

Chapter 6:

DAITSS Workflow Interface

Topics covered in this chapter:

- ✓ [A brief glossary](#)
- ✓ [A DAITSS workflow diagram](#)
- ✓ [DAITSS Processing Status vs. Operations Events](#)
- ✓ [Suggested workflow steps for managing repository processing](#)
- ✓ [The Package Details page](#)
- ✓ [DAITSS Task processing steps](#)
- ✓ [A tour of the DAITSS Workflow Interface](#)

DAITSS Workflow Interface

DAITSS Operators interact with a DAITSS repository primarily through the DAITSS Workflow Interface. The only exception is the command-line submission function that enables Operators to submit batches of SIPs to DAITSS with one command.

The DAITSS preservation database maintains a record of all actions taken on packages within the DAITSS system, including those performed by DAITSS Operators via the Workflow Interface.

This chapter provides an overview of the DAITSS Workflow Interface and instructions for specific work flow steps.

Brief Glossary

Active job: A job/WIP in the workspace that has started but not yet completed processing. This includes jobs that are running successfully with a status RUNNING or IDLE, have an ERROR status that indicates that processing has stopped due to a system or program error, have a STOPPED status that indicates that an Operator has manually stopped the job, or have a DEAD status that indicates that the task process has died or has been killed by a System Administrator. Active jobs are listed in the “workspace” tab of the Interface.

DAITSS Affiliate User: A DAITSS Affiliate is one of two user types in DAITSS, and is intended for content owners who are not staff of the repository. An Affiliate User can view only the “dashboard” and “packages” tab and can perform the following actions:

- Submit packages for the Account to which they’re assigned
- View packages in the Account to which they’re assigned

DAITSS Operator User: A DAITSS Operator is one of two user types in DAITSS, and is intended for repository staff. An Operator user has maximum privileges in the Workflow Interface.

Operator Action: An action performed on a single WIP or multiple WIPs by a DAITSS Operator. Possible actions are:

- Start: starts a WIP with a processing status of IDLE
- Stop: stops processing on a WIP with a processing status of RUNNING
- Reset: resets to IDLE a WIP with a processing status of ERROR
- Do over: starts processing of a WIP’s task from the first step (Used only in cases where a major code update requires it.)
- Stash: moves a WIP that is not running from the workspace to a stash bin
- Unstash: moves a WIP from a stash bin back into the workspace

Preservation Task: A DAITSS preservation task: ingest, dissemination or withdrawal. (Please note that a third task type, “d1refresh”, is used only by the Florida Digital Archive and will be removed from subsequent releases of the DAITSS Workflow Interface.) Each DAITSS WIP has one of these preservation tasks assigned to it. DAITSS will perform the necessary steps defined for a specific task on any given WIP. DAITSS will only perform one task on any given WIP.

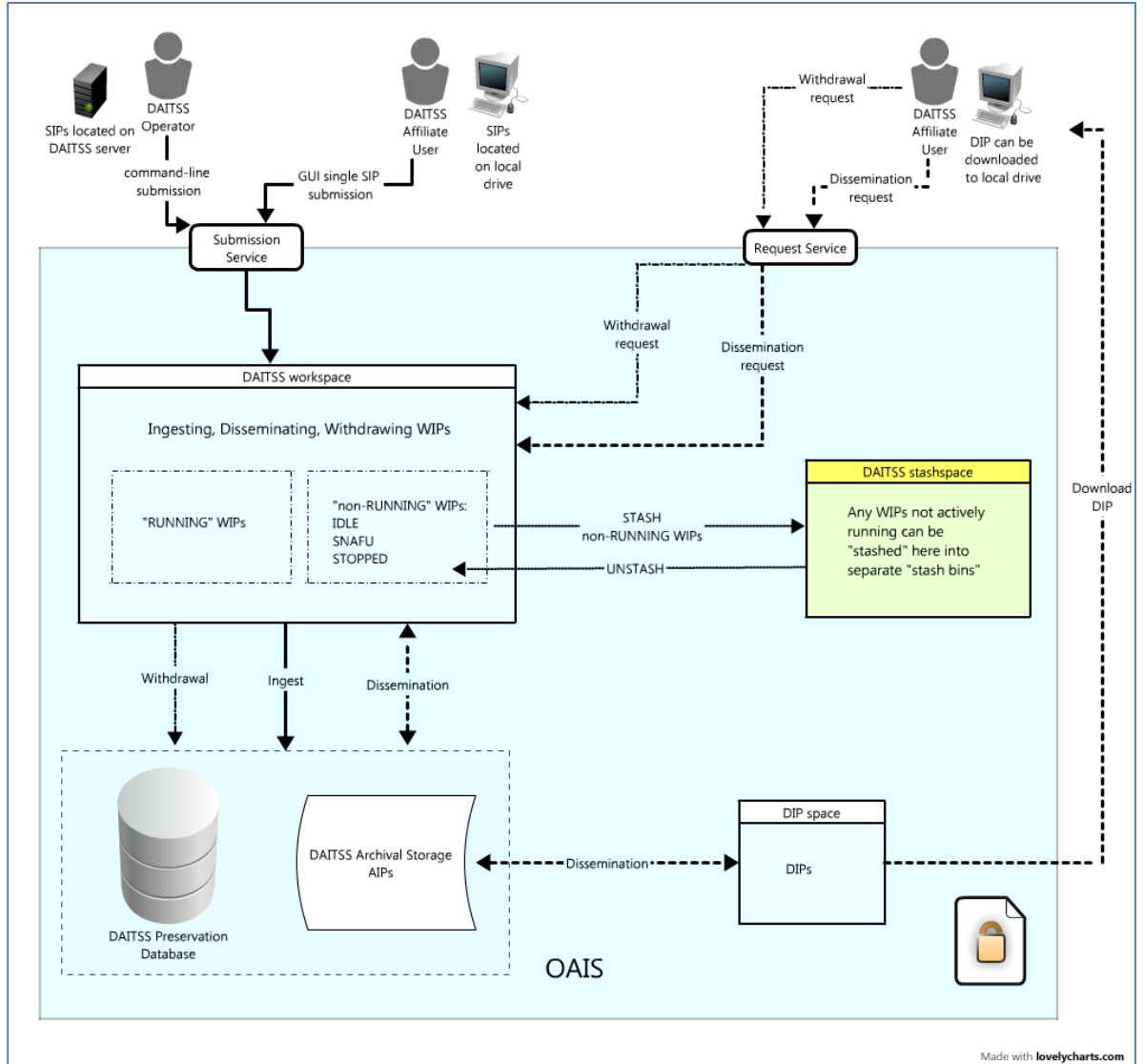
Preservation Task Steps: DAITSS preservation tasks include a number of separate processing steps that differ according to the task. During WIP processing, the Task Details page displays details of all task steps completed at the time of display. A detailed list of steps for each task is provided in this chapter.

Processing Status: The execution status of a specific job/WIP, as viewed from the “workspace” tab. This status is not the same as an Operations Event, which records both Operator actions, such as stopping running WIPs, and SYSTEM (DAITSS software) actions, such as “disseminate finished”.

Stashspace: A filesystem directory predefined to DAITSS where WIPs can be held temporarily in “stash bins” while problems are resolved. Multiple stash bins can be created via the Workflow Interface.

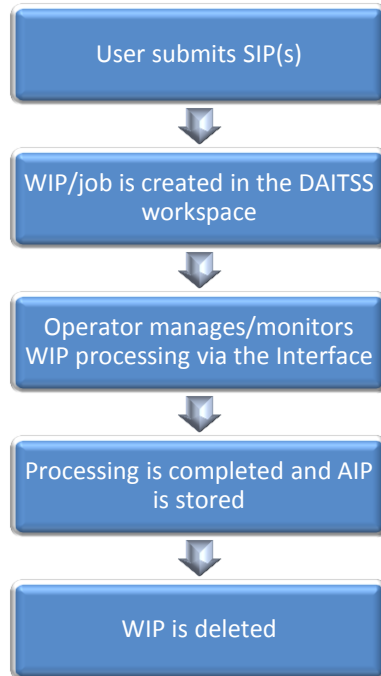
Stash Bin: A directory within the DAITSS Stashspace that contains WIPs set aside temporarily while problems are resolved. Stash Bins are created in the Workflow Interface’s “stashspace” tab.

DAITSS Work Flow Diagram



The diagram above illustrates Operator and User interaction with DAITSS and the Ingest, Dissemination, and Withdrawal workflows. Detailed recommendations for managing these workflows are provided later in the chapter.

