

2016-12 Import - Export Sprint 03 Meetings

- [Call in details:](#)
- [Sprinters](#)
- [Objectives](#)
- [Resources](#)
- [Meetings](#)
 - [Meeting 01 - 2016-12-05](#)
 - [Meeting 02 - 2016-12-07](#)
- [Standups](#)
 - [IRC Template](#)

Call in details:

- Dial-in Number: (712) 775-7035
 - Participant Code: 479307#
 - International numbers: [Conference Call Information](#)
 - You may also call in using the [VoIP dialer](#) from a web browser, or Android/iOS apps
- IRC:
 - [Join the #fcrepo chat room via Freenode Web IRC](#) (enter a unique nick)
 - Or point your IRC client to #fcrepo on [irc.freenode.net](#)

Sprinters

- Developers
 - [Esmé Cowles](#)
 - [Nick Ruest](#)
 - [Jared Whiklo](#)
 - [Danny Bernstein](#) ★
- Testing and Validation
- [Joshua Westgard](#)
 - [Nick Ruest](#)
 - [Kieran Etienne](#)
 - [Adam Wead](#)
 - [Justin Simpson](#)
- Documentation
- [Joshua Westgard](#)
 - [Nick Ruest](#)

Objectives

The objectives of sprint-03 are to:

1. Address Phase-1 issues

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due
Unable to locate Jira server for this macro. It may be due to Application Link configuration.										

2. Implement Phase-2 [requirements for verification tooling](#)
3. Implement Phase-2 requirements

type	key	summary	assignee	reporter	priority	status	resolution	created	updated	due
------	-----	---------	----------	----------	----------	--------	------------	---------	---------	-----

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

4. Plan the production of a few 3-minute videos that demonstrate import/export functionality

Resources

[Design - Import - Export](#)

Meetings

Meeting 01 - 2016-12-05

- 11am ET

Agenda

1. Introductions
2. Logistics
 - a. Virtual daily stand-ups?
 - b. IRC?
 - c. Conference calls?
 - d. Email check-in?
 - e. Sprint retrospective?
3. Sprint Priorities
4. Assign work

Minutes

1. Virtual Stand ups
 - a. IRC (freenode - #fcrepo)
 - b. Drop IRC template by 10am eastern the next day.
2. Calls
 - a. Wednesday morning meeting: 12/7 12pm Eastern - same line
 - b. Wednesday morning meeting: 12/14 12pm Eastern - same line
 - c. Friday 12/16 Retrospective Meeting: 2pm Eastern (no Andrew or Danny)
3. Sprint Priorities
 - a. Purpose of the 3 - minute video is to engage the community as well as provide an entry point for people to use the tools
4. Work Assignment
 - a. Justin to create ticket for verifying external binaries
 - Created
 - b. Can Kieran and Adam use Scholarsphere data to help test the tooling?
 - Yes. Adam to create the ticket
5. Note: The import export tooling runs against the API. Backup/Restore runs against Modeshape.
6. State of the Documentation:
 - Justin to update the import export format documentation; ensuring there are good pointers for getting to the documentation and docs actually are useful for enabling people to use the tool.
 - Justin to review Bag related tickets to make sure they cover what needs covering.

Meeting 02 - 2016-12-07

- 12pm ET

1. <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/48>

Should we support mapping repository paths, not just host/port

[Esmé Cowles](#) - This ticket came out of integration test for round tripping, export and import into a different directory.

We had some allowance for different hostname ports, but base URI and everything else must be the same. At least the base URI could be modified in case of changed deployment.

While doing that Hydra already has various sub-directories, so that could be great. But it brings up issues. How do you handle if you export a resource, but not any linked resources.

Is it a requirement to be able to export a resource/tree of resources and then import it into a different location in a repository. This means we are changing the resource.

There is more complication, but it would be nice to do it. If we run into unsolvable problems, then we can stop. But baseURI as a minimum.

Verification tooling would have to be updated as it not designed to handle this circumstance. If there is a need for this, then we can correct this in the verification tooling.

Does this resolve the Phase 1 #6 wish. "Support import into an existing empty container."

We have been assuming that the container is going to be the same, host/port is a hard one to adjust for, changing base URI in Tomcat is somewhat easier.

Don't want to get sidetracked on this issue and not achieve the sprint goals.

Base Path is generally "/fcrepo/rest"

Ingesting in to a Hydra instance under /dev and then after verifying moving it to /prod.

