

BOSC 2020

---



# Building better bioinformatics tools with batteries included

---

Peter Georgeson, Anna Syme, Jessica Chung, Michael Milton, Harriet Dashnow,  
Andrew Lonsdale, Clare Sloggett, **Bernard Pope** \*

\* Victorian Health and Medical Research Fellow  
Melbourne Bioinformatics  
The University of Melbourne, Australia  
[bjpope@unimelb.edu.au](mailto:bjpope@unimelb.edu.au)

# Scientific software crisis?

Research Evaluation 24 (2015) pp. 454–470  
Advance Access published on 27 July 2015

doi:10.1093/reseval/rvv014

## Understanding the scientific software ecosystem and its impact: Current and future measures

James Howison<sup>1\*</sup>, Ewa Deelman<sup>2</sup>, Michael J. McLennan<sup>3</sup>,  
Rafael Ferreira da Silva<sup>2</sup> and James D. Herbsleb<sup>4</sup>

## How Do Scientists Develop and Use Scientific Software?

Jo Erskine Hannay  
Dept. of Software Engineering  
Simula Research Laboratory  
Dept. of Informatics, Univ. of Oslo  
johannay@simula.no

Carolyn MacLeod  
Dept. of Computer Science  
University of Toronto  
cmacleod@cs.utoronto.ca

Janice Singer  
Software Engineering Group  
National Research Council of Canada  
j.singer@nrc.ca

Hans Petter Langtangen  
Center for Biomedical Computing

**focus** developing scientific software.....

## Dealing with Risk in Scientific Software Development

Rebecca Sanders, *Queen's University*  
Diane Kelly, *Royal Military College of Canada*

OPEN ACCESS Freely available online

PLOS COMPUTATIONAL BIOLOGY

Perspectives

## Scientific Software Development Is Not an Oxymoron

Susan M. Baxter\*, Steven W. Day, Jacquelyn S. Fetrow, Stephen

BRIEFINGS IN BIOINFORMATICS. VOL 12. NO 3. 288–300  
Advance Access published on 28 January 2011

doi:10.1093/bib/bbq084

## Case studies in reproducibility

Torsten Hothorn and Friedrich Leisch

Submitted: 3rd September 2010; Received (in revised form): 29th November 2010

ICCGI 2013 : The Eighth International Multi-Conference on Computing in the Global Information Technology

## Lack of Software Engineering Practices in the Development of Bioinformatics Software

Dhawal Verma, Jon Gesell, Harvey Siy, Mansour Zand  
Department of Computer Science  
University of Nebraska at Omaha  
Omaha, Nebraska 68182  
Email: {dverma,jgesell,hsiy,zand}@unomaha.edu

## POLICYFORUM

COMPUTATIONAL SCIENCE

## Troubling Trends in Scientific Software Use

"Blind trust" is dangerous when choosing software to support research.

Lucas N. Joppa, <sup>1\*</sup> Greg McInerney, <sup>12</sup> Richard Harper, <sup>1</sup> Lara Salido, <sup>3</sup> Kenji Takeda, <sup>1</sup>  
Kenton O'Hara, <sup>1</sup> David Gavaghan, <sup>2</sup> Stephen Emmott<sup>1</sup>

News Feature | Published: 13 October 2010

## Computational science: ...Error

Zeeya Merali

Nature 467, 775–777(2010) | Cite this article

119 Accesses | 112 Citations | 216 Altmetric | Metrics

...why scientific programming does not compute.

## Toward effective software solutions for big biology

To the Editor:

Leading scientists tell us that the problem of large data and data integration, referred to as 'big data', is acute and hurting research. Recently, Snijder *et al.*<sup>1</sup> suggested a culture change in which scientists would aim

development that need to be resolved. Biologists are not formally trained for software engineering, so much of the bioinformatics software available today has been developed by PhD biologists in relative isolation on the back of funded experimental

# Help is at hand!

## 22nd Conference on Software Engineering Education and Training Software Engineering Education for Bioinformatics

Medha Umarji, Carolyn Seaman,  
A. Gunes Koru  
Department of Information Systems,  
University of Maryland Baltimore County  
Baltimore, MD 21250 USA  
{medha1, gkoru, cseaman}@umbc.edu

Hongfang Liu  
Department of Biostatistics  
Bioinformatics and Biomathematics  
Georgetown University Medical Center  
Washington, DC 20007 USA  
hl224@georgetown.edu

