



Snakes on a Spaceship Tutorial Production Software and Version Control

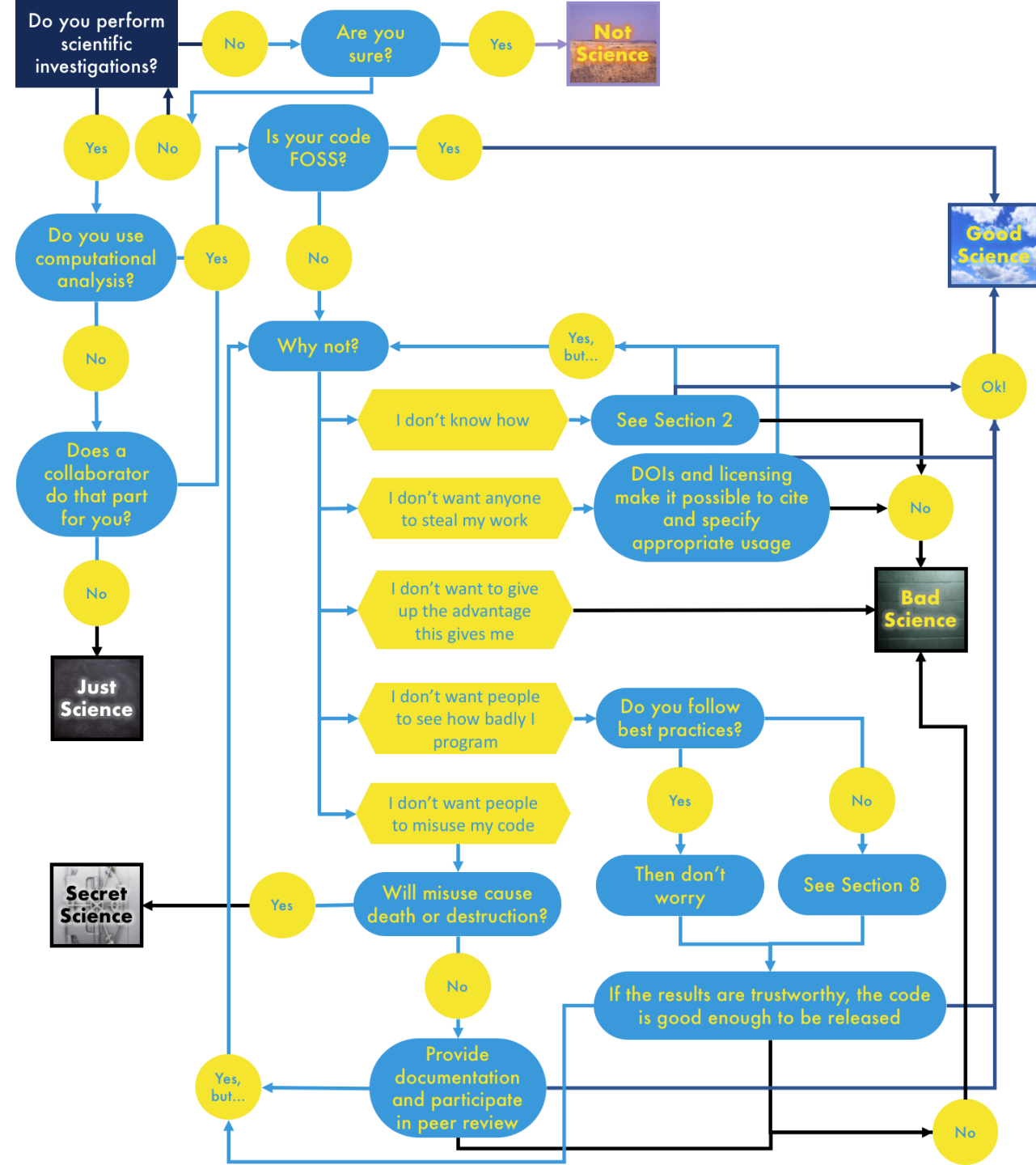


Leslie Lamarche, Angeline Burrell, Katherine Cariglia, Jessica Norrell



Good Code, Good Science

- Publishing scientific code is becoming more accepted in various communities
- DOIs and reviewers enforcing good citation practices have alleviated many fears of code being stolen
- Worries remain
 - Potential misuse
 - Shame about code quality
 - Lack of knowledge about how to go about publishing code



Moving Towards Production Code

“Paper” Code

- Needs to function only for the time it takes to prepare the manuscript
- You are probably working on it constantly
- Only the authors of the paper will ever look at the code

“Production” Code

- May need to function for long periods of time (ie, years)
- Development occurs sporadically as needed for new features and bugfixes
- Code will probably need to make sense to many people

When do you need production code?