Esmé Cowles will make a ticket or two and point to pull/48 as a first stab and move onto integration tests.

Final question "What would we want inter-resource relationships to look like." when ingesting to a new location. Principle was what would it look like if you exported the entire repository and import the entire repository. We haven't talk through all the possibilities, it would be good to make sure we agree on this.

2. Validation

<https://github.com/fcrepo4-labs/fcrepo-import-export-verify/pull/7>

Hope the PR is complete, Josh has not tested but feels it is almost there. The latest commit just changed the default logger level.

Should be able to test the latest version this afternoon and merge.

It would be great if that issue could be closed, and the config file as changed a bit and that needs to be dealt with.

Won't have working tool until the update to deal with new config file options. Should be a quick fix (possibly completed today).

Still haven't sorted out the issues around verifying RDF, filtering out server managed triples

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Next person that runs in the "Too many open files in system" should document the steps in a new JIRA ticket.

Nick has done lubm and the 10k indirects and never run into this. Josh ran the one-click on his machine and he may have had other applications which could have used up some file handles

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Probably need to walk through all the verification tickets and close tickets that might have been resolved by existing work.

summary	type	assignee	reporter	priority	status	resolution
---------	------	----------	----------	----------	--------	------------

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

3. Logging

<https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>

Unable to locate Jira server for this macro. It may be due to Application Link configuration. and Unable to locate Jira server for this macro. It may be due to Application Link configuration. It a

Unable to locate Jira server for this macro. It may be due to Application Link configuration. is a different issue but has some

4. Esmé Cowles is creating a ticket for setting Digest header when importing checksum for binary

5. Danny Bernstein is moving onto

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

6. Andrew Woods suggests we all

Only a few tickets left in progress and in review

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

Unable to locate Jira server for this macro. It may be due to Application Link configuration. - leaving for later.

Verification has some errors showing up, was kind of hard as import-export tool was constantly updating.

Perhaps not have both tools worked at same time would be good, but having lots of people around was helpful.

— Good —

IRC standups +++

3 java developers.

For some on their first sprint attending meetings and learning the processes was helpful.

Nice having for testers to have developers available to ask questions.

Nice to not be a Java developer so they could use the tool more like an end user.

4 new people jump in successfully

2 people doing user testing

— Bad —

code reviews could take some time

people in and out during sprint/travel ++

LoC BagIt stuff

JIRA made it hard to see what was going on (maybe an embedded jira chart on the sprint page).

Felt like there wasn't a lot to work on (for reviewing)

Using (clicking on) a JIRA label was handy once you figure it out.

Nice to have Andrew around

Working on import/export and verification at the same time might be more pain than its worth.

—Questions —

With import/export and verification would it have useful to have matching tags between the two products to synchronize. — dev was too fast at this point and there is no release for either.

— Next Steps —

Some stuff to get us to Stage 3

Lots of stakeholder feedback

Ready for a 0.1 release and have people try it out.

Leaving Bag profiles tickets open for now and have stakeholders verify that they are ready.

Standups

IRC Template

```
[Import/Export Standup]
Finished yesterday:
{ticket titles and associated JIRA links}
{and/or brief textual description}
Working on today:
{ticket titles and associated JIRA links}
{and/or brief textual description}
Blockers:
{brief textual description}
```

<ruebot> [Import/Export Standup]

```
19:26:17 <ruebot> [Import/Export Standup]
19:26:21 <ruebot> Finished yesterday:
19:26:25 <ruebot> (today)
19:26:29 <ruebot> - Wearing sprint leader hat
19:26:32 <ruebot> - Tested & resolved https://jira.duraspace.org/browse/FCREPO-2259
19:26:35 <ruebot> - I shouldn't have assigned myself https://jira.duraspace.org/browse/FCREPO-2252
19:26:39 <ruebot> - Testing/validation https://github.com/fcrepo4-labs/fcrepo-import-export/pull/47
19:26:45 <ruebot> Working on today (tomorrow):
19:26:50 <ruebot> - https://jira.duraspace.org/browse/FCREPO-2237
19:26:53 <ruebot> - Testing/validation https://github.com/fcrepo4-labs/fcrepo-import-export/pull/47
19:26:56 <ruebot> - Testing/validation https://github.com/fcrepo4-labs/fcrepo-import-export/pull/48
19:26:59 <ruebot> Blockers:
19:27:02 <ruebot> - I have an interview in the morning, and won't be back around until 11AM ET at the earliest
```

<awoods> [Import/Export Standup]

```
00:30:31 <awoods> Finished yesterday:
00:30:31 <awoods> Updated JIRA permissions to allow "Reviewer" to close tickets
00:30:31 <awoods> Working on today:
00:30:31 <awoods> Keeping up with sprint team activity... and helping when needed
00:30:31 <awoods> Blockers:
00:30:33 <awoods> None
```

<justinsimpson> [Import/Export Standup]

08:41:48 < justinsimpson> Finished yesterday:
08:41:54 < justinsimpson> - attended first sprint

justinsimpson> - reviewed documentation for
configuration.