A brief overview of DAITSS processing:



DAITSS Work Flow Interface Operator's Toolbar

The DAITSS Workflow Interface is accessed through a Web browser. Firefox and Safari have been tested. An operator must log on with a userid and password. After logging on, the DAITSS Operator will have access to the Operator's Toolbar, located at the top of the screen. The Operator's Toolbar displays the various tabs available to the Operator. The active tab will always be highlighted in blue.

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin Ollie_operator

FLORIDA
Digital Archive

daitss v0.17.9

[logout](#)

sip summary

count	2813
files	24858
total size	85819.00 MB

aip summary

count	2733
files	12598
total size	37553.00 MB

Detailed descriptions of all functions available in each tab are given at the end of this chapter.

DAITSS Processing Status vs. Operations Events

It is critical to understand the difference between package Processing Status and DAITSS Operations Events in order to successfully monitor DAITSS processing workflow.

Processing Status

Processing Status indicates the current state of a specific task. Processing Status is a visual cue to the Operator that a WIP/job is actively running or is in a non-running state. It does not indicate which archival task (ingest, dissemination, withdrawal) the WIP is performing, but rather indicates whether the task is being actively executed or not. Processing status is assigned by the Workflow Interface based on internal information. Processing Status is not recorded in the DAITSS Preservation Database as it is a transient state.

Processing Status options are:

IDLE: Idle status is assigned to WIPs that are in the DAITSS workspace, waiting for their preservation task to begin. Idle WIPs can either be started manually by an Operator or they can be started by the pulse background process.

RUNNING: Running status is assigned to WIPs that are in the workspace and are actively executing a task (ingest, withdrawal, dissemination).

STOPPED: Stopped status is assigned when an Operator has manually updated a RUNNING WIP's status to STOPPED, which essentially pauses processing of a WIP and retains it in the workspace until it is manually started.

ARCHIVED: Archived status is assigned when a package has been successfully archived. Archived is a special status that indicates that no processing is taking place on this AIP. No WIP exists for a package with this status.

ERROR: Error status is assigned when a WIP/job has encountered an unexpected error, e.g., one of the DAITSS web services is down and processing cannot continue. Once the problem is corrected the Operator can "reset" the WIP, which will set its Processing Status to IDLE. At that point the WIP can be started manually by an Operator or automatically by the Pulse background process.

STASH: Stash status is assigned when an Operator has “stashed” a WIP in the DAITSS workspace to a stash bin. This indicates to the Operator that no further processing of the WIP will take place until that WIP is “unstashed” from the stash bin and returned to the workspace.

Operations Events

Operations Events, in contrast to Processing Status, indicate an action that has been performed. Multiple Operations Events can be associated with a single WIP/job as it goes through processing.

The following is a list of DAITSS Operations Events which are recorded in the DAITSS database and displayed in the DAITSS User Interface. They are grouped logically by preservation task:

Submission-related events

submit: A submit event is recorded after a SIP has successfully passed the submission process and been transformed into a WIP (Workspace Information Package).

reject: A reject event is recorded when a SIP fails the submission process (is invalid). In a sense, reject indicates a failed submission attempt. Rejected packages are not transformed into WIPs and no additional processing is done on them. Rejected packages should be corrected and re-submitted.

Ingest-related events

ingest started: An ingest started event is recorded when either the pulse daemon or an Operator via the GUI “start” button has created an ingest process for a WIP.

ingest stopped: An ingest stopped event is recorded when an Operator clicks on the ‘stop’ button in the GUI. Only WIPs which are actively being ingested, i.e., have an ingest started event, can be stopped.

ingest finished: An ingest job has successfully completed and an AIP has been archived. Packages with an ingest finished event are successfully archived packages (AIPs).

ingest error: An error event and event details are recorded when the processing of a WIP fails. Reasons for errors are varied: a web service may be down or a hardware problem may be encountered, or the DAITSS programs may encounter an unexpected condition and fail. Most errors are temporary. WIP processing can be restarted when the problem is resolved.

ingest reset: A reset event is recorded when an Operator clicks on the ‘reset’ button for a single WIP or a group of WIPs. Procedurally, this is done to either retest a process to determine if a problem can be reproduced, or to resume processing of a WIP after a problem has been resolved.

ingest report mailed or **ingest report written out for ftp:** This type of event is recorded by the mailer-daemon process after an ingest report has been generated for an archived

package. The delivery method is determined by the contents of the 'report e-mail address' in the admin account record: a blank indicates that ingest reports for that account will be written to a directory on your DAITSS server, and a valid email address indicates that ingest reports will be emailed. This is the final step in the ingest process.

Dissemination-related events

disseminate requested: A disseminate requested event is recorded when an Operator clicks the disseminate 'Request' button for an AIP or the 'Disseminate' button for a batch of AIPs. This action also creates a 'request' record for the job. Request records are viewable under the 'requests' tab.

disseminate released: A disseminate released event is recorded when the Pulse daemon determines that a dissemination request can be released into the workspace as a dissemination WIP. The Pulse daemon actually creates a dissemination WIP during the process of releasing a dissemination request to the workspace. Processing of the dissemination WIP proceeds according to Pulse 'throttle' settings.

disseminate request cancelled: A disseminate request cancelled event is recorded when a dissemination request is cancelled. Dissemination requests can only be cancelled before a dissemination process is started.

disseminate started: A disseminate started event is recorded when either the Pulse daemon or an Operator via the GUI "start" button starts processing an IDLE dissemination WIP in the workspace.

disseminate finished: A disseminate finished event is recorded when the dissemination process has been completed and a DIP is available to the requestor.

disseminate error: A disseminate error event and event details are recorded when the processing of a disseminate WIP fails. Reasons for errors are varied: a web service may be down or a hardware problem may be encountered, or the DAITSS programs may encounter an unexpected condition and fail.

disseminate reset: A disseminate reset event is recorded when an Operator clicks on the 'reset' button for a single disseminate WIP or group of WIPs that have a status of ERROR. Procedurally, this is done to either restart a dissemination process to determine if a problem can be reproduced, or to resume processing of a WIP after a problem has been resolved.

disseminate do over: A disseminate do over event is recorded when an Operator clicks on the 'do over' button for a single disseminate WIP or group of WIPs that have a status of ERROR. The result is that disseminate processing of the WIP is completely redone. Procedurally, this is done to restart a dissemination process from the beginning, and should only be necessary in cases where a major DAITSS code update has been installed.

Withdrawal-related events:

Withdraw request placed: A withdraw request placed event is recorded when an Operator places a withdrawal request on an AIP. This causes the withdrawal request to be “enqueued” (waiting to be released to the workspace) pending authorization for release to the workspace by a second Operator.

Withdraw request cancelled: A withdraw request cancelled event is recorded when an Operator (this can be the same Operator who placed the request) clicks on the “Cancel” button of an “enqueued” withdrawal request, either in the package details’ page of an AIP, or in the “requests” tab. Note that once a withdrawal request is ‘released_to_workspace’ it can no longer be cancelled.

Withdraw request authorized: A withdraw request authorized event is recorded when an Operator (other than the Operator who placed the request) clicks on the “Authorize” button of a withdrawal request either in the package details’ page of an AIP, or in the “requests” tab.

Withdraw released: A withdraw released event is recorded when the requests service determines that no other requests (such as Dissemination requests) are pending on an AIP. This causes the withdrawal request to be “released_to_workspace”, which means that a WIP (a job) is created in the workspace for the task of withdrawing the AIP from the repository. Once in the workspace, the WIP/job is managed by the pulse program until the withdrawal process is completed.

Withdraw started: A withdraw started event is recorded when the pulse program or an Operator starts processing a withdrawal WIP/job from the DAITSS workspace.

Withdraw finished: A withdraw finished event is recorded when the withdrawal task is successfully completed.

Other WIP management events:

stash: A stash event is recorded when an Operator clicks on the 'stash' button of a WIP or group of WIPs in the DAITSS workspace and specifies a stash bin to which the WIP(s) should be moved. Stashing WIPs moves them from the workspace to a stash bin (directory outside of the workspace) where they are effectively on hold and will not be processed further until they are unstash and returned to the workspace.

unstash: An unstash event is recorded when an Operator clicks on the 'unstash' button of a WIP or a group of WIPs in a DAITSS stash bin. Unstashing moves WIPs from the stashspace to the workspace where they are eligible for ingesting.