- Running a model that many people might use
- Processing routinely collected data
- Toolkits that are widely used
- Anything you might want to queue up and run automatically
- Anything you might want to share widely enough that you would not have direct contact with all the users

Converting to Production Code

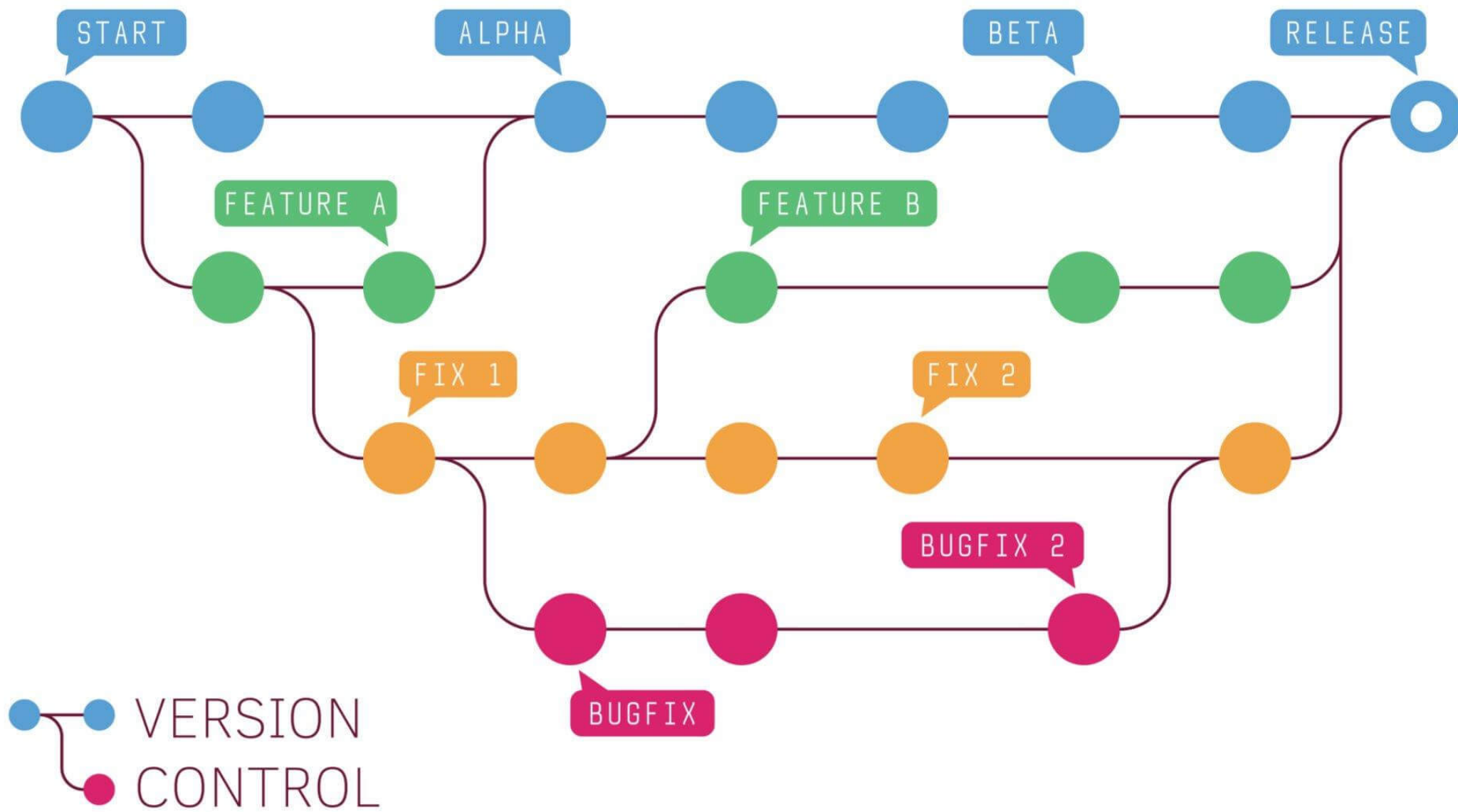
- Move away from jupyter notebooks
- Package code
- Avoid hardcoding (filepaths, configurable options, ect)
- No “manual tweaks” to source code
- Be aware of how the code will be used
- Document, document, document
- Best practices are critical
- **USE VERSION CONTROL**

What is version control?

- Software that keeps track of changes in your code
- Maintains a semipermanent record of changes
- “git” is very popular, but there are others (i.e. svn)

Why use version control?

- Maintains a semi-permanent record of all changes in code
- Can go back to any prior development state in the code
- Embolden you to make aggressive developmental changes
- Allows multiple features to be developed simultaneously



Git vs GitHub

- GitHub is an online platform for hosting git-controlled projects so multiple people can work on them
- You can use git locally WITHOUT doing anything with GitHub
- There are alternatives to GitHub (i.e. GitLab)

Workflow

Git

- Modify a file (or several)
- Stage the file(s) [git add]
- Commit changes [git commit]

Git with Remote/GitHub

- Modify a file (or several)
- Stage the file(s) [git add]
- Commit changes [git commit]
- Push changes [git commit]

Questions?