Unable to locate Jira server for this macro. It may be due to Application Link

justinsimpson> Working on today:
08:42:13 < justinsimpson> - review Jira tickets related to Bag Profiles
08:42:19 < justinsimpson> - reinstalling vagrant vm with latest code
08:42:24 < justinsimpson> Blockers:
08:42:29 < justinsimpson> - I will be on irc only sporadically due to meetings today

<awead> [Import/Export Standup]

08:46:25 < awead> Finished yesterday:
08:46:27 < awead> Approval of PR to add Hydra/Sufia dataset for testing
08:46:28 < awead> Working on today:
08:46:29 < awead> Testing 2240 (Pair-trees)
08:46:30 < awead> Blockers:
08:46:31 < awead> none

<escowles> [Import/Export Standup]

09:14:44 < escowles_> Yesterday: FCREPO-2240 Pairtrees should be Pairtrees
09:15:25 < escowles_> Today: Responding to testing, then FCREPO-2192 Adding more extensive integration tests
09:15:31 < escowles_> Blockers: None

<whikloj> [Import/Export Standup]

09:46:53 < whikloj> Finished yesterday:
09:46:56 < whikloj> nothing
09:47:02 < whikloj> Working on today:
09:47:11 < whikloj> FCREPO-2169 and FCREPO-2170
09:47:18 < whikloj> Blockers:
09:47:21 < whikloj> none

<pktienne> [Import/Export Standup]

09:51:28 < pktienne> Finished yesterday:
09:51:28 < pktienne> - Investigated PSU infrastructure for backup.
09:51:28 < pktienne> - Ran backup of Fedora 4.0 data off staging environment.
09:51:28 < pktienne> Working on Today:
09:51:28 >> OpenURL 11
09:51:28 < pktienne> - Continue setting up full PSU export (<https://jira.duraspace.org/browse/FCREPO-2326>)
09:51:31 >> OpenURL 12
09:51:31 < pktienne> - Attempt verification on PSU export (<https://jira.duraspace.org/browse/FCREPO-2190>)
09:51:33 < pktienne> Blockers:
09:51:35 < pktienne> - None

<westgard> [Import/Export Standup]

10:03:29 < westgard> Finished yesterday:
10:03:31 < westgard> Reviewing PR (<https://github.com/fcrepo4-labs/fcrepo-import-export-verify/pull/7>) for <https://jira.duraspace.org/browse/FCREPO-2284>
10:03:33 < westgard> Created ticket for verification tool to handle ".external" resources (<https://jira.duraspace.org/browse/FCREPO-2327>)
10:03:35 < westgard> Looked at <https://jira.duraspace.org/browse/FCREPO-2322> and started thinking about what to do there.
10:03:37 < westgard> Working on today:
10:03:39 < westgard> We involved in a search committee campus visit today, so time will be limited. Will try to sort out above PR today.
10:03:41 < westgard> Blockers:
10:03:43 < westgard> Question around two features of the verification tool (logging and verbose mode) and their relation to the above PR. We just need to decide what our goals for those features are and then merge or update the PR accordingly.Y

<dbernstein> [Import/Export Standup]

12:35:50 < dbernstein> Finished yesterday: Nothing
12:35:51 < dbernstein> Working on today:
12:35:53 >> OpenURL 4

12:35:53 < dbernstein> Working on today: Unable to locate Jira server for this macro. It may be due to Application Link configuration. today

12:35:54 < dbernstein> Blockers:
12:35:54 < dbernstein> None

<whikloj> [Import/Export Standup]

20:47:23 < whikloj> Finished yesterday: n/a
20:47:44 >> OpenURL 15
20:47:44 < whikloj> Working on today: Finish up <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>

< awoods> [Import/Export Standup]