Managing Processing with the DAITSS Preservation Interface

The DAITSS Workflow Interface can be used to manage your processing workflow in a number of ways. What follows is the workflow used by Florida Digital Archive Operators to monitor and manage daily operations.

Submitting SIPs

DAITSS has two mechanisms for submitting SIPs to the DAITSS Submission service: the DAITSS Workflow Interface “submit” function found under the “packages” tab, and a command-line batch submission command that can be used to submit batches of SIPs to the DAITSS Submission Service.

Note that the DAITSS Workflow Interface “submit” function requires that SIPs be in a tar or zip format but command-line batch submission does not.

Submission of single SIPs via the Interface

1. Log into the DAITSS Workflow Interface
2. Click on the “packages” tab
3. Click on the “browse” button to find SIPs accessible from your PC.
4. Browse your PC to find a tarred or zipped SIP directory, click on that tar or zip file, and click on “open”
5. Optionally: add a batch to which this package will be added upon submission.
6. Optionally: add a submission note that will be recorded in the DAITSS preservation database with the submission event.
7. Click on the “Submit” button. Depending on the size of the SIP this process may take up to a minute or so as the SIP is transferred to DAITSS and passes through the Submission Service.

dashboard : [packages](#) : errors : workspace : stashspace

submit

packages must be in *tar* or *zip* format

File	<input type="text" value="C:\Users\Lydia\Desktop\temp\FDAtest4.zip"/> <input data-bbox="1047 556 1193 598" type="button" value="Browse..."/>
Batch	<input type="text" value="test_batch"/>
Note	<input type="text" value="This is a test SIP for the new ETD project."/>
	<input type="button" value="Submit"/>

After the submitted SIP passes through the DAITSS Submission Service you will see the results of the submission. The SIP will either be successfully submitted (event type 'submit'), or an invalid SIP will be rejected (event type 'reject', with a more detailed reject message.)

An example of a successful submission attempt:

dashboard : [packages](#) : errors : workspace : stashspace : [batches](#) : requests : log : admin

ETAQOIF30_4KKBZT : ingesting (idle)

sip name	account	project
FDAtest4	FDA	FDA

batches

batch name
test_batch

events

type	time	agent	comments	notes
submit	Mon Sep 26 2011 04:54:54 PM	Ollie_operator	0 comment(s)	This is a test SIP for the new ETD project.

An example of an unsuccessful submission attempt (a rejected package):

The screenshot shows a web interface with a navigation bar at the top containing links: dashboard, packages, errors, workspace, stashspace, batches, requests, log, and admin. Below the navigation bar, the main content area displays the package ID **E2DWRSKU9_2DFAC1 : rejected**. Underneath, there is a table with columns: sip name, account, and project. The values are unnamed, SYSTEM, and default respectively. Below this is a section titled **events** with a table containing columns: type, time, agent, comments, and notes. A single event is listed with type 'reject', time 'Tue Sep 13 2011 02:03:24 PM', agent 'root', and notes: 'cannot extract sip archive, must be a valid tar or zip file containing directory with sip files'.

The above example illustrates a case where a submission attempt has been made for a file other than a tar or zip file. The reject notes indicate “must be a valid tar or zip file containing a directory with sip files”.

Note that both in cases of successful submission and rejected submission attempts a DAITSS Intellectual Entity ID (IEID) is issued and the submission attempt is recorded in the DAITSS preservation database. In addition, DAITSS generates a Reject Report upon rejecting a submission attempt. Successful submission attempts will result in Ingest Reports being generated once ingest is finished and the package is archived.

Submitting SIPs with the command-line submission program

Single SIPs and batches of SIPs can also be submitted from the DAITSS server by using the command-line submission. See Chapter 3 for details.

Retrieving and Viewing DAITSS Reports

DAITSS generates reports of all successful Ingests and rejected packages, providing details about each.

To retrieve a DAITSS Ingest report:

- Identify a package whose report you want to view
- Go to the “aips” section of the Package Details page, where you’ll find links to the report and to the report stylesheet

aip

title	12003_1979_Index
aip descriptor	mets descriptor
descriptor errata	
copy url	E28BLWNRR_1ILGVU.tar
report	ingest report
report stylesheet	report stylesheet
aip size	48527360
copy sha1	a5a86e3b2ac5b4b9c62e03ef7be90b6f8b42cc20
copy md5	74fac123646e31e3344a6a018792a874
number of datafiles	8
number of submitted datafiles	8

- Right click on each link, and select the "Save as" option, and save both files to the same directory.
- You'll then be able to view the report with a browser.

Chapter 6: DAITSS Workflow Interface

Example of an Ingest Report

Ingest

Package name	Int. Entity ID	Ingest time
FDAE001863	ECXFF0BY7_794ICS	2011-09-16T11:51:06-04:00

Agreement Info

Account:	FDA
Project:	ETD

Files

Archival Attributes

Id	Name	Size	Origin	Message Digests	Events	Broken Links	Warnings
daitss-test://ECXFF0BY7_794ICS/file/0	sip-files/FDAE001863.xml	10850	DEPOSITOR	<u>2</u>	<u>3</u>	0	0
daitss-test://ECXFF0BY7_794ICS/file/1	sip-files/Arevalo_Jorge_M_200712_PhD.pdf	3296344	DEPOSITOR	<u>2</u>	<u>2</u>	0	<u>1</u>

Message Digests

daitss-test://ECXFF0BY7_794ICS/file/0

Message Digest	Algorithm
f4bf7e902accf4aefa256af3aa8b90a1	MD5

Retrieving a DAITSS Reject Report

DAITSS Reject Reports contain the information displayed on a rejected package's Package Details page in a downloadable XML file. To download a DAITSS Reject Report:

- Identify a package whose report you want to view
- Click on the "reject report" link from the Package Details page:

EZBNFGZUD_7JLQY7 : rejected

sip name	account	project
FDA6151501	FDA	DZA

batches

batch name
disseminate_for_Smith

events

type	time	agent	comments	
reject	Fri Sep 02 2011 10:30:44 AM	childree	0 comment(s)	sip path: /var/daitss/ missing content file:

legacy events

type	time	agent	comments	notes
show fixity events				

reject report

[reject report](#)

Example of a Reject Report

Error

sip path: /var/daitss/incoming/benchmarks/DZA/FDA6151501, submission note: Testing DZA.; missing content file: FDA6151501.wav missing content file: FDA6151501.doc
 FDA6151501 (EZBNFGZUD_7JLQY7) rejected 2011-10-04T14:15:08-04:00

Monitoring the DAITSS Workspace

If a number of SIPs or requests have been submitted to the workspace, the DAITSS Operator will want to monitor the progress of their processing by monitoring the “workspace” tab. The workspace tab (page?) will provide the following information:

- **How much is waiting to be processed:** the type, number, and size of IDLE jobs/WIPs in the workspace awaiting processing are displayed in the “idle jobs” section of the display.
- **How many simultaneous tasks can be processed:** The “throttle” settings for the number of simultaneous ingests, disseminations, and withdrawals are displayed. These throttles allow each type of task to be run concurrently so that no task category has a higher priority than another and each task category effectively has its own “queue”. Note that setting a throttle limit for a category of task to zero effectively prevents any tasks of that category from being executed.
- **How much is currently being processed:** the number and total size of ACTIVE jobs/WIPs that have begun but not yet completed their processing tasks is displayed.
- **A list of WIPs being processed:** A list of ACTIVE jobs/WIPs, along with their package names, task, status of the job and timestamp of the status.
- **Details of specific WIPs being processed:** Clicking on the IEID of a specific active job will display details of the processing steps completed and any ERROR messages and stack trace.

Monitoring a RUNNING WIP/job

An Operator can track the processing progress of a specific WIP as follows:

- From the “workspace” tab, click on the IEID of a running WIP.

dashboard : packages : errors : **workspace** : stashspace : batches : requests : log : admin

batch	account	project	start date	end date	status	set scope
<input type="text"/>	<input type="text"/>	<input type="text"/>	09/24/2011	09/29/2011	<input type="text"/>	<input type="button" value="Set Scope"/>

update jobs

start
 stop
 reset
 do over
 stash

idle jobs: 0

size: 0.00 B

task	count
ingest	0
d1refresh	0
disseminate	0
withdraw	0

active jobs: 1

running: 1 throttles: 2 ingests; 2 disseminations; 2 withdrawals; 2 d1 refreshes;

size: 1.20 GB

intellectual entity id (ieid)	package	task	status	time
EA3U8PIXM_UI5JBE	FDA00028308_00127	ingest	running	Wed Sep 28 12:18:06 -0400 2011

- Clicking on the IEID of a running WIP displays the Task Details page that lists the progress of a specific DAITSS task (ingest, dissemination, withdrawal) by displaying the name of each task step (including, for steps performed on each content file, the number of the file being processed), the time that the step was started, and the number of seconds required for each step to complete. Note that this display records task steps *only after they are completed*. Should WIP processing be stopped, it will resume processing *on the next consecutive step*. A list of processing steps for each DAITSS task is included in this chapter.

