

Tapestry - Feature #8701

Arvados job/pipeline generates ClinVar HTML report from VCF

03/15/2016 05:59 PM - Tom Clegg

| | |
|--|----------------------------------|
| Status: Closed | Start date: 03/22/2016 |
| Priority: Normal | Due date: |
| Assigned To: Abram Connelly | % Done: 100% |
| Category: | Estimated time: 0.00 hour |
| Target version: Interpretation automation | |
| Description | |
| Subtasks: | |
| Task # 8768: Review #8701 | Resolved |

History

#1 - 03/15/2016 08:30 PM - Tom Clegg

- Subject changed from Arvados job/pipeline generates ClinVar HTML report to Arvados job/pipeline generates ClinVar HTML report from VCF

#2 - 03/15/2016 09:30 PM - Ward Vandewege

- Story points set to 2.0

#3 - 03/15/2016 09:56 PM - Ward Vandewege

- Assigned To set to Abram Connelly

#4 - 03/18/2016 05:54 PM - Abram Connelly

- Status changed from New to In Progress

I think the best way to go about this is to separate the pipeline into two jobs. The first will produce a JSON output file and the second will take the JSON file and create an HTML page. I guess we'll need to duplicate the report in the second job so it's available from the same collection?

The first leg of that pipeline is created and looks to be running successfully. I've based it off of Madeleine's [vcf2clinvar](#). I've used [abe.vcf](#) from a project I've created called [ClinVar](#) that holds relevant data for this pipeline.

The pipeline is called [ClinVar Report Pipeline](#). The [test run](#) ran successfully and produced a [JSON 'report'](#) as expected.

#5 - 03/21/2016 07:43 PM - Abram Connelly

I would still like to put these repositories in a GitHub repo to use them instead of using the Arvados Git repo but a test run using the Arvados repos works. The results of running the 'abe.vcf' input file can be seen at:

<https://workbench.su92l.arvadosapi.com/collections/4ff0aefdc5c9d187fd27b57f78c3a2f+1742/clinvar-report.html?disposition=inline&size=1895>

Original pipeline template:

https://workbench.su92l.arvadosapi.com/pipeline_templates/su92l-p5p6p-oopbp7ikcyckcw

The [Bootstrap Table \(MIT/expat license?\)](#) was used to generate the table, have filtered and sortable columns. It's 'bare bones' but should be rich enough to get useful feedback on.

I've tested the other test VCF and had it run with success.

#6 - 03/22/2016 03:23 PM - Abram Connelly

This pipeline along with it's corresponding code has been created and can be reviewed.

Two GitHub repositories have been created:

- <https://github.com/abeconnelly/ClinVarReport>
- <https://github.com/abeconnelly/ClinVarReportHTML>

Where the first generates the JSON report and the second creates the HTML for display. Thought ClinVarReportHTML doesn't do anything other than copy the relevant portion of the git repo and pass the JSON report data through, I think it's still good to separate the report generation from the

report display portions into separate jobs. Though we don't do it here, we could imagine some post processing taking place for presentation and that can be done in the display portion (ClinVarReportHTML here).

A sample run with a pipeline using a test VCF file and using the above two GitHub repos can be found at:

https://workbench.su92l.arvadosapi.com/pipeline_instances/su92l-d1hrv-0fq6djbem3pcrvo#Components

I won't link to it here but the pipeline has also been run with the other test data we've been given.

#7 - 03/31/2016 05:11 PM - Abram Connelly

Copied from the review task:

Review code and pipeline:

- <https://github.com/abeconnelly/ClinVarReport>
- <https://github.com/abeconnelly/ClinVarReportHTML>
- https://workbench.su92l.arvadosapi.com/pipeline_templates/su92l-p5p6p-oopbp7ikcycckcw

Sample run:

- https://workbench.su92l.arvadosapi.com/pipeline_instances/su92l-d1hrv-0fq6djbem3pcrvo

Based on feedback from [#8700](#), I've added a spinner to the ClinVar report as well.

The following sample run should have the updated changes:

- <https://workbench.su92l.arvadosapi.com/collections/42de9014c53b5487c069a5b86909afc8+1911/clinvar-report.html>

#8 - 03/31/2016 06:55 PM - Ward Vandewege

Thanks for addressing that. This looks good to go for me.

#9 - 04/14/2016 03:55 PM - Abram Connelly

- Status changed from *In Progress* to *Closed*