OPEN ACCESS Freely available online

### Community Page

## Best Practices for Scientific Computing

Greg Wilson<sup>1\*</sup>, D. A. Aruliah<sup>2</sup>, C. Titus Brown<sup>3</sup>, Neil P. Chue Hong<sup>4</sup>, Matt Davis<sup>5</sup>,  
Steven H. D. Haddock<sup>7</sup>, Kathryn D. Huff<sup>8</sup>, Ian M. Mitchell<sup>9</sup>, Mark D. Plumbley<sup>10</sup>,  
Ethan P. White<sup>12</sup>, Paul Wilson<sup>13</sup>

PLOS BIOLOGY

frontiers in  
GENETICS

OPINION ARTICLE  
published: 02 July 2014  
doi: 10.3389/fgene.2014.00199

## On best practices in the development of bioinformatics software

Felipe da Veiga Leprevost<sup>1,2\*</sup>, Valmir C. Barbosa<sup>3</sup>, Eduardo L. Francisco<sup>2</sup>, Yasset Perez-Riverol<sup>4</sup> and

Biophys Rev (2015) 7:343–352  
DOI 10.1007/s12551-015-0177-3

### REVIEW

## How to test bioinformatics software?

Amir Hossein Kamali<sup>1,2</sup> · Eleni Giannoulatou<sup>1,3</sup> · Tsong Yueh Chen<sup>4</sup> ·  
Michael A. Charleston<sup>5</sup> · Alistair L. McEwan<sup>2</sup> · Joshua W. K. Ho<sup>1,3</sup>

PLOS COMPUTATIONAL BIOLOGY

### EDITORIAL

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

Yasset Perez-Riverol<sup>1\*</sup>, Laurent Gatto<sup>2</sup>, Rui Wang<sup>1</sup>, Timo Sachsenberg<sup>3</sup>,  
Julian Uszkoreit<sup>4</sup>, Felipe da Veiga Leprevost<sup>5</sup>, Christian Fufezan<sup>6</sup>, Tobias Ternent<sup>1</sup>,  
Stephen J. Eglén<sup>7</sup>, Daniel S. Katz<sup>8</sup>, Tom J. Pollard<sup>9</sup>, Alexander Kononov<sup>10</sup>, Robert  
M. Flight<sup>11</sup>, Kai Blin<sup>12</sup>, Juan Antonio Vizcaino<sup>1\*</sup>

and Protein Engineering, Carlos Chagas Institute - Fiocruz, Curitiba, Brazil  
Curitiba, Brazil  
Computer Science Program, COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil  
European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge, UK  
leprevost.com.br

## Engineering bioinformatics: building reliability, performance and productivity into bioinformatics software

PLOS COMPUTATIONAL BIOLOGY

### EDITORIAL

## Ten simple rules for biologists learning to program

Maureen A. Carey<sup>1</sup>, Jason A. Papin<sup>2\*</sup>

<sup>1</sup> Department of Microbiology, Immunology, and Cancer Biology, University of Virginia School of Medicine, Charlottesville, Virginia, United States of America, <sup>2</sup> Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, United States of America

### EDITORIAL

## Ten simple rules for documenting scientific software

Benjamin D. Lee<sup>1\*</sup>

School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, United States of America

COMPUTATIONAL BIOLOGY

### EDITORIAL

## Ten simple rules for making research software more robust

Morgan Taschuk<sup>1\*</sup>, Greg Wilson<sup>2\*</sup>

<sup>1</sup> Genome Sequence Informatics, Ontario Institute for Cancer Research, Toronto, Ontario, Canada  
<sup>2</sup> Software Carpentry Foundation, Austin, Texas, United States of America

PLOS COMPUTATIONAL BIOLOGY

### EDITORIAL

## Ten Simple Rules for Developing Usable Software in Computational Biology

Markus List<sup>1\*</sup>, Peter Ebert<sup>1,2\*</sup>, Felipe Albrecht<sup>1,2</sup>

<sup>1</sup> Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarland Informatics Campus, Saarbrücken, Germany, <sup>2</sup> Graduate School of Computer Science, Saarland Informatics Campus, Saarbrücken, Germany

Seemann GigaScience 2013, 2:15  
http://www.gigasciencejournal.com/content/2/1/15

GIGA SCIENCE

### COMMENTARY

### Open Access

## Ten recommendations for creating usable bioinformatics command line software

Torsten Seemann<sup>1,2</sup>



Briefings in Bioinformatics, 19(4), 2018, 693–699

doi: 10.1093/bib/bbw134  
Advance Access Publication Date: 14 January 2017  
Paper

## Top considerations for creating bioinformatics software documentation

Mehran Karimzadeh and Michael M. Hoffman

# Help is at hand!

There are lots of recommendations in the literature about how to address this problem.