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin

ingest of EA3U8PIXM_UI5JBE (running)

start
 stop
 reset
 do over
 stash
 idle

progress

step	time	seconds
virus check 48	2011-09-28T12:13:45-04:00	001.94
virus check 83	2011-09-28T12:13:47-04:00	004.73
virus check 36	2011-09-28T12:13:52-04:00	000.56
virus check 30	2011-09-28T12:13:53-04:00	000.48
virus check 84	2011-09-28T12:13:53-04:00	004.57
virus check 55	2011-09-28T12:13:58-04:00	004.20
virus check 37	2011-09-28T12:14:02-04:00	000.45
virus check 86	2011-09-28T12:14:02-04:00	004.88
virus check 88	2011-09-28T12:14:07-04:00	004.55
virus check 124	2011-09-28T12:14:12-04:00	000.02
virus check 19	2011-09-28T12:14:12-04:00	000.07
virus check 102	2011-09-28T12:14:12-04:00	000.03
virus check 14	2011-09-28T12:14:12-04:00	000.06
virus check 68	2011-09-28T12:14:12-04:00	000.05

Monitoring any outstanding errors

All WIPs with unresolved errors (i.e., WIPs that have not completed processing because of an error condition) can be found by clicking on the "errors" tab in the Operator's toolbar. The resulting list will include an excerpt from the error message.

dashboard : packages : **errors** : workspace : stashspace : batches : requests : log : admin Ollie_oper

errors

filter scope

batch	account	project	start date	end date	status	error message	set scope
	FDA			09/02/2011			<input type="button" value="Set Scope"/>

results

2 packages 1.36 GB 10 files

intellectual entity id (ieid)	package	account	project	size (MB)	# of datafiles	status	error date	error message
EHKFVHPGH_6M7NQV	FDA6151101	FDA	DZA	1338.92	3	stashed	Fri Sep 02 2011 12:08:50 PM	private method 'gsub' called for nil:NilClass
ERWN6Z9E4_JUP51Q	UF00100139_00015	FDA	default	23.06	7	stashed	Tue Jul 19 2011 11:31:33 AM	undefined method 'xml' for nil:NilClass

Save as Batch

batch name

- Click on the package IEID to see detailed information about processing, and to execute Operator Actions such as “reset”, then “start” to resume package processing.
- Once the WIP has been fully processed the package will be removed from this display

Searching for Packages

To retrieve detailed information about a specific package, use the “search” section of the “packages” tab. Packages can be retrieved either by package name (the directory name of the submitted SIP) or by the IEID assigned by DAITSS upon submission. Multiple IEIDs or package names (or a mixture of both) can be submitted for retrieval in the same search operation.

Click on the “Search” button to begin the search.

Note that the resulting list of packages can be downloaded as a CSV file for importing into a spreadsheet or saved as a batch.

Managing the jobs throttle

The DAITSS jobs or tasks throttle settings allow for separate processing queues for each type of DAITSS task: ingest, dissemination, and withdrawal. The throttle settings specify to the Pulse daemon how many simultaneous processes will be allowed for each task so that no specific type of task has priority over another. In this way the repository can simultaneously process ingests, disseminations, and withdrawals, all of which may be equally important to repository workflow, without requiring an Operator to set the processing priority of individual WIPs.

For example, while the bulk of a repository’s processing may consist of ingesting/archiving new materials, periodic dissemination or withdrawal requests can be promptly processed without requiring additional Operator intervention.

DAITSS is configured with throttle defaults of "1" for each task category. These settings are displayed in the "workspace" tab in the "active jobs" section.

Using Stash Bins to set aside Packages

Operators can set aside WIPs from the workspace into stash bins in the DAITSS stashspace. Only WIPs that are in the DAITSS workspace and have a workspace status other than RUNNING (e.g., IDLE, ERROR, STOPPED) can be stashed to a stash bin. Any stashed WIPs can be unstashed from a stash bin, which will cause them to be returned to the DAITSS workspace with their processing status unchanged.

Stashing WIPs is a two-step process:

- From the "stashspace" tab, create a new stash bin. (Note that existing stash bins can be used by skipping to the next step.)
- From the "workspace" tab, or the Package Details Page, select the desired stash bin from the pull-down list and click on the "stash" radio button.
- Optionally: add a note that will be stored and displayed with the "stash" event.
- Click on the "Update" button to complete the stash action.

Stopping and Starting WIP processing

DAITSS Operators can manually override the start of package processing either by starting or stopping a group of WIPs from the "workspace" tab, or individual WIPs can be started or stopped from the Task Details Page. Operators may occasionally need to stop/start WIP processing for scheduled system maintenance, particularly if very large WIPs are being processed that will take some time to complete.

Stopping running WIPs from the "workspace" tab

1. Click on the "workspace" tab.
2. Click on the "stop" radio button, optionally add a note explaining the action, and click on the "Update" button to complete the action. (Note that only running processes will be stopped.)

dashboard : packages : errors : **workspace** : stashspace : batches : requests : log : admin

batch	account	project	start date	end date	status	set scope
<input type="text"/>	<input type="text"/>	<input type="text"/>	09/24/2011	09/29/2011	<input type="text"/>	<input type="button" value="Set Scope"/>

update jobs

start
 stop
 reset
 do over
 stash

idle jobs: 0

size: 0.00 B

task	count
ingest	0
d1refresh	0
disseminate	0
withdraw	0

active jobs: 1

running: 1 throttles: 2 ingests; 2 disseminations; 2 withdrawals; 2 d1 refreshes;

size: 1.08 GB

intellectual entity id (ieid)	package	task	status	time
EBN7MQFW1_MVGJLB	FDA00028308_00128	ingest	running	Wed Sep 28 11:55:40 -0400 2011

Stopping running WIPs from the Task Details Page

1. Click on the IEID of a running WIP to access the Task Details Page.
3. Click on the "stop" radio button, optionally add a note explaining the action, and click on the "Update" button to complete the action.

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin

ingest of EBN7MQFW1_MVGJLB (running)

start
 stop
 reset
 do over
 stash
 idle

Stopping to allow system maintenance.

progress

step	time	seconds
virus check 48	2011-09-28T11:55:22-04:00	004.36
virus check 83	2011-09-28T11:55:26-04:00	013.59
virus check 36	2011-09-28T11:55:40-04:00	000.46
virus check 30	2011-09-28T11:55:40-04:00	000.43
virus check 84	2011-09-28T11:55:41-04:00	004.78

The same procedure can be followed to start stopped WIPs. WIP processing will resume *on the next consecutive step* for that task and *no completed steps will be repeated*. This enables Operators to easily stop and restart running WIPs as necessary without adding to the total processing time. (Note that starting packages manually will override the task throttle settings.)

Changing Processing Order

The default DAITSS processing order is FIFO (First In First Out) order, so that packages are processed in the order submitted. Under some circumstances, such as rush projects, an Operator may want to stop the pulse program, submit the rush project packages, and restart the pulse program so that it processes the new packages first (LIFO order, Last In First Out). Once the new packages are processed pulse can be restarted in FIFO order.

1. Stop the pulse program that maintains the DAITSS processing queue (link to that process)
2. If the processing order change is temporary, override the default processing order set in the configuration file (daitss-config.yml), by using either the `-fifo` or `-lifo` flag when restarting pulse.

Using Batches to monitor processing

If your repository's policy or procedures dictate that packages should be deposited and processed in batches, you should use the DAITSS batch function from start to finish, as follows:

1. Every SIP submission should include the option to include the SIP(s) in a batch specific to the archiving project. If a user negotiates a submission agreement that specifies the deposit of a certain number of SIPs on a certain schedule, we recommend that each

group of SIPs deposited on that schedule be included in a separate DAITSS batch and named appropriately. For example, "GSN_Sept_2011" might be the name of the batch of SIPs submitted for the "GSN" project during September 2011. An operator could then easily verify that submission of the 200 SIPs scheduled during September 2011 has been completed.

