Reid McIlroy-Young

April 5, 2017

Question

How do the tools used by scientists shape their work?
Specifically:

What causes new scientific software to be adopted by a community and what communities are most likely to integrate a new tool?

Data

- Scientific meta-data collections
  - Web of Science, pubmed, etc
- Open source portals
  - CRAN, Github, PyPi, etc

Method

- Identify communities / disciplines to study
  - e.g. statistics community
- Develop means of identifying new software tools
  - some subset of journals publish new software: *journal of statistical software*, *The R Journal*, *Journal of Multiscale Modelling and Simulation*
  - Pubmed has some of the linking already done
  - Generalize to other sources by hand and ML techniques

Analysis

- Can we identifying accurately publications introducing new software packages, libraries, code snippets, etc?
- How are new tools distributed within the network?
- What predicts their success, both in the literature and in usage?
- What causes them to adopted by their scientific community, by other communities and/or non-scientists?
- What communities are most frequently adopting new tools?