

# Experience on a Brazilian Workshop of Python for Biological Data (BWPBD) in 2017

Renato Augusto Corrêa dos Santos\*, Leonardo Martins Silva, Franciele Grego Esteves, Diego M. Riaño-Pachón, Alisson Hayasi da Costa, Ana Ciconelle, Bruno F. de Souza, Flávia V. Winck, and Raniere G. C. da Silva

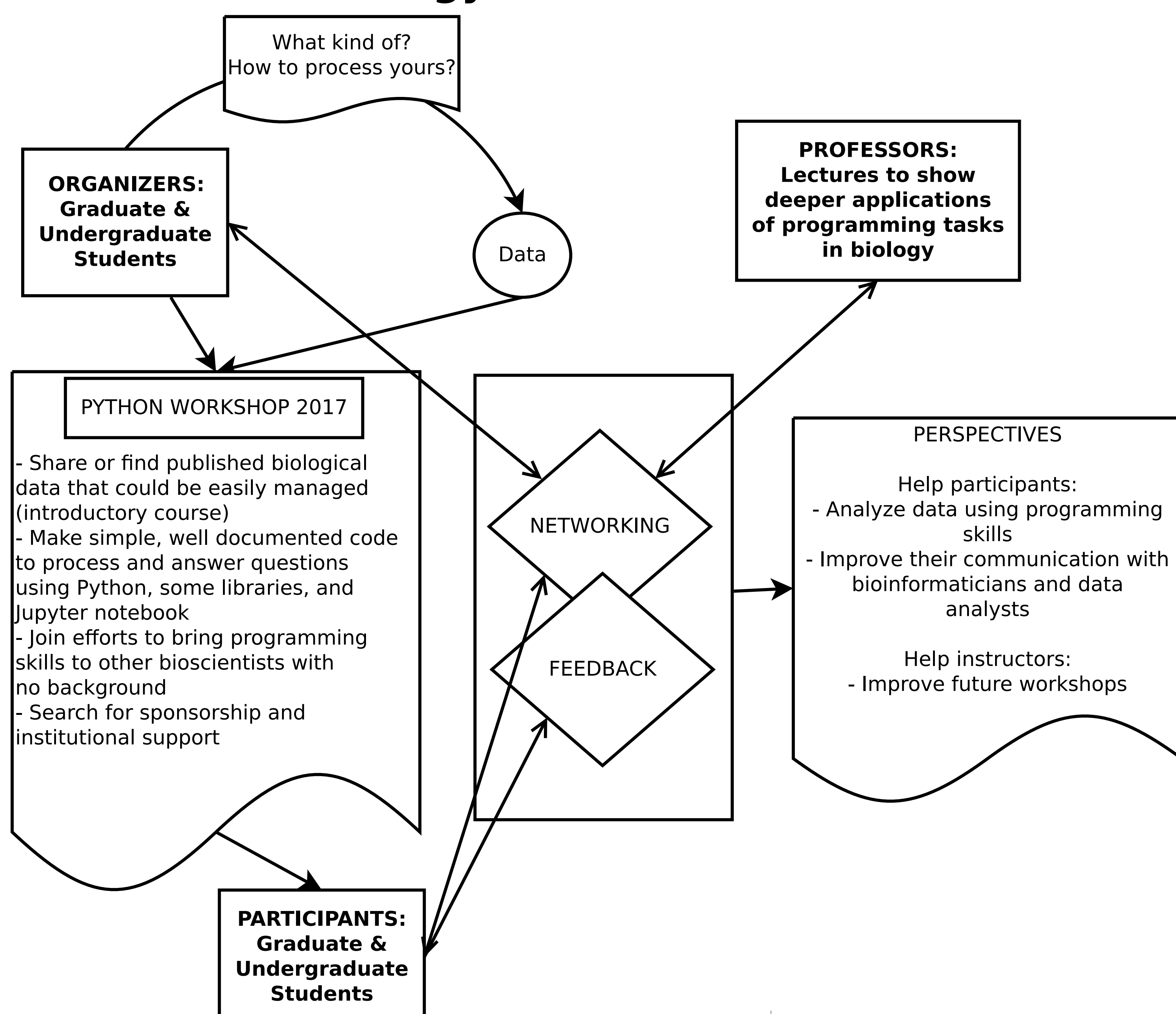
\* renatoacsantos@gmail.com

## Introduction

In Brazil, there is increasing trend in the number of scientific projects involving “Omics” analyses. Omics is the global analysis of each class of biological components, such as all proteins from an organism. These approaches require bioscientists with a little knowledge of algorithms and computer programming to use more advanced computing practices. Python has been widely employed in the development of software in Biology for different applications, including the analysis of the movement of atoms in DNA molecules and in the application of Machine Learning to extract their features. In addition to it, Python is also very useful for solving daily tasks in biology and easy for biologists to teach and learn compared to other programming languages like C++. In this case, short courses covering computer science skills are relevant to bioscientists not only because the course is a starting point for people interested in data science, but also increases their knowledge of basic concepts that improves their communication skills with bioinformaticians and data analysts. Therefore, we taught a workshop on python at the University of São Paulo on November 10 - 11, 2017.



## Educational strategy



## Institutional Support



## Sponsorship



## Overall results and outputs

Several languages are available for biologists to handle their data (see **Carey** et al. 2018), but we decided to use Python in our workshop inspired by a recent PLoS Computational Biology paper (**Ekmekci** et al. 2016) that introduces its syntax, with applications for Biological Sciences. The event took two days.

We consider this first workshop an amazing and enriching experience. Here, we highlight the main achievements. We:

- set up a team of enthusiast graduate and undergraduate students with prior experience in data handling (background on either biology, computer science, or both);
- selected participants whose research projects include analysis of biological data and that lack knowledge in computer science/programming skills;
- broadcasted on social networks (mainly on Facebook), in particular institutions in São Paulo State;
