cannot be changed by resubmitting the SIP. The FDA 
has no ability to update archived content. To correct a package without adding a 
duplicate, the existing AIP must first be withdrawn before submitting a SIP 
containing the corrected content. (See item 5. Withdrawal, below).