2. By periodically viewing the "activity summary" of a batch from the "batches" tab, an Operator can determine how many packages in the batch have been rejected from submission, successfully archived, or have processing errors.
3. Once the packages in an entire batch have been successfully archived, the Operator can choose to either delete the batch, or the batch can be retained for tracking purposes. Batches are simply lists of package IEIDs, so they do not use much space in the preservation database.

Disseminating Packages

Dissemination Information Packages (DIPs) are currently created by DAITSS as a result of Operator-placed dissemination requests. Future DAITSS development will include the ability for authorized Affiliate Users to request dissemination of packages from within their own accounts. The DAITSS Dissemination Information Package (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.

Dissemination WIPs (that is, WIPs created in the process of reprocessing the AIP to create any new normalized or migrated versions of original AIP files) can be managed from the DAITSS workspace. [Note that while a disseminate WIP is in process the original AIP is still stored in the DAITSS repository. After the reprocessed AIP is successfully stored to the repository the original AIP is deleted and a new AIP with the same IEID is stored in the repository. The new AIP will contain both the original content files submitted in the SIP and the 'latest, best' version of content files. The DIP will also contain both the original content files and the 'latest, best' version of content files.]

Dissemination Steps (single AIP)

- Identify the package you want to disseminate from the "packages" tab.
- Click on the IEID of the package.
- Add a note that will be recorded in the dissemination request (optional) and click on the 'Disseminate' button. (Clicking on the 'disseminate' button creates a DAITSS request record for the dissemination job.)

ingest stats

name	value
elapsed time	20.0 sec

requests

submit request

disseminate ▾ Requested by P. Tam | Request

aip

title	The (fd)A Team
aip descriptor	mets descriptor
descriptor errata	

- If the AIP is available for dissemination and the pulse process is running, the dissemination request job will cause a dissemination WIP to be created in the workspace, and the dissemination process will update the WIP in the workspace.
- Once the dissemination is completed and an operations event of 'dissemination finished' is recorded, a DIP will be available for download from a link at the bottom of the package details page.
- Retrieve the DIP by clicking on the link in the "dips" section of the package detail page (at the bottom of the screen):

requests

id	request type	request submitter	note	authorized	
1048	disseminate	Ollie_operator	Requested by P, Tam	yes	released

submit request

disseminate

aip

title	The (fd)A Team
aip descriptor	mets descriptor
descriptor errata	
copy url	EZI6PG6S8_PA71MZ.tar
report	ingest report
report stylesheet	report stylesheet
aip size	1966080
copy sha1	72eb5ab6a05212131d43ada466706a5a2ea50e4f
copy md5	0a49c3350e7008a854455bac3de8d303
number of datafiles	2
number of submitted datafiles	2

dips

[EZI6PG6S8_PA71MZ-0.tar](#)

Dissemination Steps (a batch of AIPs)

1. Identify the IEIDs of set of packages you want to disseminate.
2. From the "batches" tab, use the "create a new batch" function to create a batch that includes those IEIDs.

- From the “batches” tab list of existing batches, add a note that will be recorded in the dissemination request (optional) and click the “Disseminate” button next to your new batch.

dashboard : packages : errors : workspace : stashspace : **batches** : requests : log : admin

batches

create new batch

batch name insert whitespace separated list of IEIDs here

existing batches

batch name	# of packages	disseminate	withdraw	peek	delete
ateams	7	<input type="text"/> <input type="button" value="Disseminate"/>	<input type="text"/> <input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
BENCHMARK_20110916	1	<input type="text"/> <input type="button" value="Disseminate"/>	<input type="text"/> <input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
CF_fixed	2	<input type="text"/> <input type="button" value="Disseminate"/>	<input type="text"/> <input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
disseminate_for_Smith	77	Requested by J. Smith <input type="button" value="Disseminate"/>	<input type="text"/> <input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>

- If an AIP in the batch is available for dissemination and the pulse process is running, the dissemination request job will cause a dissemination WIP to be created in the workspace, and the dissemination process will update the WIP in the workspace.
- Once the dissemination is completed and an operations event of 'dissemination finished' is recorded, a DIP will be available for download from a link at the bottom of the package details page.
- Retrieve the DIPs by copying them from the DAITSS dissemination directory/DIP space (/var/daitss/data/disseminate) to a desired target directory or server. Alternately, single DIPs can be downloaded from the Interface by clicking on the package IEID, and downloading from the “dips” section of the package details page.

Withdrawing AIPs

DAITSS does not allow direct deletion of AIPs from archival storage once a DAITSS repository has taken responsibility for long-term preservation of a package. To remove an AIP from the repository, the DAITSS withdraw function must be used. As an additional safeguard against unintentional removal of AIPs from the repository, DAITSS requires authorization of a withdrawal request by a second Operator to ensure that AIPs are never removed from the repository in error. Currently only DAITSS Operators can initiate withdrawal requests, but a

Chapter 6: DAITSS Workflow Interface

future version of DAITSS will include the ability for Affiliate Users with “withdraw” permission to initiate their own withdrawal requests.

The withdraw function includes the following steps/Operations Events that document each step in the process of requesting and executing a withdrawal. Identify the package or packages to be withdrawn from the “packages” tab.

1. Click on the IEID link of the package to access the Package Details Page.
2. In the “submit request” section of the page, select “withdraw” from the pull-down menu and enter a note that documents owner authorization of withdrawal of the AIP. (For example, at the Florida Digital Archive we require that withdrawal requests be made via email by authorized individuals, and the original email is permanently retained in a “ticketed” support tracking system.)
3. Click on the “Request” button to initiate the withdrawal request.

submit request

withdraw ▼ Requested by Jim Smith Request

aip

title	The Ocala evening star.
aip descriptor	mets descriptor
descriptor errata	

A second Operator, possibly a manager, must now authorize the withdrawal request:

- Click on either the “requests” tab to view all requests.
- Note that withdrawal requests requiring authorization display an “Authorize” button and “false” in the “authorized?” column. While withdrawal requests can be authorized from this screen, a more conservative approach is described in Step 3 below.

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin lydia

requests

filter request scope

batch	account	project	user	request type	request status	set scope
all ▼	all ▼	all ▼	all ▼	all ▼	all ▼	Set Scope

request list (top 100 displayed)

display all display 100

id	intellectual entity id (ieid)	package	project	request type	requesting user	request status	authorized?	request timestamp	authorize	cancel
1050	E9T2JTFM4_G9M2OY	FDA00075908_00714	FDA	withdraw	Ollie_operator	enqueued	false	Wed Sep 28 2011 11:27:48 AM	Authorize	Cancel

Chapter 6: DAITSS Workflow Interface

- A more conservative approach: click on the IEID of the withdrawal request requiring authorization and review the details of the request note prior to clicking on the "Authorize" button. (Note that a "Cancel" button is also available, for use in cases where the withdrawal request is inappropriate or has been made in error.)

requests								
id	request type	request submitter	note	authorized	status	timestamp	authorize	cancel
1050	withdraw	Ollie_operator	Requested by Jim Smith	no	enqueued	Wed Sep 28 2011 11:27:48 AM	<input type="button" value="Authorize"/>	<input type="button" value="Cancel"/>

The Package Details Page

Details about a specific package can be accessed from within any listing of packages in the Workflow Interface by clicking on the Intellectual Entity ID (IEID) of the package. (The IEID is a 16 character ID consisting of 9 alphanumeric characters, an underscore, and 6 additional alphanumeric characters, e.g., E7TOIYCGD_YALOZZ.)

