Lessons learned from neuroimaging tool development in Python

Stephan Gerhard

May 30, 2011

The large complexity of the systems neuroscientific researchers aim to investigate demands sophisticated infrastructure and technology. To tackle the challenges associated with these technical developments, the field of neuroinformatics is emerging.

For a large variety of problems in neuroscience and their associated computational requirements, the Python programming language has been adopted for tool and library development.

Python thus may play a central role in the evolution of the developments necessary to tacke the upcoming technical challenges neuroscience researchers are faced with. Python is apt to be the glue for different legacy systems, and to build tools and libraries to support research questions in a straightforward fashion.

In my talk, I would like to present my subjective experience on Python tool development for the Diffusion MRI community. A number of technical and non-technical aspects of the process will be highlighted, a number of questions are raised which are important to be considered early on in the developmental process and a number of decisions will be outlined which I took in my own projects which from the more experienced perspective I share now, could have been taken differently.

The information I would like to convey aims to serve other researchers starting with Python development to arrive at improved decisions pertaining their own projects. Five relevant aspects shall be touched upon:

Research Environment & Community Reflect on your commitment: Is it tool development à la "Getting-The-Job-Done-For-Me" vs. a community-targeted, reusable, documented, tested tool with a stable API? Do you see a chance of building a community around your tools? Does a community already exist where you could possibly contribute? Do you have close collaborates with a similar vision that care for the tools your are going to develop? What are the constraints of your supervisors and your institution? How can you overcome concerns of making the code available as open source? What license to choose? Who will own the copyright? What are the current practices and tools employed in your lab, and are you able and willing to educate people in order to change them?

The User Perspective How to distinguish between mere users (that do not care about software design or methodological concerns), and user-developers (doing method development and applied research), methods people (focusing on the improvement or creation of new methods, the user interface is very secondary). How to best approach different groups adequately? What are the advantages and disadvantages of building a GUI on top of libraries? Similarly, user-Dialogs vs. writing a scripts? Do you want your code to be used at all by others and can you live with the implications thereof? How can you overcome obstacles of proper software distribution and multi-platform support? What happens if you do not develop according to user requirements, but according to imagined user requirements?

Recommended supporting tools How to adapt to your particular research environment? What are the habits of your collaborators? Do you have ideas to change their habits for the better? What IDE and code editors would you recommend to lab mates? Why to encourage to write unit tests? How to write documentation fast and properly? How to interact with your users?

My tool contributions (so far) I will give a short overview of the tools I have been working on for the last two and a half year, and how they evolved. How modularization and working on proper interfaces turned out to be very beneficial. Why writing things up in a paper is important, and which journals accept software submissions. Why sometimes reinventing the wheel is good, and when to contribute to existing open source projects?

Open Science What would be a vision for "Python in Neuroscience", and how would it tie into the emerging open science movement? Why do we need reusable, well-documented and tested libraries, especially in neuroscience? What could be your contribution?