But they are spread over many papers, and only partially overlap.

Advice is useful, but action is better.

22nd Conference on Software Engineering Education and Training  
Software Engineering Education for Bioinformatics

OPEN ACCESS Freely available online

PLOS BIOLOGY

Community Page

Best Practices for Scientific Computing

frontiers in  
GENETICS

OPINION ARTICLE  
published: 02 July 2014  
doi: 10.3389/fgene.2014.00199

Medha Umarji, Carolyn Seaman,  
A. Gunes Koru  
Department of Information  
University of Maryland Balti  
Baltimore, MD 21250  
{medhal, gkoru, cseaman}@

Hongfang Liu

Biophys Rev (2015) 7:343–352  
DOI 10.1007/s12551-015-0177-3

REVIEW

How to test bioinform

Amir Hossein Kamali<sup>1,2</sup> • Eleni Gianno  
Michael A. Charleston<sup>5</sup> • Alistair L. M

EDITORIAL

Ten simple  
software

Benjamin D. Lee  
School of Engineeri  
of America

of bioinformatics

isco<sup>2</sup>, Yasset Perez-Riverol<sup>4</sup> and

Rio de Janeiro, Brazil  
nbridge, UK

building reliability,  
ty into bioinformatics

s for biologists learning to

apin<sup>2\*</sup>

inology, and Cancer Biology, University of Virginia School of Medicine,  
is of America, 2 Department of Biomedical Engineering, University of  
ted States of America

(GIGA)<sup>n</sup>  
SCIENCE

Open Access

creating usable  
software

Top considerations for creating  
bioinformatics software documentation

Mehran Karimzadeh and Michael M. Hoffman

EDITORIAL

Ten Simple Rules for Developing Usable  
Software in Computational Biology

Markus List<sup>1,2\*</sup>, Peter Ebert<sup>1,2\*</sup>, Felipe Albrecht<sup>1,2</sup>

1 Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Saarland Informatics  
Campus, Saarbrücken, Germany, 2 Graduate School of Computer Science, Saarland Informatics Campus,  
Saarbrücken, Germany

# What is Bionitio?

---

- A **tool** for starting new bioinformatics software projects following recommended best practices
- Supports 12 different programming languages

# What is Bionitio?

---

- One command starts a new bioinformatic project:

```
bionitio-boot.sh -n skynet -i python
```

# What is Bionitio?

- One command starts a new bioinformatic project:

```
bionitio-boot.sh -n skynet -i python
```

new project name

# What is Bionitio?

- One command starts a new bioinformatic project:

```
bionitio-boot.sh -n skynet -i python
```

programming  
language



# What is Bionitio?

- One command starts a new bioinformatic project:

```
bionitio-boot.sh -n skynet -i python
```

choices: C, C++, C# , Clojure, Java, Javascript,  
Haskell, Perl 5, Python 3, R, Ruby, Rust

programming  
language

# What is Bionitio?

---

- The result is a functional software artefact:
  1. working example of best practices
  2. template for making new tools

# What does the new project do?

```
$ skynet -h
```

```
usage: skynet [-h] [--minlen N] [--version] [--log LOG_FILE]
           [FASTA_FILE [FASTA_FILE ...]]
```

Read one or more FASTA files, compute simple stats for each file

positional arguments:

FASTA_FILE	Input FASTA files
------------	-------------------

optional arguments:

-h, --help	show this help message and exit
--minlen N	Minimum length sequence to include in stats (default 0)
--version	show program's version number and exit
--log LOG_FILE	record program progress in LOG_FILE

# What does the new project do?

```
$ skynet file1.fasta file2.fasta
```

FILENAME	NUMSEQ	TOTAL	MIN	AVG	MAX
file1.fasta	1	237	237	237	237
file2.fasta	2	357	120	178	237

