

Tapestry - Feature #8688

Accept authenticated API calls from data providers to add datasets to a public profile

03/14/2016 08:20 PM - Ward Vandewege

Status:	Resolved	Start date:	03/24/2016
Priority:	Normal	Due date:	
Assigned To:	Tom Clegg	% Done:	100%
Category:	Third party integration	Estimated time:	0.00 hour
Target version:	Interpretation automation		
Description			
The API call will have these arguments:			
a) a unique identifier for the external data provider			
b) the 3rd party activity identifier for the pgp participant which was sent to the data provider when the participant signed up for the activity			
c) a payload which indicates what file(s) to import. Details of the format of that payload are TBD, but at a minimum it should probably be a collection of URLs, ideally associated with a file type for each individual file to download. The URLs <i>should</i> include either a manifest file with MD5 digests (like http://cdimage.debian.org/debian-cd/8.3.0/amd64/iso-cd/MD5SUMS), or an *.md5 file for each data file. The size of each file should also be included.			
Acceptance criteria			
When a researcher or admin shows/edits a third party study, the unique identifier for the study is shown. (This is passed as api_key.)			
Given a POST /third_party/add_dataset request:			
1. Ensure api_key corresponds to an approved third-party study			
2. Ensure participant_code param corresponds to a participant who is participating in the study (otherwise HTTP 4xx)			
3. Ensure data_sources_tsv param is a TSV document, where each line has a URI, size, and hex md5 digest of the corresponding file data			
4. Respond HTTP 4xx if any parameters are invalid or there is an error submitting the download job; otherwise 200			
Subtasks:			
Task # 8792: Review 8688-pull-data-api			Resolved
Related issues:			
Related to Tapestry - Feature #8689: Trigger an arvados job to retrieve the d...		In Progress	03/14/2016
Related to Tapestry - Feature #8707: Arvados job: download data from remote s...		In Progress	03/15/2016

Associated revisions

Revision c15f3953 - 03/25/2016 05:44 PM - Tom Clegg

Merge branch '8688-pull-data-api'

refs #8688

Revision 8b161241 - 04/07/2016 07:03 PM - Tom Clegg

8688: Rename param from participant_code to participant_id to match "clickthrough to participate" link.

refs #8688

History

#1 - 03/14/2016 08:26 PM - Ward Vandewege

- Description updated

#2 - 03/14/2016 08:39 PM - Ward Vandewege

- Description updated

#3 - 03/15/2016 08:40 PM - Tom Clegg

- Description updated

#4 - 03/15/2016 09:00 PM - Tom Clegg

- Description updated

#5 - 03/15/2016 09:01 PM - Tom Clegg

- Description updated

#6 - 03/15/2016 09:36 PM - Abram Connelly

- Description updated

#7 - 03/15/2016 09:49 PM - Tom Clegg

- Story points set to 2.0

#8 - 03/18/2016 12:33 AM - Tom Clegg

- Assigned To set to Tom Clegg

- Category set to Third party integration

#9 - 03/18/2016 03:17 AM - Tom Clegg

- Description updated

#10 - 03/19/2016 05:24 AM - Tom Clegg

- Status changed from New to In Progress

#11 - 03/24/2016 03:02 PM - Tom Clegg

- Description updated

- Status changed from In Progress to New

#12 - 03/24/2016 03:28 PM - Tom Clegg

- Status changed from New to In Progress

#13 - 03/24/2016 03:43 PM - Tom Clegg

- Description updated

Branch 8688-pull-data-api @ [4969df6](#)

- Accepts valid API calls: POST /third_party/add_dataset with params api_key, data_sources_tsv, participant_code
- Rejects invalid API calls with 4xx and a JSON response body
- Tries to submit an Arvados pipeline using arv-run-pipeline-instance (arvados-cli must be installed, but this dependency can't be in Gemfile because it needs Ruby 1.9.3) (this is part of [#8689](#))
- Introduces a new config knob ruby19_shell_args (default "rvm-exec 1.9.3")

#14 - 03/25/2016 01:04 AM - Ward Vandewege

I looked at 8688-pull-data-api, and I think it's OK to merge. When merged I can deploy to my-dev and we can test functionality. Thanks for adding a journal table!

#15 - 04/09/2016 05:20 PM - Tom Clegg

- Status changed from In Progress to Resolved