# **Manuscript Title**

This manuscript (<u>permalink</u>) was automatically generated from <u>lubianat/codigobonito@c20d547</u> on September 27, 2020.

# **Authors**

- John Doe

Department of Something, University of Whatever  $\cdot$  Funded by Grant XXXXXXXX

- Jane Roe

Department of Something, University of Whatever; Department of Whatever, University of Something

# Reunião 24/09/2020

# A Quick Guide to Organizing Computational Biology Projects

[1]

## **Highlights**

Curriculum fails to address many of the day-to-day organizational challenges

Someone unfamiliar with your project should be able to look at your computer files and understand in detail what you did and why. Everything you do, you will probably have to do over again.

data directory for storing fixed data sets, a results directory for tracking computational experiments peformed on that data, a doc directory with one subdirectory per manuscript, and directories such as src for source code

In parallel with this chronological directory structure, I find it useful to maintain a chronologically organized lab notebook. In practice, I ask members of my research group to put their lab notebooks online, behind password protection if necessary.

driver script (I usually call this runall) that carries out the entire experiment automatically

Avoid editing intermediate files by hand. Doing so means that your script will only be semi-automatic Use relative pathnames to access other files within the same project

Even in a simple script, you should check for bogus parameters, invalid input, etc.

how much effort to put into software engineering ### Comments

It goes on to a lot of personal notes, which are okay, but far from statndard. GitHub logs work well.

# **Template**

# **Ten Simple Rules for Effective Computational Research**

[2]

# **Highlights**

the present guide is aimed specifically at the vast majority of scientific researchers: those without formal training in computer science. conduct a software literature review to ascertain what software is available.

The program code itself can be made more understandable by using meaningful variable names and formatting the code consistently. you should maintain a record of the "big picture" functionality (i.e., interconnectivity of components and input/output formats) keep a record of your work

all scientists could be better educated in design, so any investment will be rewarded

Simple tests that the software behaviour matches expectations are essential for ensuring robust results. Get the computer to run tests for you automatically and alert you to problems, using a suitable testing framework.

we advocate an open approach of sharing source code, data, and results as freely as possible.

prioritise in a way that suits you and your projects and career aspirations

#### **Comments**

Gostei mais que o outro. Lista de links ao fim é interessante.

# Ten Simple Rules for Taking Advantage of Git and GitHub

[3]

# **Highlights**

It provides developers and researchers with a dynamic and collaborative environment, often referred to as a social coding platform, that supports peer review, commenting, and discussion [7] GitHub hooks can be used to automate numerous tasks to help improve the overall quality of your project

GitHub now integrates with archiving services such as Zenodo and Figshare, enabling DOIs to be assigned to code repositories.

#### **Comments**

# **Template**

#### **Title**

[???]

# **Highlights**

#### **Comments**

# Ten Simple Rules for a Bioinformatics Journal Club

[<u>4</u>]

# **Highlights**

Don't be a "lonely bioinformatician"[1], create a journal club

We recommend you find a time that is typically free for the majority and suits both students and early career researchers, who may have time commitments during "normal" working hours. For us, that time was 8 A.M.

-> Claramente nao sao brasileiros.

a self-selection bias that we can live with.

ruled out formal presentations with PowerPoint slides.

for every paper, there will be someone who knows nothing about the topic and others who are nearexperts

open membership approach and encourage attendance from different cohorts. Mixing research students, coursework students, and early career researchers

We have discussed overcoming procrastination, writing more effectively, online learning, editing Wikipedia, student groups, Software Carpentry, document preparation systems for sci- ence, and even the PLOS Ten Simple Rules collection itself

Although we didn't originally envisage journal club as a platform for collaborations, we've found that meeting regularly to write amongst a group of supportive peers has led to a publication, with a style that suits the informal atmosphere of the club.

Keep the barrier to entry low.