# Key features built into the template

---

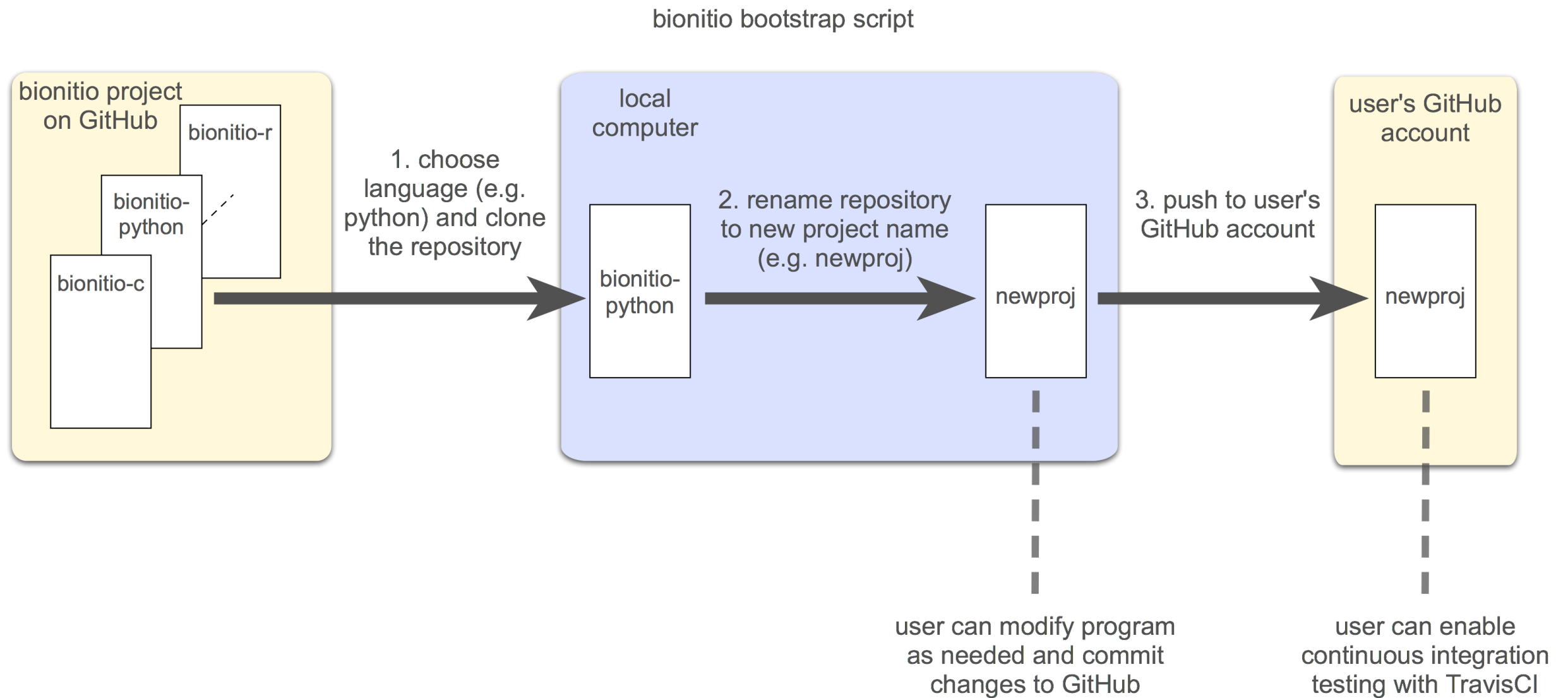
- Command-line argument parsing
- Progress logging
- Defined exit status values
- Test suite
- Version number
- Software packaging and a Docker container
- Standard open source license
- Documentation
- Revision control with Git and (optionally) GitHub
- Wrapper for the Common Workflow Language

# Create a new project

---

- `bionitio-boot.sh` creates new projects
- It can be run from:
  - Docker
  - from GitHub, via curl, or downloaded manually

# Create a new project



# Create a new project

---

```
$ URL=https://git.io/bionitio-boot
```

```
$ curl -sSfL $URL | bash -s -- -i python -n skynet
```



# Contents of new project directory

```
$ tree -a skynet
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

# Contents of new project directory

```
$ tree -a skynet
```

```
skynet
```

```
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

Git repository (contents abbreviated for the sake of this slide)

# Contents of new project directory

```
$ tree -a skynet
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   ├── unit-test.sh
│   └── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

Files for continuous  
integration testing with  
Travis CI

# Contents of new project directory

```
$ tree -a skynet
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

Support for a Docker  
container

# Contents of new project directory

```
$ tree -a skynet
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

A standard open source license. Defaults to MIT, but you can choose a different option with the -c flag.

# Contents of new project directory

```
$ tree -a skynet
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

User documentation explaining how to install and use the program.