- established institutional support (coffee-break, room for practical sessions, gifts) ;
- developed practical sessions with Python programming and lectures from invited speakers;
- received feedback from participants through a Google form.

## Feedback

One learner wrote "the workshop was 100% practical allowing that the new concepts could be immediately related with real daily problems that I encounter when working with large scale biological data." Other learners provided similar positive review.

## Future work

We are organizing a second edition of the BWPBD (2018). Our expectations this year include: i) providing clearer biological problems and solutions with Python; ii) including a basic revision statistical concepts while introducing aspects of programming and data structures; iii) explore data visualization (making plots) during the practices; iv) covering broader aspects of the biological data used in the workshop, as well as including more experts working with data in the fields (transcriptomics and proteomics) in the organizing committee; v) improvement of the registration form to reflect the expectations of participants (to avoid frustration), as well as catch their previous knowledge on "omics" data (to avoid topic specific questions to allow more space for learning programming); vi) early general organization and production of a course material (ten months); vii) migration from Github to Gitlab (free, private project); viii) improvement of how the workshop is broadcasted (mailing lists, social media trying to reach students from other Brazilian states; ix) increasing the number of days (four days).

## Supporting material

We included the code (Jupyter notebooks are not available for the entire course), course materials (available in Portuguese), and this poster in a Github repository:

<https://github.com/SantosRAC/Python4BioDataBrazilianWorkshop>

## Acknowledgments

We thank the graduate student (PhD) Marian Schmidt (University of Michigan, MI, USA) for the opportunity to have an online workshop debriefing session through the Carpentries Mentoring Subcommittee.

## References

- Carey, Maureen A., and Jason A. Papin. "Ten simple rules for biologists learning to program." PLoS computational biology 14.1 (2018): e1005871.
- Ekmekci, Berk, Charles E. McAnany, and Cameron Mura. "An introduction to programming for bioscientists: A python-based primer." PLoS computational biology 12.6 (2016): e1004867.