00:14:11 < awoods> Finished yesterday:
00:14:11 < awoods> - Reviewed/ReOpened: "Rip out the --binDir and --descDir"
00:14:11 < awoods> ** <https://jira.duraspace.org/browse/FCREPO-2226>
00:14:11 < awoods> - Reviewed/OK'd: "Can't import into sub-path"
00:14:13 < awoods> ** <https://jira.duraspace.org/browse/FCREPO-2328>
00:14:15 < awoods> Working on today:
00:14:17 < awoods> - Reviewing tickets as they come up
00:14:19 < awoods> Blockers:
00:14:21 < awoods> - None

< dbernstein> [Import/Export Standup]

00:22:28 < dbernstein> Finished yesterday:
00:22:30 < dbernstein> NA
00:22:31 < dbernstein> Working on today:
00:22:33 >> OpenURL 18
00:22:33 < dbernstein> Finishing up <https://jira.duraspace.org/browse/FCREPO-2226>
00:22:34 >> OpenURL 19
00:22:34 < dbernstein> Starting on <https://jira.duraspace.org/browse/FCREPO-2296>
00:22:35 < dbernstein> Blockers:
00:22:36 < dbernstein> None

< awead> [Import/Export Standup]

08:48:36 < awead> Finished yesterday:
08:48:38 < awead> <https://jira.duraspace.org/browse/FCREPO-2240>
08:48:39 < awead> Working on today:
08:48:40 < awead> Nothing. Tag me as a reviewer on anything you need.
08:48:41 < awead> Blockers:
08:48:42 < awead> I'll be offline for a little while today.

< ruebot> [Import/Export Standup]

08:50:27 < ruebot> Finished yesterday:
08:50:30 < ruebot> - Test, merged <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/47>
08:50:34 < ruebot> - Tested, merged <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/50>
08:50:37 < ruebot> - Moved <https://jira.duraspace.org/browse/FCREPO-2237> to phase 3
08:50:42 < ruebot> - Testing, reviewing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/48>
08:50:52 < ruebot> - Testing, reviewing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>
08:50:57 < ruebot> Working on today:
08:51:01 < ruebot> - Testing, reviewing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/48>
08:51:04 < ruebot> - Testing, reviewing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>
08:51:07 < ruebot> - Sprint check-in call
08:51:10 < ruebot> - Dive-in to <https://github.com/fcrepo4-labs/fcrepo-import-export-verify>
08:51:13 < ruebot> Blockers:
08:51:18 < ruebot> - Nothing

< justinsimpson> [Import/Export Standup]

08:57:32 < justinsimpson> Finished yesterday:
08:57:38 < justinsimpson> -installed local vagrant box with FC 4.7.0
08:57:45 < justinsimpson> -reviewed docs and jira tickets on bag profiles
08:57:50 < justinsimpson> Working on today:

08:58:05 < justinsimpson> - updating docs in confluence to reflect changes in Application Link configuration.
08:58:05 < justinsimpson> - updating bag profile for Archivematica

Unable to locate Jira server for this macro. It may be due to

08:58:05 < justinsimpson> Blockers:

08:58:10 < justinsimpson> -I will miss the first half or 3/4 of the check-in call today, possibly the whole thing. I will ping ruebot after for a link to notes

< escowles> [Import/Export Standup]

09:13:43 < escowles> - yesterday: finished FCREPO-2240 (skipping Pairtree resources) and worked on FCREPO-2328 (importing into subpath)
09:14:18 < escowles> - today: working on FCREPO-2192 (adding more rigorous integration tests)
09:14:34 < escowles> - blockers: none, but leaving early to take my daughter to the orthodontist

< pktienne> [Import/Export Standup]

09:27:35 < pktienne> Finished Yesterday:
09:27:35 < pktienne> - Completed a few attempts at export of PSU production data.
09:27:35 < pktienne> - Resolved HTTP 500 errors (halted export process) arising from some records' bad content.
09:27:35 < pktienne> Working on Today:
09:27:35 < pktienne> - Completing prod export & PSU prod repo.
09:27:37 < pktienne> - Beginning prod export & PSU prod repo.
09:27:39 < pktienne> Blockers:
09:27:41 < pktienne> - None

Unable to locate Jira server for this macro. It may be due to Application Link configuration.

utilizing PSU

< escowles> [Import/Export Standup]