#### **Comments**

o cdb eh um grupo de estudo

# **Template**

# Ten simple rules to increase computational skills among biologists with Code Clubs

[<u>5</u>]

# **Highlights**

participants need regular opportunities to engage in repeated deliberate practice that encourage them to practice their retrieval and application of the material

Initially, the Code Club was used to review code from trainee projects. Instead of a presentation, the presenter would project their code onto a screen and the participants would go through the code, stating the logic behind each line.

someone may have an R script with a repeating chunk ofcode. The chal-lenge for the session would be to convert the code chunk into a function to be called through- out the script to make it "DRY" (i.e.,

Don't Repeat Yourself [12]). The presenter leaves the session with several partial or working solutions to their problem, and the importance of writ- ing DRY code is reinforced.

We generally find that preparing a Code Club session takes similar effort to preparing a Journal Club presentation "think-pair-share" approach.

It is critical that the presenter and participants respect each other and that a designated indi-vidual (e.g., the lab director) enforces a code of conduct.

After the Code Club, try to incorporate that material into new code or refactor old code.

Set specific goals.

A typical schedule for Code Club is 10 minutes ofintroduction and instruction, 30 minutes ofpaired program- ming, 5 minutes to get groups to wrap up, and 10 minutes to report back to the group.

Using GitHub repositories for each Code Club can help make information, scripts, and data easily available to participants.

Perhaps a first Code Club could introduce using git and GitHub to engage in collaborative coding.

Central to the Code Club format is the use ofpaired programming.

Regardless ofhow partners are selected, consider asking the pairs to identify a navigator and a driver [16] ### Comments Table 1. Examples of successful Code Club topics. !! This is RICH.

### References

#### 1. A Quick Guide to Organizing Computational Biology Projects

William Stafford Noble

PLoS Computational Biology (2009-07-31) https://doi.org/fbbpkn

DOI: 10.1371/journal.pcbi.1000424 · PMID: 19649301 · PMCID: PMC2709440

#### 2. Ten Simple Rules for Effective Computational Research

James M. Osborne, Miguel O. Bernabeu, Maria Bruna, Ben Calderhead, Jonathan Cooper, Neil Dalchau, Sara-Jane Dunn, Alexander G. Fletcher, Robin Freeman, Derek Groen, ... Charlotte Deane *PLoS Computational Biology* (2014-03-27) <a href="https://doi.org/gdb9wq">https://doi.org/gdb9wq</a>

DOI: 10.1371/journal.pcbi.1003506 · PMID: 24675742 · PMCID: PMC3967918

#### 3. Ten Simple Rules for Taking Advantage of Git and GitHub

Yasset Perez-Riverol, Laurent Gatto, Rui Wang, Timo Sachsenberg, Julian Uszkoreit, Felipe da Veiga Leprevost, Christian Fufezan, Tobias Ternent, Stephen J. Eglen, Daniel S. Katz, ... Juan Antonio Vizcaíno

PLOS Computational Biology (2016-07-14) https://doi.org/gbrb39

DOI: 10.1371/journal.pcbi.1004947 · PMID: 27415786 · PMCID: PMC4945047

#### 4. Ten Simple Rules for a Bioinformatics Journal Club

Andrew Lonsdale, Jocelyn Sietsma Penington, Timothy Rice, Michael Walker, Harriet Dashnow *PLOS Computational Biology* (2016-01-28) <a href="https://doi.org/ghcwtj">https://doi.org/ghcwtj</a>

DOI: <u>10.1371/journal.pcbi.1004526</u> · PMID: <u>26820645</u> · PMCID: <u>PMC4731213</u>

#### 5. Ten simple rules to increase computational skills among biologists with Code Clubs

Ada K. Hagan, Nicholas A. Lesniak, Marcy J. Balunas, Lucas Bishop, William L. Close, Matthew D. Doherty, Amanda G. Elmore, Kaitlin J. Flynn, Geoffrey D. Hannigan, Charlie C. Koumpouras, ... Patrick D. Schloss

PLOS Computational Biology (2020-08-27) <a href="https://doi.org/gg92xw">https://doi.org/gg92xw</a>

DOI: <u>10.1371/journal.pcbi.1008119</u> · PMID: <u>32853198</u> · PMCID: <u>PMC7451508</u>