ITCR Training and Outreach Working Group

Thursday, June 8, 2017 Meeting Notes

4pm Eastern

Toll-free: 855-259-6342; \*\*Conference Code: 40152#; \*\* Security PIN: 256871#

**Google Doc for the Meeting:**

<http://bit.ly/itcr-tow-june2017>

**Attendees: Please sign in!!**

Mervi Heiskanen, NCI

Andrey Fedorov, BWH/HMS

Martin Morgan, RPCI

Juli Klemm, NCI

Leah Mechanic, NCI

Bradley Broom, UT MD Anderson Cancer Center

Nathalie Pochet, BWH/HMS/Broad

Michael Reich, UCSD

Anand S. Merchant, NCI CCR

Rao Divi, NCI

Hiro Yoshida, MGH/HMS

Rudi Pillich, NDEx - UCSD

Isaiah Norton, BWH

**Agenda/Minutes**

1. Updates on training activities from the group
	* *Bioconductor*: new MultiAssayExperiment [video](https://www.youtube.com/watch?v=w6HWAHaDpyk); two-day course at [OMRF](https://github.com/Bioconductor/BiocIntro/blob/OMRF-2017/README.md), May 8-9. Boston [meetup](https://github.com/Bioconductor/BiocTalks/blob/Boston-meetup-2017/README.md) (meetup [group](https://www.meetup.com/Boston-R-Bioconductor-for-genomics/) page), April 20. Forthcoming: [CSAMA](http://www.huber.embl.de/csama2017/) week-long R / Bioconductor training.
2. Planning of the WG activities based on the F2F discussions
	* Full meeting notes are [here](https://docs.google.com/document/d/1KDP67NXI9BF-b6lGR_S59NLcuHkN1eIPBnijBYhgB9o/edit)

**TOW Lunch Breakout Session** <http://bit.ly/2rEYkSJ>

Should this group provide outreach to the general public? This would be best done in collaboration with OCPL. E.g., connect popular topics such as AI to real activities in cancer informatics

Office of advocacy/outreach to get patient advocates involved. Consider inviting Tony Dickherber to an ITCR meeting to discuss.

Link concepts such as “data sharing is good” to real data analysis tools and activities.

Question: What are the pros/cons of a “phone home” functionality?

A: Not a huge deal, as long as people can opt out. Much of the software we use every day has features like this.

Maybe ok to do “lightweight” (e.g. version) checking

Summary: This topic did not raise any significant concerns

GitHub Topics: Label tools supported in some way through ITCR

Use “nci-itcr” as the topic for GitHub

Topic: nci-itcr

This is really targeted for the technical audience

Discussion:

* Establish a list of best practices from the working groups on some public site. Where? Perhaps separate Google Doc where we provide updated recommendations.
* This can also be a reference for new grantees e.g., make a video, fill out surveys, contribute to the usage metrics document, tag your github repositories with the recommended topic
* Phone home functionality: The group can share what the benefits are for this type of service, but not necessarily make a recommendation.
* Usage metrics: None are perfect but what is important is how they change over time.
* GitHub Topics: Announce/remind at the next monthly meeting, include in the meeting GoogleDoc.
	+ What granularity should be used to tag?
	+ Can we have simple workflows for end user (a researcher or a clinician and not the bioinformatics specialist) and the tutorials for using the tools at one place?  Probably, linked on ITCR resources website. Github is amazing, but it is little bit intimidating for a biologist/clinician.
	+ Outreach to the general public, example: understanding patient data <https://wellcome.ac.uk/news/understanding-patient-data-launches-today>
	+ Should tutorials/training materials be linked from the ITCR website? The risk is that links can break. Also, there may be multiple tutorials, so multiple links. There may be no one tutorial that covers everything.
	+ How findable are the tools? Can we make this come up higher in a Google search?
	+ Tie in with the Cancer Research special issue? Cross posting, for more visibility.
		- Social media opportunity. LinkedIn, Twitter
		- Thunderclap campaign? Create a communications toolkit, including individual divisions.
* Granularity of the tools reported to the ITCR web site
	+ Example: in QIICR, we have a variety of tools we developed, of different levels of significance and importance. How to decide which ones should be “advertised” in the ITCR table? <https://github.com/QIICR>
		- Should it be 1 grant, 1 entry? Seems to be project dependent.
* Input for ITCR Renewal
	+ **How many active users do you have? How many web access, downloads, and citations of your tools have been recorded?  It would be helpful to show the trend in the last year or two in graphical form. Among the active users, how many of them are in the cancer research field?**
	+ **How is your product disseminated?  Do you have any mechanism of user support and user training?  How is the source code disseminated? Is it being used by others and how?**
	+ **How do you involve the user community, and cancer researchers in particular, in gathering feedback to improve user experience and to promote adoption of the tools for cancer research?**
	+ **How do you collaborate with other ITCR projects and/or outside informatics projects to improve software interoperability?**
	+ **Provide examples and citations of how their tools have been used to support scientific discoveries.**
	+ Which tools are being clinically validated? What problems the investigators are facing in clinical validation?  What do they need to validate the tools for clinical use?

Discussion:

* Quantification is challenging. Can we make it more qualitative?
	+ Consider binning the possible answers
* Yes/no questions about resources supporting your tool
	+ End-user oriented documentation? Tutorials? >> emphasize these metrics
	+ Videos? Sample datasets? Developer oriented documentation?
	+ Is there a public resource to access source code?
	+ Do you have a mechanism for users to report issues, ask questions about the tool?
	+ For desktop tools, do you provide a ready to use downloadable binary of the tool?