2013-12-05 Research computing needs of grad students: UCB focus group & beyond

Announcement and readings:

We'll discuss the results of a recent focus group held with UCB graduate students on the topic of research computing needs. The focus group was convened as part of the Research & Academic Engagement (RAE) Portfolio Planning project, a joint effort between Academic Engagement/ETS, Research IT/IST, and The Library. The summary for the focus group is written up at:

November Graduate Focus Group Findings

Our goal for the reading group is to engage in discussion across broad areas of the campus about the computing needs of students-as-researchers at Berkeley (graduate students in particular though not exclusively). In this vein, we hope you'll also review the home page of Princeton's Graduate Certificate in Computational and Information Science as background to our discussion:

http://www.princeton.edu/researchcomputing/education/graduate-certificate-CIS/

Finally, as supplemental reading, the RAE project also held a focus group with undergraduates. That group focused much more on teaching and learning. The summary is written up at:

2013-10-31 Focus Group Findings

Notes:

Questions to consider:

- How can we get a broader and more representative set of findings about graduate student needs? What are the best ways to engage this group?
- What do UCB graduate students need the most in terms of computing needs to support their research?
- The Princeton certificate in Computation and Information Science has 3 core courses (software engineering, numerical algorithms, interacting with data), of which two must be selected along with a discipline-focused elective. Are these the right courses? Could Berkeley use something like this?

ATTENDING:

Noah Wittman, Claudia von Vacano, Jon Stiles, Nick Adams, Chris Hoffman, Steve Masover, Richard Millet, George Atala, Aron Culich, Rick Jaffe, Allison Bloodworth, Patrick Schmitz

Allison spoke on the RAE focus group. Many possible follow-up questions that, in the single session, we haven't gotten to yet.

Nick Adams (with Claudia von Vacano and Jon Stiles) spoke about D-Lab (social science focused) offerings in support of research -- from orientation to research methods to areas of study to support for using particular tools (R, Stata, et al.). Addressing the 'where do I look for stuff' problem by consolidating information about research methods courses on the D-Lab web site. Asking people:

- what kind of data they work with (qual/mixed/quant);
- what are the key research challenges they face (being oriented to data sources available; training in cleaning and organizing data); and,
- what kind of software and other IT support they need (e.g., for Qualitative data, which package is the right one of six available commercial packages -- and how to stitch together projects given that current qualitative data packages don't scale to handling data/documents from dozens or hundreds of cases).

Cluster consulting to deal with "small" but show-stopping problems students encounter in their research (e.g., a line of code on which their software is crashing), where they don't want to burn one of the rare and short (~20 min once a month) face-time sessions they get with a professor (advisor).

Need for space for collaboration and for workshopping.

Social sciences need for courses that formalize advanced level qualitative methods.

Over 50% of students are interested in more training in R and Stata. Depts. don't offer courses in topics that fill less than a semester ... so focused trainings, or brush-ups when they're ready to use research methods they took a course about a couple years before.

Need for orientation: what you don't know about what you don't know

Some workstations at D-Lab on which folks can run software; some problems can be run on a laptop's power. D-Lab staff have certainly found
people are excited about the possibility of running software from home on virtual machines -- looking into how that might be addressed.

What would be good next steps to flesh out research needs of graduate students?

- Allison: user data research sharing would be v. helpful, not just within this group but across the campus; build a network
- Chris: D-Lab has done great work in this area in social sciences, how can we reach grad students in other disciplines
- Aaron: How does D-Lab scale, how can sibling efforts around the campus 'normalize' what they ask, offer; how can they leverage lessons learned in other disciplines
- Jon: Restricted use data, how to handle ... standardized models for handling data in a way that data producer/owners can approve. D-Lab working with Dave Reiger and Leon Wong to explore what IST can do to support this need.

Chris: another conversation along these lines in January......