# Contents of new project directory

```
$ tree -a skynet
```

```
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

← Testing scripts and data



# Contents of new project directory

```
$ tree -a skynet
```

```
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

Python-specific  
installation scripts and  
data.



# Contents of new project directory

```
$ tree -a skynet
```

```
skynet
```

```
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

Source code

# Contents of new project directory

```
$ tree -a skynet
skynet
├── .git
│   ├── COMMIT_EDITMSG
│   ├── HEAD
│   └── ... abbreviated ...
├── .travis
│   ├── install-dependencies.sh
│   └── unit-test.sh
├── .travis.yml
├── Dockerfile
├── LICENSE
├── README.md
├── functional_tests
│   ├── skynet-test.sh
│   └── test_data
│       ├── empty_file
│       ├── empty_file.expected
│       ├── no_header
│       ├── one_sequence.fasta
│       ├── one_sequence.fasta.expected
│       ├── single_greater_than.fasta
│       ├── two_sequence.fasta
│       ├── two_sequence.fasta.expected
│       ├── two_sequence.fasta.minlen_1000.expected
│       ├── two_sequence.fasta.minlen_200.expected
│       └── two_sequence.fasta.minlen_200.stdin.expected
├── requirements-dev.txt
├── setup.py
├── skynet
│   ├── .gitignore
│   ├── __init__.py
│   ├── skynet.py
│   └── skynet_test.py
└── skynet.cwl
```

Support for the  
Common Workflow  
Language

# Automatically create a GitHub remote

---

```
$ URL=https://git.io/bionitio-boot  
$ curl -sSfL $URL | bash -s -- -i python -n skynet \  
    -g cyberdyne \  
    -a 'Miles Bennett Dyson' \  
    -e 'miles@cyberdyne.com'
```

# Automatically create a GitHub remote

Same as before

```
$ URL=https://git.io/bionitio-boot  
$ curl -sSfL $URL | bash -s -- -i python -n skynet \  
-g cyberdyne \  
-a 'Miles Bennett Dyson' \  
-e 'miles@cyberdyne.com'
```

# Automatically create a GitHub remote

