seqeralabs

Phil Ewels
phil@seqera.io

Adventures in Open Science

NextGenBUG 2022

Open Science

Open Science

Open Source

Open Data

Open Contributions

Open Community

Open Core

Open Source

Open Source

Labrador

Search

Search

Create New Project

Log In / Register

Filters

TEXT FILTER

PROJECT NAME

A B C D E F G H I J K L M N

O P Q R S T U V W X Y Z 0-9

SPECIES

Mus musculus

Homo sapiens

Harpegnathos saltator

Mouse

Saccharomyces cerevisiae

Camponotus floridanus

Polistes canadensis

abidopsis thalia

mellifera

Labrador Dataset Browser A database of datasets processed by the BI Bioinformatics group.

You can use labrador to find and download processed data or request new datasets. Projects are annotated with how they were processed.

Labrador Dataset Browser

Video Tutorial

You can use the table below to browse the projects and datasets. You can filter the visible data using the tools on the left. If you're looking for something really specific, try the search bar at the top of the page.

Key: [x] Processing Complete [x] Currently Processing [x] Not Started [x] Directory not found

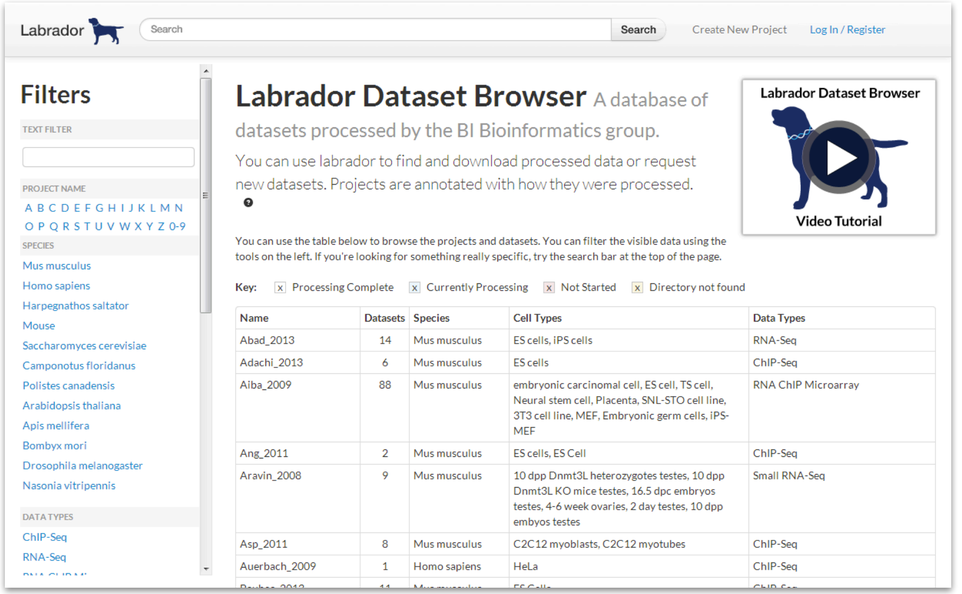
Name	Datasets	Species	Cell Types	Data Types
Abad_2013	14	Mus musculus	ES cells, iPS cells	RNA-Seq
Adachi_2013	6	Mus musculus	ES cells	ChIP-Seq
Aiba_2009	88	Mus musculus	embryonic carcinomal cell, ES cell, TS cell, stem cell, Placenta, SNL-STO cell line, e, MEF, Embryonic germ cells, iPS-	RNA ChIP Microarray