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin

E23V6EVLX_8H1CK4 : archived

sip name	account	project
FDA00028419_02639	FDA	FDA

batches

batch name
disseminate_for_Smith
TESTING_AD_HOC

events

type	time	agent	comments	notes
submit	Fri Sep 02 2011 10:27:21 AM	childree	0 comment(s)	sip path: /var/daitss/incoming/benchmarks/AD_HOC/FDA00028419_02639, submission note: Testing AD_HOC.
ingest started	Fri Sep 02 2011 02:02:20 PM	SYSTEM	0 comment(s)	
ingest finished	Fri Sep 02 2011 02:13:10 PM	SYSTEM	0 comment(s)	

legacy events

type	time	agent	comments	notes
------	------	-------	----------	-------

[show fixity events](#)

ingest stats

name	value
elapsed time	650.0 sec

requests

submit request

aip

title	The star.
aip descriptor	mets descriptor
descriptor errata	
copy url	E23V6EVLX_8H1CK4.tar
report	ingest report
report stylesheet	report stylesheet
aip size	489953280
copy sha1	705fd9912d609c131ca694c7d665b108cc8a8ff0
copy md5	fae2df317c9b0b590db8dd59916b4d58
number of datafiles	100
number of submitted datafiles	100

The Sections of the Package Details Page:

Package header

The Package Header section of this page provides the following information:

Chapter 6: DAITSS Workflow Interface

- The Intellectual Entity ID (IEID) of the package followed by the Processing Status of the package. (Detailed descriptions of all Processing Statuses are included elsewhere in this chapter.)
- The SIP Name (i.e., the name of the SIP directory)
- The Account and Project specified in the SIP descriptor.

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin

ECWRHR3YT_O485L4 : ingesting (running)

sip name	account	project
FDAtest4	FDA	FDA

- In addition, any packages with a Processing Status of RUNNING, STOPPED, or ERROR (in other words, WIPs that have not completed a processing task) will display the Task Details page if the Processing Status link is clicked (e.g, running):

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin

ingest of ECWRHR3YT_O485L4 (running)

start stop reset do over stash

progress

step	time	seconds
virus check 36	2011-09-30T11:26:12-04:00	000.13
virus check 30	2011-09-30T11:26:12-04:00	000.09
virus check 37	2011-09-30T11:26:13-04:00	000.13
virus check 19	2011-09-30T11:26:13-04:00	000.12
virus check 14	2011-09-30T11:26:13-04:00	000.03

- Clicking on the IEID link returns you to the Package Details page.

Batches section

If the IEID of the package being displayed is contained within one or more DAITSS batches the names of all batches of which this package is a member are displayed here:



Events section

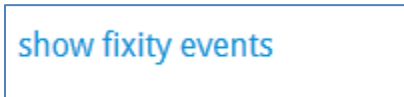
This section lists all Operations Events recorded for this package, including the timestamp of the event, the user/agent that initiated the event, links to any additional comments associated with the event, as well as the event note:

The image shows a screenshot of a web interface. At the top, the word "events" is displayed in a blue font. Below it is a table with the following data:

type	time	agent	comments	notes
submit	Fri Sep 30 2011 11:26:09 AM	Ollie_operator	0 comment(s)	
ingest started	Fri Sep 30 2011 11:26:12 AM	SYSTEM	0 comment(s)	
ingest finished	Fri Sep 30 2011 11:27:02 AM	SYSTEM	0 comment(s)	
disseminate request placed	Fri Sep 30 2011 11:39:39 AM	Ollie_operator	0 comment(s)	request id: 1051; Requested by Jim Smith
disseminate released	Fri Sep 30 2011 11:39:44 AM	SYSTEM	0 comment(s)	request_id: 1051
disseminate started	Fri Sep 30 2011 11:39:44 AM	SYSTEM	0 comment(s)	
disseminate finished	Fri Sep 30 2011 11:42:03 AM	SYSTEM	0 comment(s)	

"show fixity events" link

This link will display the most recent successful fixity check on an AIP, and all fixity or integrity failures recorded for that AIP. By default these are not displayed until the link is clicked.



"hide fixity events" link

To hide the display of fixity events, click on this link.

Chapter 6: DAITSS Workflow Interface

The AIP section includes key information about the AIP:

- the title of the package (if provided in the SIP descriptor),
- a link to the AIP descriptor,
- a link that will recreate the ingest report,
- a link to the actual stored AIP, where appropriate. (Note: this function should only be used by DAITSS Operators for troubleshooting purposes; the dissemination function should be used to produce DIPs for AIP owners.)
- The size in bytes of the stored AIP
- The sha1 and md5 checksums of the stored AIP
- The number of datafiles contained in the AIP
- The number of datafiles submitted in the SIP (Note: datafiles included in a SIP but not described in the SIP descriptor are deleted by DAITSS during submission.)

aip	
title	12003_1979_Index
aip descriptor	mets descriptor
descriptor errata	
copy url	E28BLWNRR_1ILGVU.tar
report	ingest report
report stylesheet	report stylesheet
aip size	48527360
copy sha1	a5a86e3b2ac5b4b9c62e03ef7be90b6f8b42cc20
copy md5	74fac123646e31e3344a6a018792a874
number of datafiles	8
number of submitted datafiles	8

DIPs section

The DIPs section of the Package Details page will provide a link to a DIP, if a dissemination task has been completed. The DIP file is a tarfile with the naming convention <Package IEID>-<LATEST VERSION>.tar. Clicking on the DIP link will download the DIP to your local workstation:

dips

E20080313_AAAAAK-0.tar

DAITSS Preservation Task Processing Steps

Preservation Task Processing Steps are displayed from the Task Details page of an Active WIP (a WIP that has not completed processing). Each Preservation Task has its own sequence of steps, and an Operator can monitor WIP processing by viewing completed steps. Note that Task Steps are displayed *after completion of a step* and not during *execution of the step*. The step currently being executed is not displayed.

Task: Ingest

Steps

- **Virus check:** Perform virus checking on each content file
- **Describe-migrate-normalize:** Perform preservation functions on each content file. The preservation functions include format identification, validation and characterization, and format normalization and/or format migration according to format Action Plans.
- **Xml-resolution:** Resolve the schemas used in xml files and create a tarfile of all schema files used in the package.
- **Ingest digipro:** Write digital preservation provenance for this ingest, i.e. PREMIS events and agents describing the ingest process
- **Make aip descriptor:** Create the METS AIP descriptor which contains all preservation metadata and digital provenance for the entire package.
- **Validate aip descriptor:** Ensure that the created AIP descriptor conforms to the Guidelines for using PREMIS in METS (<http://www.loc.gov/standards/premis/guidelines-premismets.pdf>)
- **Make tarball:** Create an AIP tarball (tarfile) containing all depositor content files, normalized/migrated files, xml schema used in the package, and the AIP descriptor.
- **Make aip:** Parse the AIP descriptor and create database records for the preservation database. Write the tarball to storage.

An example of ingest steps as displayed in a package containing two content files:

step	time	seconds
virus check 0	2011-09-02T11:05:46-04:00	003.92
virus check 1	2011-09-02T11:05:50-04:00	004.03
describe-migrate-normalize-0	2011-09-02T11:05:54-04:00	001.97
describe-migrate-normalize-1	2011-09-02T11:05:56-04:00	000.11
xml-resolution	2011-09-02T11:05:56-04:00	000.87
ingest digiprov	2011-09-02T11:05:57-04:00	000.00
make aip descriptor	2011-09-02T11:05:57-04:00	000.05
validate aip descriptor	2011-09-02T11:05:57-04:00	000.73
make tarball	2011-09-02T11:05:58-04:00	000.07
make aip	2011-09-02T11:05:58-04:00	000.55
total duration	012.31	012.30

Task: Disseminate

Steps

- **Load aip datafiles:** Download the AIP tarball from storage into the disseminate WIP
- **Load aip metadata:** Parse and load AIP metadata which includes package descriptive metadata, prior digital provenance records for each file and prior package provenance records.
- **Describe-migrate-normalize:** Perform preservation functions on each content file. The preservation functions include format identification, validation and characterization, and format normalization and/or format migration according to current format Action Plans.
- **Xml-resolution:** Resolve the schemas used in xml files and create a tarfile of all schema files used in the package.
- **disseminate digiprov:** Write digital preservation provenance for this dissemination according to PREMIS, i.e. premis events and agents describing the dissemination process.

- **Make aip descriptor:** Create the PREMIS-IN-METS AIP descriptor which contains all preservation metadata and digital provenance for the entire package.
- **Validate aip descriptor:** Ensure that the created AIP descriptor conforms to Guidelines for using PREMIS in METS (<http://www.loc.gov/standards/premis/guidelines-premismets.pdf>)
- **Make tarball:** Create a new AIP tarball (tarfile) containing all of the original AIP's depositor content files, any new normalized/migrated files, xml schema used in the package, and the AIP descriptor.
- **Make aip:** Parse the AIP descriptor and create database records for the preservation database. Write the new AIP tarball to storage.
- **Delete old copy:** Remove the obsolete copy of the AIP from storage.
- **Deliver DIP:** download a copy of the re-ingested AIP (the "last, best" copy) to the DAITSS DIP space (/var/daitss/data/disseminate).

