Agenda

16 September 2014

1-866-910-4857, pass code 323690

1. Charge for the group
2. Short presentations from groups discussing training and outreach approaches they have used. We probably can only get through a few groups, and will continue next time.
3. Discussion with Juli Klemm and Maggie Cam about potential collaboration with CBIIT and CCR – resources for training users
4. Next steps
5. Meeting frequency and time

Attendees

Andriy Fedorov, BWH (QIICR / 3D Slicer project)

Ishwar Chandramouliswaran, NCI

Rebecca Crowley Jacobson (TIES, DeepPhe-cTAKES), University of Pittsburgh

Martin Morgan, Bioconductor / FHCRC

Mike Ryan, JHU (Cravat/MuPIT project)

Bobbie-Jo Webb-Robertson, PNNL

Timothy Tickle, Broad

Juli Klemm, NCI CBIIT

Hiro Yoshida, Mass General Hospital

Jerry Li, NCI

Charge

Jerry Li

* Impact of ITCR will be measured by (for example) # of manuscripts that reference the software.
* We need to be proactive in reaching out to the community
* We will develop a program-wide ITCR training and outreach initiative
* We will F/U on recommendations made at the June ITCR PI F2F, but we are not limited to them
* Our goal is to set priorities for creating impact
* Program will identify how to garner resources and make priorities happen

Short Presentations -- see [discussion forum threads](https://nciphub.org/groups/itcr/forum/default-section/discussions/104)

**Bioconductor / Martin Morgan**

Approaches

* 'Primary' documentation ('man' pages) and higher-level synthetic 'vignettes' essential
* Mailing list, face-to-face training and conferences; published training material
* Recently: work-flows, videos, support site (replaces mailing list)
* Outreach: tutorials / workshops at scientific and technical conferences; scientific publication; twitter presence; StackOverflow / Biostar / etc monitoring

Lessons learned

* Wealth of material requires effective organization
  + Easily browsable, with not too much infrastructure (e.g., tags, not formal ontology; title but not full-text search)
* Videos
  + 'expensive' to produce, time-wise
  + hard to edit, rapidly becoming out of date
  + low information content, difficult to reproduce complex operations
  + net: short videos on focused topics, rather than comprehensive expositions
* Bootstrapping: new users tend not to know where to find training, because they're not familiar with the project!

Needs with respect to centralized resources / approaches

* Discovery
* Technical assistance, e.g., video streaming

**Trinity / Brian Haas and Tim Tickle**

Workshop characteristics:

Sequence based analysis:

- Often characterized by command line calls to modular tools that have long runtimes.

- Machine image with tools pre-installed

- Data installed, a toy data set that can run through tools in a reasonable time

- Then one can step through the process with the toy data set

Downstream statistical analysis:

- Although can be tool / script based (and so characterized and handled as above), often can be scripting within an IDE for a statistical package like R with focus on hands-on interactive sessions.

- All materials can be posted in Git and pulled down to machines and then ran locally (or an image can be used)

-- This way users can write code as you go but also have code available that works

-- Can see in real time

--- Code running

--- Expected output

--- Thier output

--- Built-in Help

--- And work on their scripting skills :-)

A word on images (VM):

- Images can be available for download

-- VirtualBox

- Could be deployed to computers before workshop

-- If made correctly can also be deployed on Amazon Web Services

* (MTM) Bioconductor also uses machine images (AMIs) to make training more effective; sometimes it feels like the trainee is ‘duped’ into not having to deal with technical issues that they will necessarily encounter in real life
* [RCJ] We just started making TIES VM available, and I worry about the above issue. But so far we’ve found this to be useful to user community. We can’t use AWS or any external processing unfortunately as we deal with EMR data derived from patients..

User support (Primary docs and user interaction):

- Mailing list (user and dev)

-- Currently moving things over to google groups.

-- Free

- Videos

-- Youtube

- Publish protocols

-- Journals

- Can leverage markdown in versioning systems / R packages

-- So the documentation is where the code is online and is downloaded with the code

- Have an image for download with data loaded

-- Coincides with published protocols

Lesson-learned:

Tiered Users = Tiered Solutions

- If you can make your tools module / command-line centric

- This integrates into front-ends like galaxy

- In this case, can expose:

-- command-line (for expert users and pipelines)

-- galaxy modules

-- server instances

**3D Slicer / Andriy Fedorov**

* Wikipage structured by releases, can edited by anyone to crowd-source documentation
  + Structuring and organization is also an issue (wiki does not support tags, although is indexed by google)
* Documentation for each module for 3D Slicer is a requirement
* Active mailing lists - separate ones for users and developers
  + Based on content, changes are made to Wiki
* Powerpoint slides and sample datasets accompany software
* Video tutorials sometimes requested - but hard to resource. Used infrequently for specific items
* Tutorials at major conferences (eg RSNA)
* Outreach through conferences, introduction to web resources
* Google plus community for 3D Slicer, post announcements
* Google Hangouts useful, free weekly developer hangouts, attended both by experienced developers and new developers, also 1-to-1 training supported by google hangout. Wiki page used to summarize contents discussed on Google Hangouts
* Project Week - informal, groups of interest are self formed,
* Webinars periodically

Discussion

Juli Klemm

* NCI recognizes needs for training across community
* NCI may be able to help experienced groups make training more widely available, and also help newer groups to develop training methods and activities.
* Are there ways that NCI can support coordination and dissemination?
* Many different needs represented will require many different approaches

Jerry Li

* Monthly webinars and online training were rated highly on our list. Cancer Informatics training summer course (1 week?) Would that be feasible? Impactful?
* (MTM) Is the primary goal to increase awareness of diverse software tools (as opposed to learning specific tools)? Is it useful to combine presentation of, e.g., 3D Slicer and Trinity? Would short and necessarily superficial presentations serve user community needs, or not go into enough depth to be useful in practice?
* (MTM) Do NCI personnel / others make use of existing week-long courses, e.g., CSHL? Maybe the wheel doesn’t need to be re-invented, or obstacles to participation identified as lessons learned for going forward

Maggie Cam

* Topic specific RNAseq, CHiPseq, etc. Use of tools centered around analysis.
* Intramural scientists are hungry for training, CCR events are oversubscribed. CCR hasn’t used Hangouts, but they are thinking about ways to bring in more training. Working on ways to include super-users to guide hands-on training.

(MTM) Different roles for different training resources

* Short videos for *introduction* to specific topics
* Webinars to provide overviews
* Hangouts / static training for delving in-depth
* Challenges associated with need for trouble-shooting individual problems during video / web presentations

(MTM) Need for carrots to encourage developers to provide training via well-supported and well-defined grant resources

Next steps

Contribute URLs to existing training resources (and perhaps a very quick description  into ITCR collection - Ishwar will identify location for these on NCIPHub.

Contribute cleaned up Research Plans as pdfs that will be added to collections in ITCR available to others in the ITCR group (remember it’s closed to those involved in the project)

[RCJ] Would it make sense to create a matrix of methods for training and outreach and get different groups to identify things they have tried, things they currently do, and things they want to do? I will try to do that before the next meeting.

[MTM] For next steps it might be valuable to have the 'newer' projects go through a similar exercise, outlining how they plan (and are supported to!) reach out to nascent users. Perhaps a useful step after that (two meetings from now) is some thoughtful reflection from NCI staff on what realistic steps -- support both practical and financial -- they think might effectively leverage the lessons / approaches of both new and established projects.

Meeting time and frequency

Monthly, at the same time Tuesday 12-1 pm ET

Next meeting Tuesday, October 14, 2014 12-1 PM ET