Open Source

2010 First release

2012 Moved to GitHub

2015 Stopped active development



Open Source

2010 First release

2012 Moved to GitHub

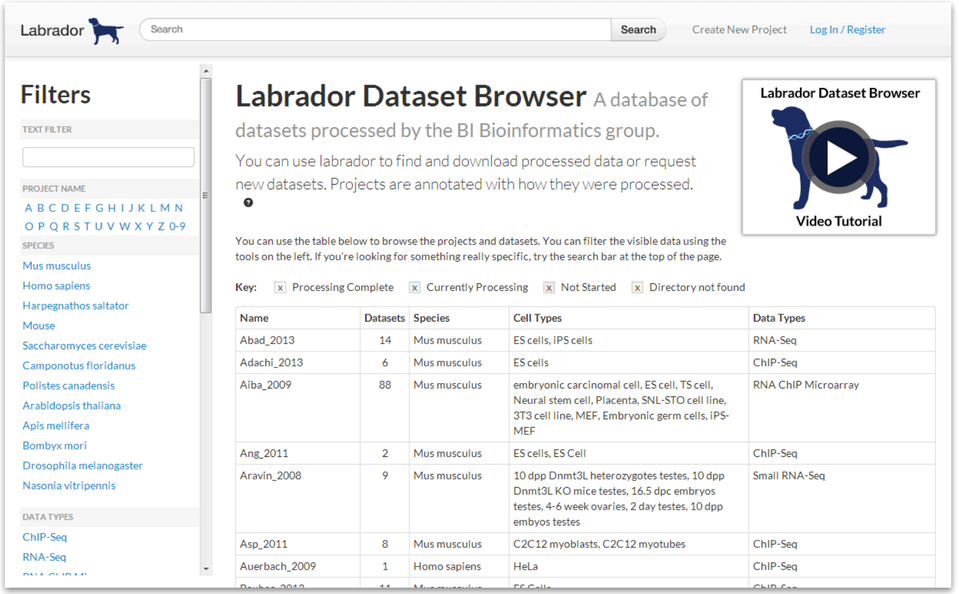
2015 Stopped active development

2017 Contributions from Russell Hamilton

2018 Fixes from Simon Andrews

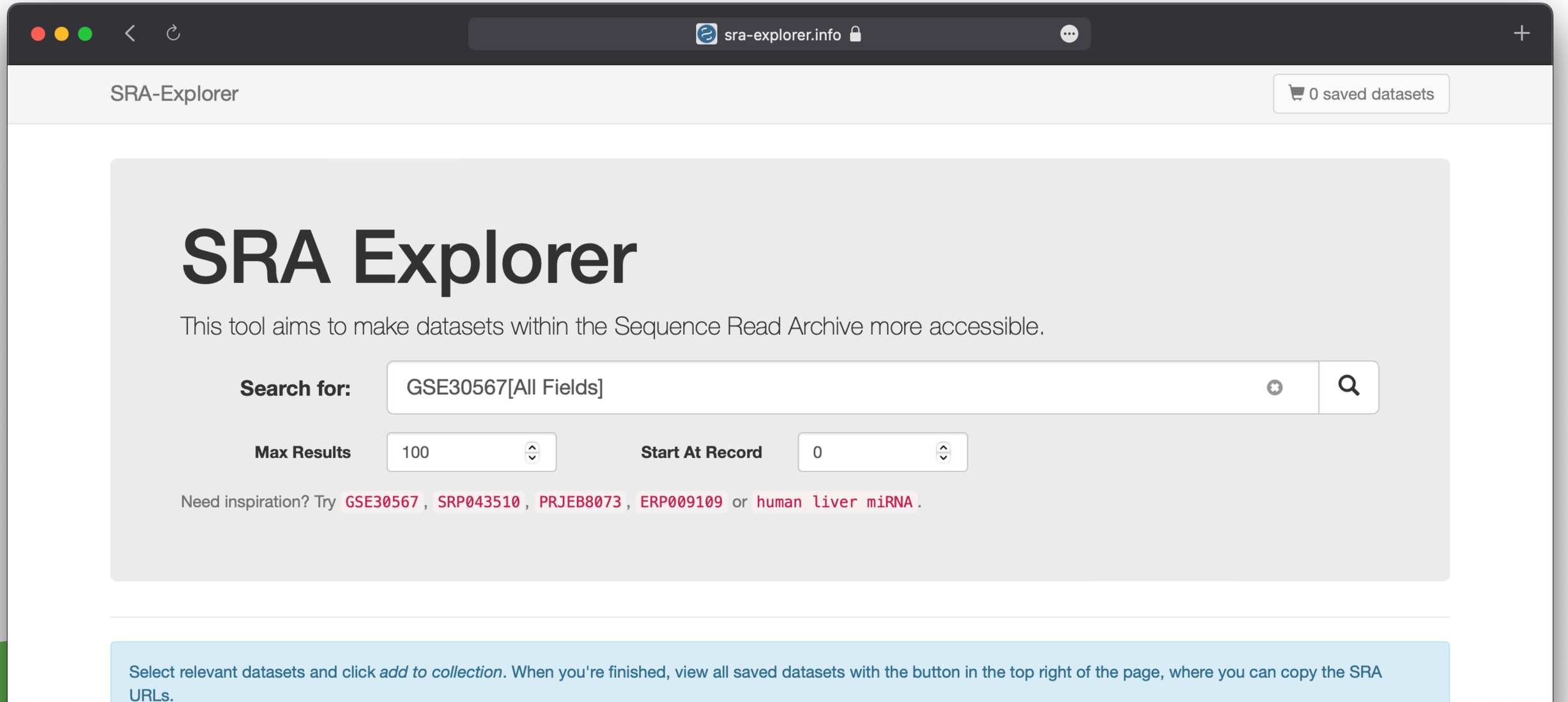
2021 Updates from Steven Wingett