An example of disseminate steps:

step	time	seconds
load aip datafiles	2011-10-04T10:44:12-04:00	000.49
load aip metadata	2011-10-04T10:44:12-04:00	000.09
describe-migrate-normalize-F20101214_AAAAAB	2011-10-04T10:44:12-04:00	034.68
describe-migrate-normalize-F20101214_AAAAAA	2011-10-04T10:44:47-04:00	004.77
xml-resolution	2011-10-04T10:44:52-04:00	001.51
disseminate digipro	2011-10-04T10:44:53-04:00	000.01
make aip descriptor	2011-10-04T10:44:53-04:00	000.24
validate aip descriptor	2011-10-04T10:44:54-04:00	001.85
make tarball	2011-10-04T10:44:55-04:00	000.10
make aip	2011-10-04T10:44:56-04:00	023.91
delete old copy	2011-10-04T10:45:19-04:00	001.45
deliver dip	2011-10-04T10:45:21-04:00	000.10
total duration	069.20	069.19

Task: Withdraw

Steps

- **Load AIP metadata:** Parse and load AIP metadata, which includes the package descriptive metadata, prior digital provenance records for each file and prior package provenance records.
- **Withdraw digiprov:** Write digital preservation provenance for this withdrawal according to PREMIS, i.e. premis events and agents describing the withdrawal process.
- **Make aip descriptor:** create a new AIP descriptor for the withdrawn AIP, retaining only package-level information.
- **Validate aip descriptor:** Ensure that the created AIP descriptor conforms to Guidelines for using PREMIS in METS (<http://www.loc.gov/standards/premis/guidelines-premismets.pdf>)
- **Withdraw AIP:** delete the AIP from storage and replace the existing AIP descriptor in the DAITSS preservation database with the new AIP descriptor.

A Tour of the Operator's Toolbar Tabs

The following is an overview of the functionality and contents of each of the tabs in the Operator's Toolbar:

"dashboard" tab

- The "dashboard" tab displays your user ID in the upper right-hand corner of the screen.
- The "dashboard" tab gives an overview of all SIPs and AIPs recorded in the preservation database.
- "sip summary" provides a summary of all submission attempts and includes statistics on SIPs rejected by submission.
- "aip summary" provides a summary of all AIPs stored in the repository.
- Your own branding or logo can easily be substituted for the default "Florida Digital Archive" logo.
- The DAITSS software version is displayed on the dashboard.
- To log out of the DAITSS Preservation Interface you must be in the "dashboard" tab

The screenshot shows the DAITSS v0.17.9 dashboard. At the top, there is a navigation bar with links: [dashboard](#), [packages](#), [errors](#), [workspace](#), [stashspace](#), [batches](#), [requests](#), [log](#), and [admin](#). The user is logged in as [Ollie_operator](#). The main content area features the Florida Digital Archive logo and the version number **daitss v0.17.9**. A [logout](#) button is visible. Below this, there are two summary sections:

sip summary

count	2813
files	24858
total size	85819.00 MB

aip summary

count	2733
files	12598
total size	37553.00 MB

Figure 1

“packages” tab

- The "packages" tab lists all recent package processing activity, displaying the most recent Operations Event and timestamp of that event.
- Subsets of all packages can be scoped/filtered and listed here, and the resulting list can be downloaded in a CSV file (comma separated values file) for importing into spreadsheets, etc.
- SIPs can be submitted one at a time for processing via this tab.
- Individual package information can be retrieved by Internal Entity ID (IEID) or package/SIP name.
- Clicking on an Internal Entity ID (IEID) of a package will retrieve detailed information about that package.
- Clicking on a package name will retrieve a list of all instances of a package by that name which have been processed by the repository.
- A batch can be created from the packages listed in this display.

dashboard : [packages](#) : [errors](#) : [workspace](#) : [stashspace](#) : [batches](#) : [requests](#) : [log](#) : [admin](#)

submit

packages must be in *tar* or *zip* format

File

Batch

Note

search

aip or *sip* separated by spaces.

ieid or package name

batch	account	project	start date	end date	activity	set scope
<input type="text"/>	FDA	<input type="text"/>	09/02/2010	09/13/2011	<input type="text"/>	<input type="button" value="Set Scope"/>

73 packages **36.95 GB** **4992 files**

intellectual entity id (ieid)	package	account	project	size (MB)	# of datafiles	lastest activity	time
EWFd60BRF_XK3P7K	FDA00067232_00001	FDA	FDA	115.16	800	ingest finished	Fri Sep 02 2011 01:07:29 PM
ETUJ0CBF1_XGRSJC	FDA00028308_00128	FDA	FDA	1077.53	138	ingest finished	Fri Sep 02 2011 01:37:33 PM
ESWZQ6LOW_OT7KCY	FDA00028419_02637	FDA	FDA	489.13	100	ingest finished	Fri Sep 02 2011 02:37:43 PM
EL4R9077D_E5R5V3	FDA00028308_00127	FDA	FDA	1198.39	148	ingest finished	Fri Sep 02 2011 01:22:55 PM

“errors” tab

- The “errors” tab provides a way to track any WIPs that have not been fully processed due to system errors. (Note: a system error indicates a hardware or software problem and does not imply that the original SIP is invalid. Invalid SIPs are rejected by the submission process and are never transformed into WIPs.)
- The “status” column displays the most recent processing event for each package.
- The “error message” column provides a brief display of the error message.
- Clicking on the “intellectual entity id (ieid)” of a package will display details about an individual package.
- Clicking on the “package” (name) column will list all packages by that name. (Note: the package name is not required to be unique.)
- The list of packages with errors can be scoped/filtered.
- Any packages displayed on this page can be saved as a batch.

dashboard : packages : **errors** : workspace : stashspace : batches : requests : log : admin Ollie_opera

errors

filter scope

batch	account	project	start date	end date	status	error message	set scope
<input type="text"/>	FDA	<input type="text"/>	<input type="text"/>	09/02/2011	<input type="text"/>	<input type="text"/>	Set Scope

results

2 packages 1.36 GB 10 files

intellectual entity id (ieid)	package	account	project	size (MB)	# of datafiles	status	error date	error message
EHKFVHPGH_6M7NQV	FDA6151101	FDA	DZA	1338.92	3	stashed	Fri Sep 02 2011 12:08:50 PM	private method 'gsub' called for nil:NilClass
ERWN6Z9E4_IUP51Q	UF00100139_00015	FDA	default	23.06	7	stashed	Tue Jul 19 2011 11:31:33 AM	undefined method 'xml' for nil:NilClass

Save as Batch

batch name

“workspace” tab

The DAITSS workspace tab provides Operators with information about the processing status of all WIPs in the DAITSS workspace and enables Operators to manage package processing.

Some key features of the workspace tab:

- All WIPs or jobs with a Processing Status of “idle” (waiting to be processed) are summarized, providing the Operator with information about WIPs in the workspace that are waiting to be processed, and displays the number of idle WIPs for each processing task.
- All running, stopped, or error packages in the workspace are listed in the “active jobs” display, and details about each active job can be viewed by clicking on the Intellectual Entity ID (IEID) of a WIP.
- The “throttle” settings that control the number of simultaneous processes by type are displayed on this page.
- Allows an Operator user to scope/filter the list of jobs by batch name, account, project, start date, end date, and WIP status.
- Allows an Operator to change the Processing Status of one WIP, the entire Workspace, or a scoped list of WIPs. For example: by scoping the workspace for all jobs with a status of “error”, jobs with errors can be set aside in a stashspace to await resolution of the underlying problem.
- Any “running” jobs can be stopped by using the “update jobs” section of the workspace display.

Chapter 6: DAITSS Workflow Interface

- Any ERROR jobs can be “reset” to IDLE for resumption of processing after a problem has been corrected.
- Any jobs, with the exception of RUNNING jobs can be “stashed” to an existing stashspace.

dashboard : packages : errors : **workspace** : stashspace : batches : requests : log : admin

batch	account	project	start date	end date	status	set scope
			09/22/2011	09/27/2011		<input type="button" value="Set Scope"/>

update jobs

start stop reset do over stash

idle jobs: 0
size: 0.00 B

task	count
ingest	0
d1refresh	0
disseminate	0
withdraw	0

active jobs: 1
running: 1 throttles: 2 ingests; 2 disseminations; 2 withdrawals; 2 d1 refreshes;
size: 1.34 GB

intellectual entity id (ieid)	package	task	status	time
EYJRE10PU_KZ025U	FDA6161902	ingest	running	Mon Sep 26 14:47:20 -0400 2011

“stashspace” tab

The DAITSS “stashspace” is a secure area within DAITSS that can be used to set aside WIPs during processing. The DAITSS “stashspace” consists of separate “stash bins” that can contain one or many WIPs.