```
$ URL=https://git.io/bionitio-boot  
$ curl -sSfL $URL | bash -s -- -i python -n skynet \  
-g cyberdyne \  
-a 'Miles Bennett Dyson' \  
-e 'miles@cyberdyne.com'
```

GitHub  
username, author  
name, email  
address

# Automatically create a GitHub remote

The screenshot shows the GitHub repository page for 'bjpop / skynet'. The repository has 1 commit, 1 branch, 0 packages, 0 releases, and 0 contributors. The 'LICENSE' file is highlighted in the file list. The 'README.md' file is also visible, showing the 'Overview' section with a description of the program and the 'Licence' section with the MIT License.

Repository: bjpop / skynet

Unwatch 1 Star 0 Fork 0

Code Issues 0 Pull requests 0 Actions Projects 0 Wiki Security Insights Settings

No description, website, or topics provided. Edit

Manage topics

1 commit 1 branch 0 packages 0 releases 0 contributors MIT

Branch: master New pull request Create new file Upload files Find file Clone or download

BIONITIO Initial commit of skynet; starting from bionitio (python) Latest commit 9362c08 2 days ago

File	Commit	Time
.travis	Initial commit of skynet; starting from bionitio (python)	2 days ago
functional_tests	Initial commit of skynet; starting from bionitio (python)	2 days ago
skynet	Initial commit of skynet; starting from bionitio (python)	2 days ago
.travis.yml	Initial commit of skynet; starting from bionitio (python)	2 days ago