09:07:53 < escowles> yesterday: stopped work on <https://jira.duraspace.org/browse/FCREPO-2328> (importing into subpath), started on <https://jira.duraspace.org/browse/FCREPO-2192> (better integration tests)
09:08:25 < escowles> today: finish up integration tests, few hours of meetings, start on <https://jira.duraspace.org/browse/FCREPO-2225> (recursive retrieval by custom predicates)
09:08:30 < escowles> blockers: none

< ruebot> [Import/Export Standup]

09:12:05 < ruebot> Finished yesterday:
09:12:08 < ruebot> - Reviewing, testing, discussing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/48>
09:12:10 < ruebot> - Reviewing, testing, discussing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>
09:12:14 < ruebot> Working on today:
09:12:16 < ruebot> - Hopefully merging:
09:12:19 < ruebot> - Reviewing, testing, discussing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/48>
09:12:22 < ruebot> - Reviewing, testing, discussing <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>
09:12:25 < ruebot> - Writing some Java and Python for <https://jira.duraspace.org/browse/FCREPO-2332> -- yaml for java, configparser for python
09:12:29 < ruebot> Blockers:
09:12:32 < ruebot> - My skillset combined with the amount of time in a single day.

< awead> [Import/Export Standup]

09:20:38 < awead> Finished yesterday:
09:20:39 < awead> nothing, I was out most of the day
09:20:41 < awead> Working on today:
09:20:42 < awead> Nothing. :(Tag me as a reviewer on anything you need.
09:20:44 < awead> Blockers:
09:20:45 < awead> none

< pktienne> [Import/Export Standup]

09:28:39 < pktienne> Finished Yesterday:
09:28:39 < pktienne> - Finally resolved all http 500 issues which were preventing export of PSU data.
09:28:39 < pktienne> - Completed a full export of PSU production data (pre FCREPO-2226 PR)
09:28:39 < pktienne> - Completed a full export of PSU production data (post FCREPO-2226 PR)
09:28:41 < pktienne> Working on Today:
09:28:43 < pktienne> - Run verification tool against full PSU production data set (pre 2226 PR)
09:28:45 < pktienne> - Run verification tool against full PSU production data set (post 2226 PR)
09:28:47 < pktienne> Blockers:
09:28:49 < pktienne> - None at the moment.

< justinsimpson> [Import/Export Standup]

09:51:59 < justinsimpson> Yesterday:
09:52:06 < justinsimpson> very little, bogged down by meetings
09:52:11 < justinsimpson> Today:
09:52:18 < justinsimpson> Updating docs for <https://jira.duraspace.org/browse/FCREPO-2226>
09:52:27 < justinsimpson> try to follow discussion on <https://jira.duraspace.org/browse/FCREPO-2332>
09:52:32 < justinsimpson> Blockers:
09:52:33 < justinsimpson> none

< westgard> [Import/Export Standup]

09:58:32 < westgard> Finished yesterday:
09:58:34 < westgard> finished reviewing tool integration (FCREPO-2284) PR and merged
09:58:36 < westgard> Working on today:
09:58:38 < westgard> started working on better error handling (FCREPO-2329)
09:58:40 < westgard> Blockers:
09:58:42 < westgard> FCREPO-2327 (handle external resources) and FCREPO-2329 (handle missing resources) both depend somewhat on sorting out the config file change issue, but I have a local workaround in the meantime

< dbernstein> [Import/Export Standup]

10:17:39 < dbernstein> Finished yesterday:
10:17:41 < dbernstein> <https://jira.duraspace.org/browse/FCREPO-2226>
10:17:42 < dbernstein> (Rip out the --binDir and --descDir)
10:17:44 < dbernstein> Working on today:
10:17:45 < dbernstein> Putting the finishing touches on
10:17:47 < dbernstein> <https://jira.duraspace.org/browse/FCREPO-2296>
10:17:48 < dbernstein> (implement binary redirects approach)
10:17:50 < dbernstein> Mods + new unit tests are done and passing.
10:17:51 < dbernstein> Will issue pull request later this morning.
10:17:52 < dbernstein> Blockers:
10:17:53 < dbernstein> None.

< whikloj> [Import/Export Standup]

10:19:59 < whikloj> Finished yesterday: Most of <https://github.com/fcrepo4-labs/fcrepo-import-export/pull/49>
10:20:13 < whikloj> Working on today:
10:21:13 < whikloj> Take audit log out of logback and back into code, turn off by default and make log directory a config parameter
10:21:17 < whikloj> Blockers:
10:21:38 < whikloj> None, but would be good to get some stakeholder feedback on the current work