- “stash bins”, or holding areas for WIPs can be created/modified/deleted here.
- WIPs in “stash bins” will not be processed until they are “unstashed” and returned to the DAITSS workspace.

Chapter 6: DAITSS Workflow Interface

- Only WIPs that are not actively "running", i.e., those with a status of "IDLE", "STOPPED", "RESET", "DO OVER" or "ERROR" can be moved to a "stash bin".
- Only one operation is allowed on the contents of an entire stash bin: "unstash all". However, additional actions are permitted on individual WIPs within a stash bin.
- "Stash bins" are commonly used to organize WIPs removed from the Workspace. WIPs may be removed from the Workspace for a variety of reasons, including:
 - requests to delay processing of submitted WIPs
 - "ERROR" WIPs that require additional troubleshooting



"batches" tab

"batches" are essentially lists of packages stored in a table in the DAITSS preservation database and viewable via the DAITSS "batches" tab. Creating a batch, or list of packages can enable an Operator to easily track the processing status and progress of a logical group of packages.

- Batch lists can be created/modified/deleted within this tab. (Note: deleting a batch only deletes a batch list and does *not* delete actual packages.)
- Batch lists can be used to Disseminate or Withdraw a group of AIPs.
- SIPs can be assigned to a batch during the submission process for ease of tracking. (The processing status and progress of a batch of WIPs can be tracked via the "batches" tab.)
- A single package can be listed in multiple batches. For example: upon submission a group of packages can be included in a single batch, but rejected packages from that group may also be included in an additional batch for ease of tracking.
- Packages can be added or removed from a batch at any time, by editing the list of IEIDs in the "modify batch" section of a batch display.
- The "activity summary" within each batch lists up to 400 packages provides an overview of the processing status of all packages with the batch.

dashboard : packages : errors : workspace : stashspace : **batches** : requests : log : admin

batches

create new batch

batch name insert whitespace separated list of IEIDs here

existing batches

batch name	# of packages	disseminate	withdraw	peek	delete
ateams	7	<input type="button" value="Disseminate"/>	<input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
BENCHMARK_20110916	1	<input type="button" value="Disseminate"/>	<input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
CF_fixed	2	<input type="button" value="Disseminate"/>	<input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
FTPDL_20110720	399	<input type="button" value="Disseminate"/>	<input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>
FTPDL_20110916	395	<input type="button" value="Disseminate"/>	<input type="button" value="Withdraw"/>	coming soon	<input type="button" value="Delete"/>

- Within a batch the “detailed package list” can be downloaded as a CSV file (comma-separated values file) for importing into a spreadsheet.

dashboard : packages : errors : workspace : stashspace : batches : requests : log : admin

ateams

modify batch

E7TOIYCGD_YALO2Z
EG7KGRVRF_4T44I5
EH5E1ULGY_6QBEU1

Submit

submit request

request note disseminate

activity summary

latest event	count
withdraw finished	2
ingest report written out for ftp	1
reject report written out for ftp	1
fixity success	3

detailed package list (7 total, up to 400 displayed)

[display all](#) [display 400](#)

intellectual entity id (ieid)	package	account	project	size (MB)	# of datafiles	latest activity	time
E7TOIYCGD_YALO2Z	ateam	ACT	PRJ	0.92	2	withdraw finished	Tue Jul 12 2011 03:52:00 PM
EG7KGRVRF_4T44I5	ateam	ACT	PRJ	0.92	2	withdraw finished	Tue Jul 12 2011 03:51:45 PM
EH5E1ULGY_6QBEU1	ateam	ACT	PRJ	0.92	2	disseminate finished	Tue Jul 12 2011 04:08:07 PM

“requests” tab

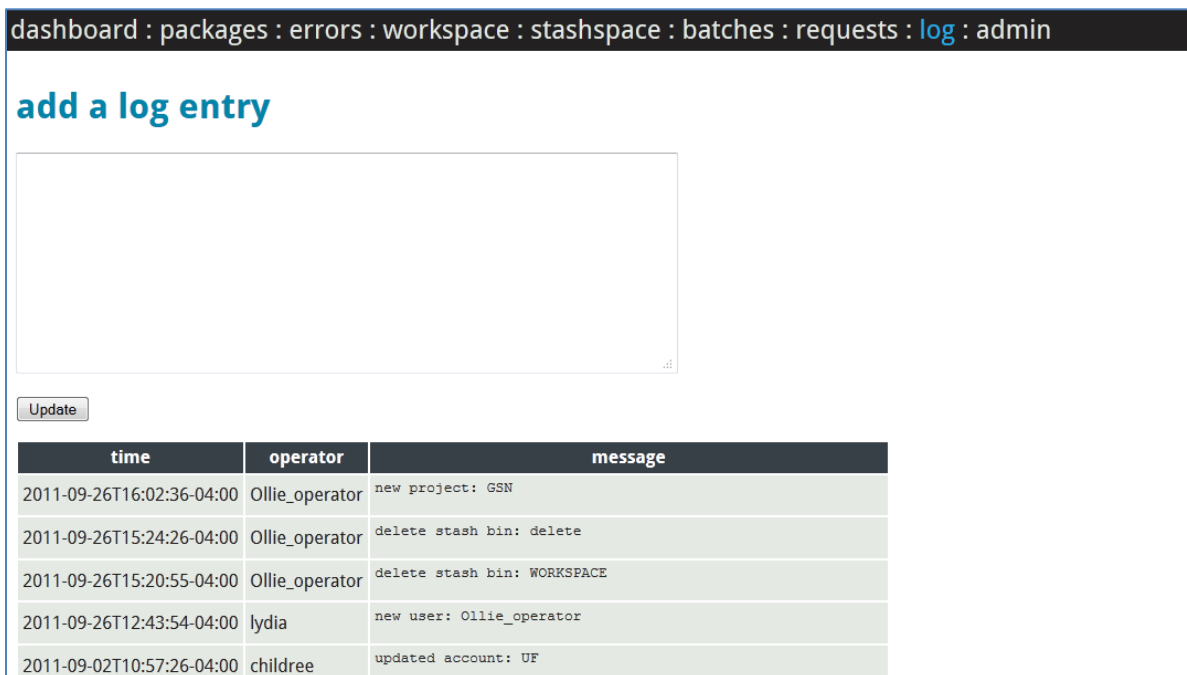
The “requests” tab displays DAITSS requests made on AIPs. Currently DAITSS supports two types of requests: withdrawal requests and dissemination requests. To prevent inadvertent withdrawal of potentially large numbers of AIPs from the repository, withdrawal requests require authorization by an Operator user other than the requesting Operator. (Small repositories can create a second Operator user just for this purpose.) Other features of the “requests” tab:

- Allows cancellation of requests not yet released to the workspace. This can be very useful in cases of erroneously placed withdrawal requests.
- The “request list” displays only the status of the actual request, and not the details of the actual execution of the task, which is performed in a WIP created when a request is “released_to_workspace”. To view details of the execution of the task, click on the IEID of a specific package.

“log” tab

The DAITSS Administrative Log creates a permanent record of key administrative operations:

- The DAITSS Administrative Log automatically records entries for the following activities:
 - Creation/modification/deletion of DAITSS Users
 - Creation/deletion of stashspaces
- In addition, the Administrative Log can be used to record events important to the running of your repository, such as:
 - Installation of software updates
 - System downtime for maintenance or other reasons



dashboard : packages : errors : workspace : stashspace : batches : requests : **log** : admin

add a log entry

Update

time	operator	message
2011-09-26T16:02:36-04:00	Ollie_operator	new project: GSN
2011-09-26T15:24:26-04:00	Ollie_operator	delete stash bin: delete
2011-09-26T15:20:55-04:00	Ollie_operator	delete stash bin: WORKSPACE
2011-09-26T12:43:54-04:00	lydia	new user: Ollie_operator
2011-09-02T10:57:26-04:00	childree	updated account: UF

“admin” tab

- Users, Accounts and Projects can be created/modified/deleted here.
- Click on "projects" or "users" to display a full list of Projects and Users.
- You must create a new Operator type user before archiving materials.