Dockerfile	Initial commit of skynet; starting from bionitio (python)	2 days ago
LICENSE	Initial commit of skynet; starting from bionitio (python)	2 days ago
README.md	Initial commit of skynet; starting from bionitio (python)	2 days ago
requirements-dev.txt	Initial commit of skynet; starting from bionitio (python)	2 days ago
setup.py	Initial commit of skynet; starting from bionitio (python)	2 days ago
skynet.cwl	Initial commit of skynet; starting from bionitio (python)	2 days ago

README.md

build unknown

## Overview

This program reads one or more input FASTA files. For each file it computes a variety of statistics, and then prints a summary of the statistics as output.

In the examples below, `$` indicates the command line prompt.

## Licence

This program is released as open source software under the terms of [MIT License](#).

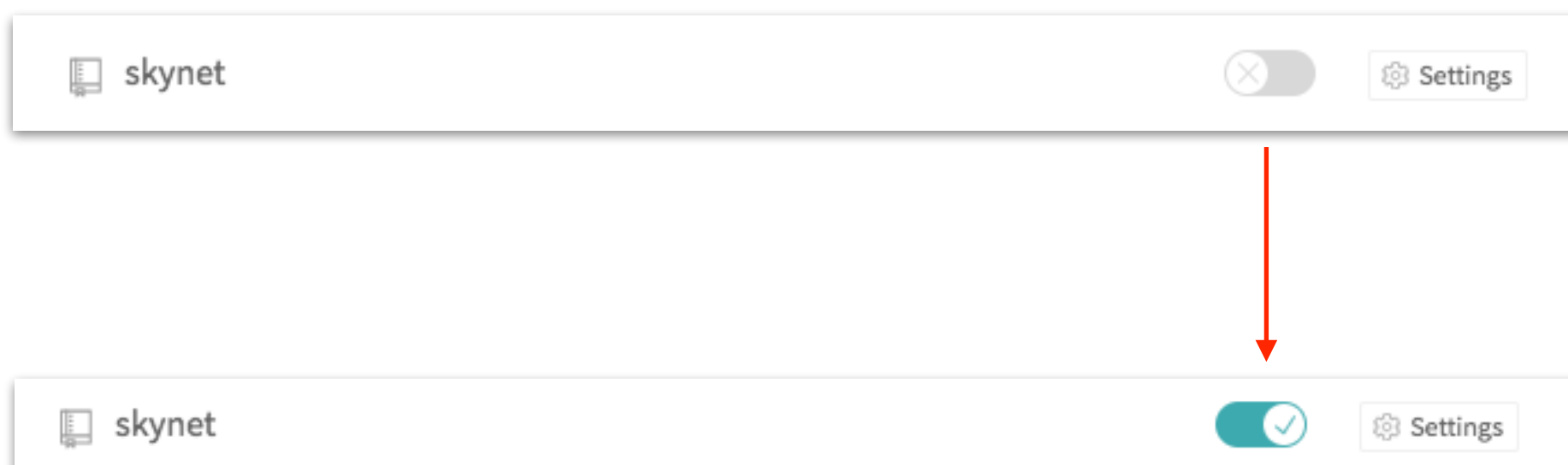
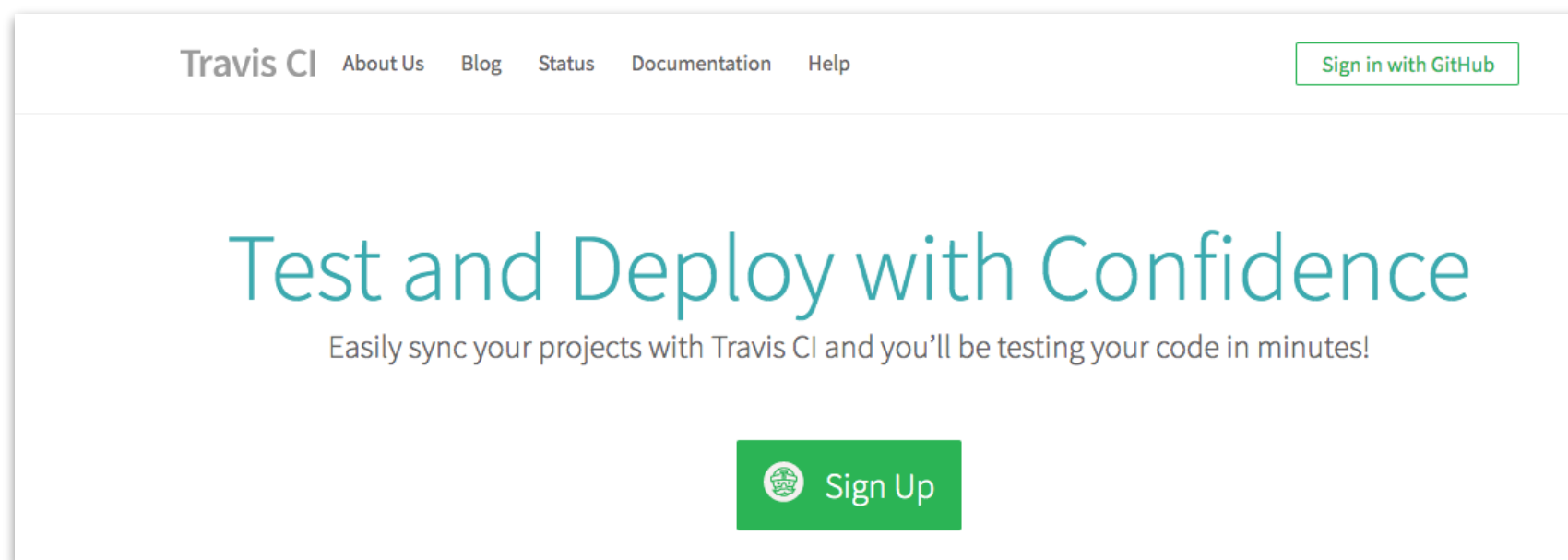
## Installing

You can install skynet directly from the source code or build and run it from within Docker container.

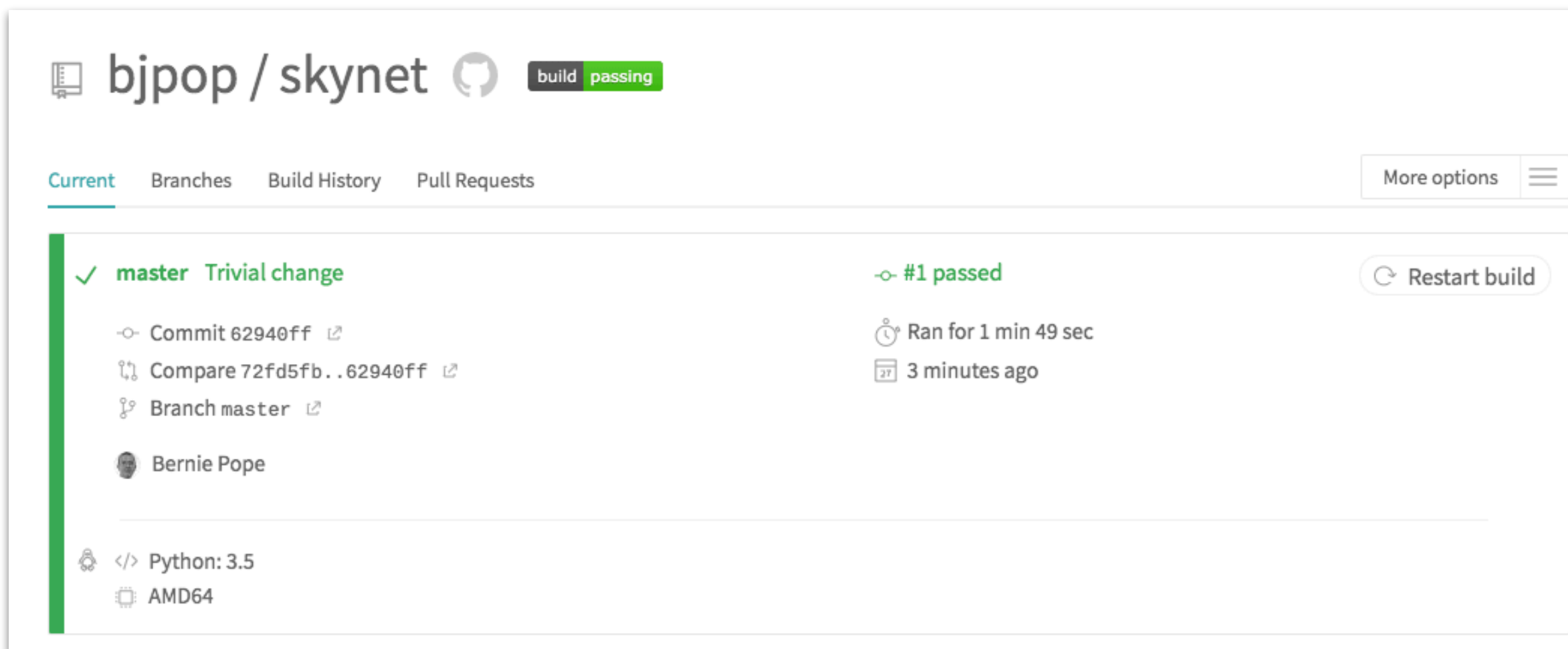
license is easy to find



README.md provides user documentation

# Integration with Travis CI



# Integration with Travis CI






 **bjpop / skynet**  build passing



Current Branches Build History Pull Requests More options

✓ **master** Trivial change #1 passed Restart build

Commit 62940ff [↗](#)  
Compare 72fd5fb...62940ff [↗](#)  
Branch master [↗](#)

 Bernie Pope

 </> Python: 3.5  
 AMD64

 Ran for 1 min 49 sec  
 3 minutes ago



# Modify the program to suit your own needs

---

- Once a new project has been created we expect users to modify it to suit their own purposes.
- This might involve rewriting large parts of the code, documentation and test suite.
- However, this can be done incrementally.
- The important point is that they are starting from a working project that already has batteries included.
- It is much easier *and faster* to modify an existing project than to start from scratch.

# Conclusion

---

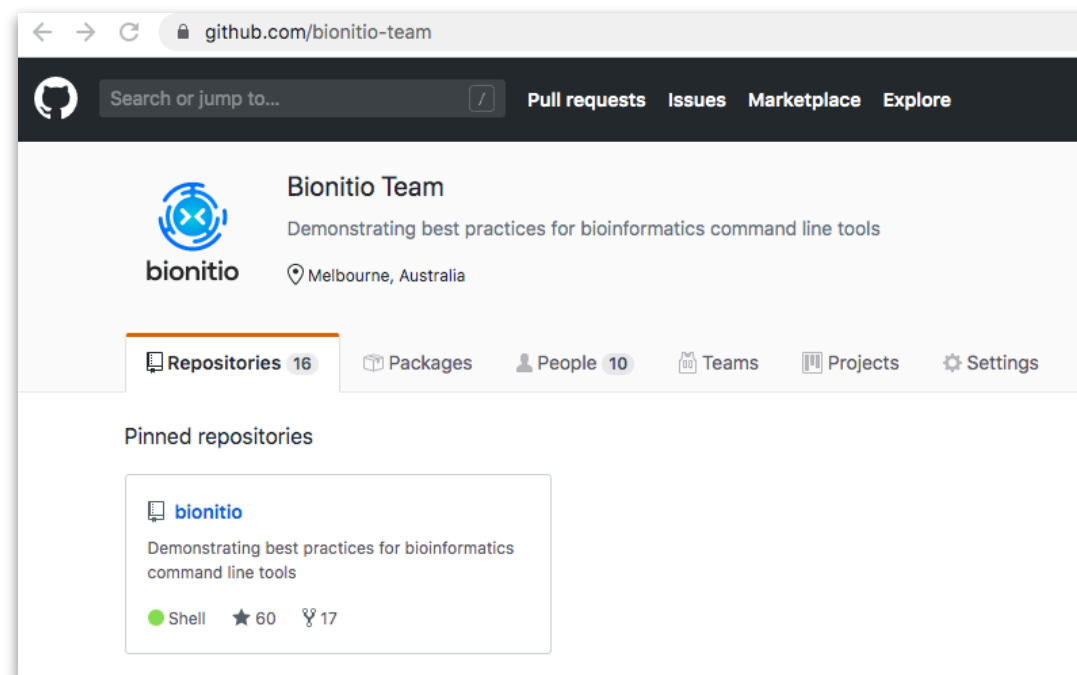
- Bionitio takes a pragmatic approach to solving the scientific software crisis.
- We want to help bioinformaticians to develop good habits early on, and to use them all the time - even for small scripts.
- Bionitio both *illustrates* good practices and makes them *easy to use*.

# Conclusion

---

- We have also found that Bionitio is a good tool for training beginner and intermediate bioinformatics software developers

# More information



- <https://github.com/bionitio-team/bionitio>

# Acknowledgements

---

## **Bionitio authors**

- Peter Georgeson
- Anna Syme
- Clare Sloggett
- Jessica Chung
- Harriet Dashnow
- Michael Milton
- Andrew Lonsdale
- David Powell
- Torsten Seemann

**Melbourne Bioinformatics**

**Department of Health and Human  
Services, Victoria**