2022 Security patches from Altos Labs

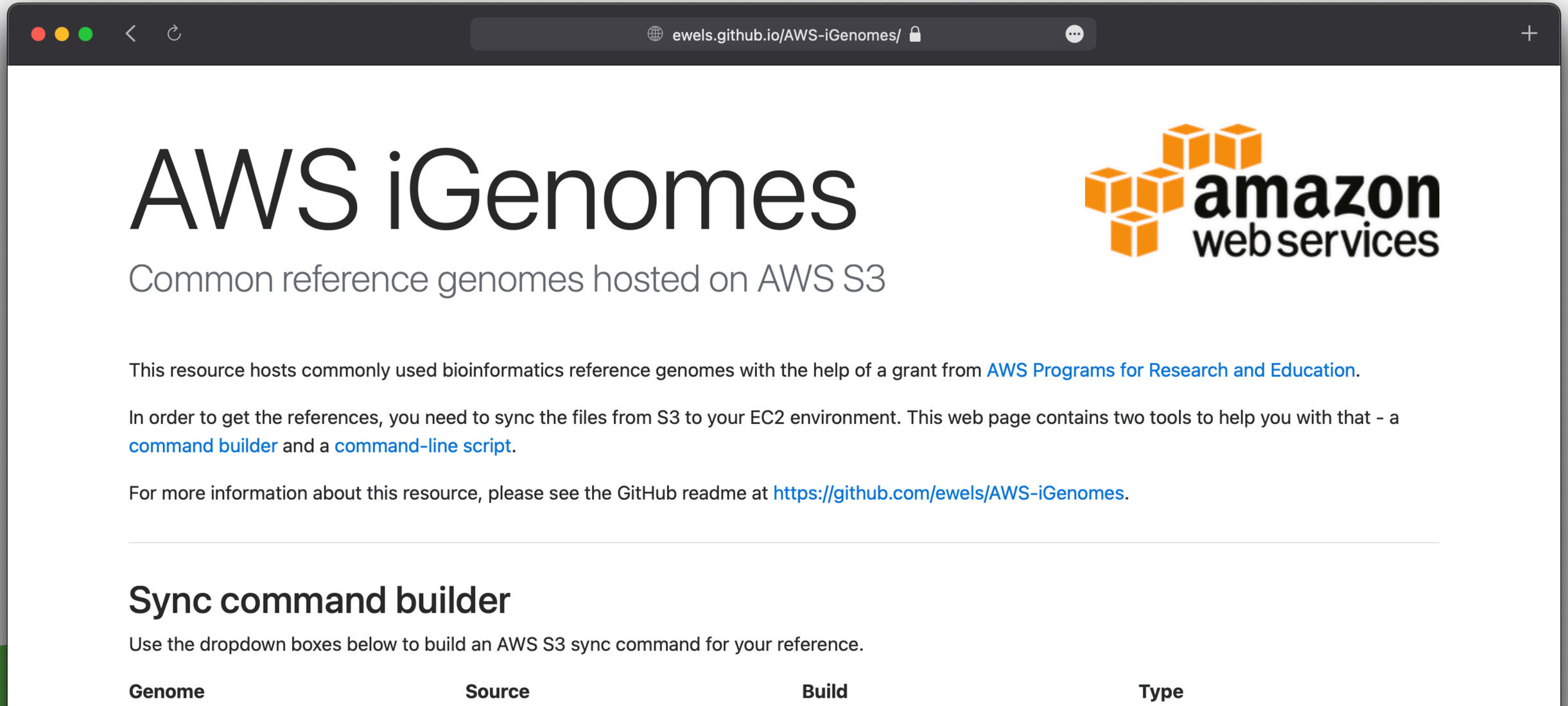


Open Data

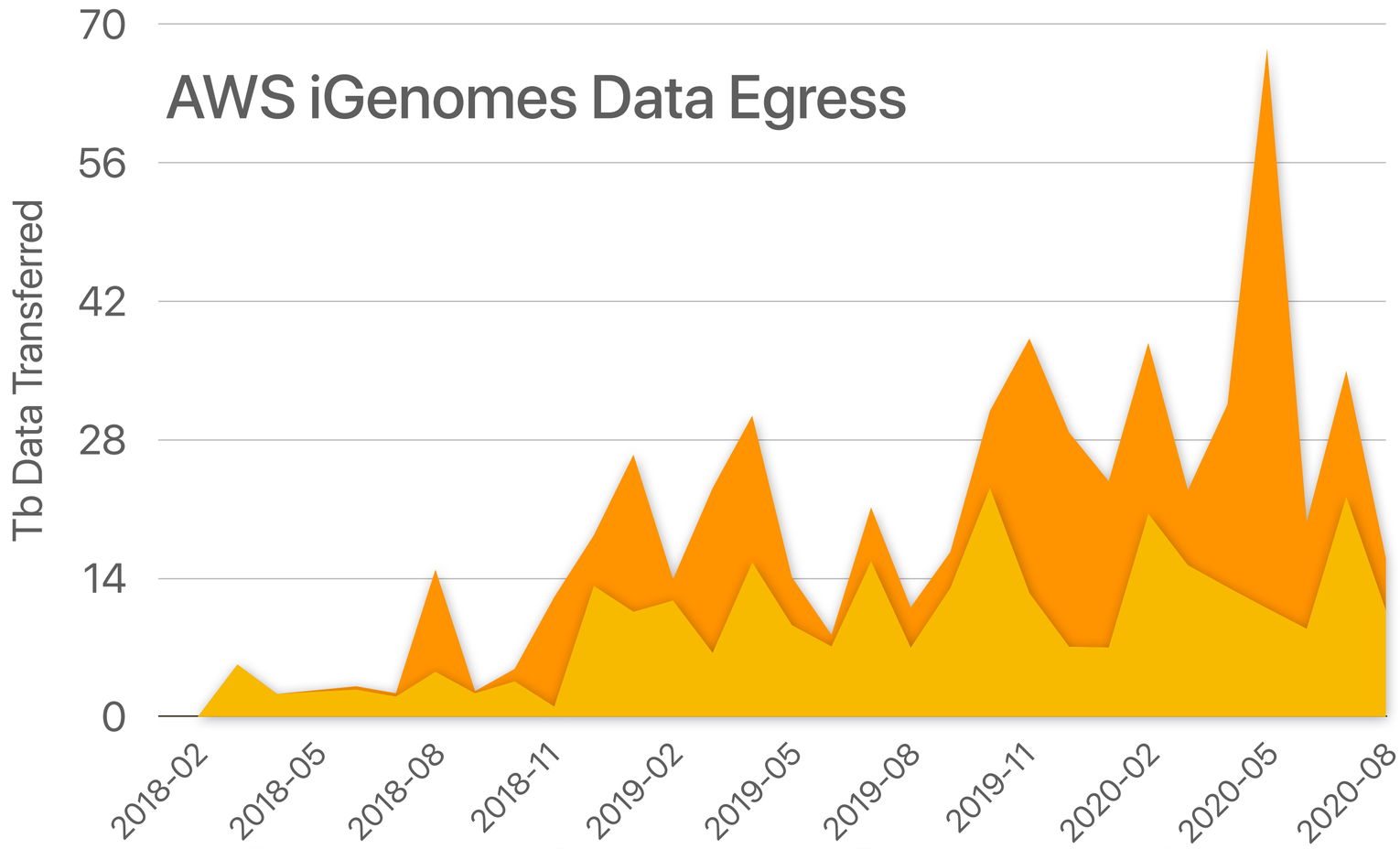
Open Data

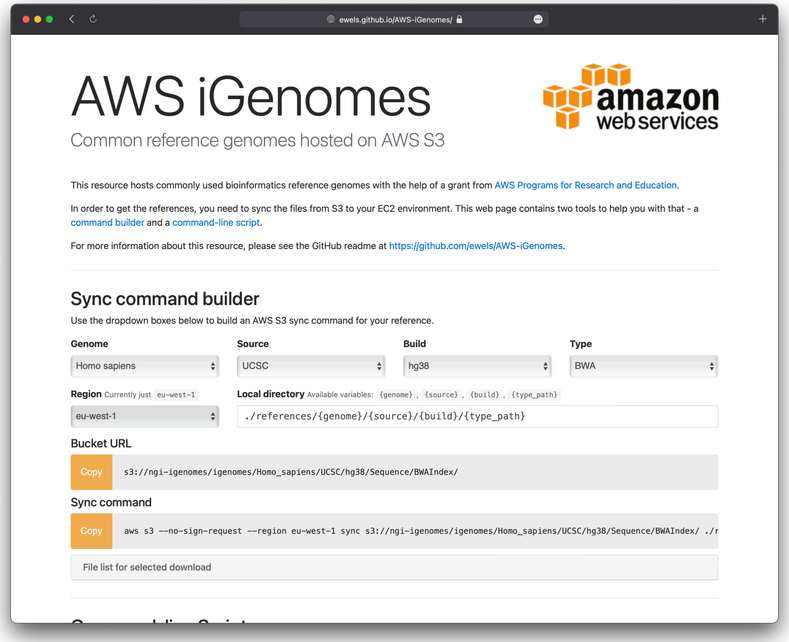

sra-explorer.info
SRA-Explorer
0 saved datasets
SRA Explorer
This tool aims to make datasets within the Sequence Read Archive more accessible.
Search for:
GSE30567[All Fields]
Max Results
100
Start At Record
0
Need inspiration? Try GSE30567, SRP043510, PRJEB8073, ERP009109 or human liver miRNA.
Select relevant datasets and click add to collection. When you're finished, view all saved datasets with the button in the top right of the page, where you can copy the SRA URLs.

Open Data


ewels.github.io/AWS-iGenomes/
AWS iGenomes
Common reference genomes hosted on AWS S3
amazon web services
This resource hosts commonly used bioinformatics reference genomes with the help of a grant from AWS Programs for Research and Education.
In order to get the references, you need to sync the files from S3 to your EC2 environment. This web page contains two tools to help you with that - a command builder and a command-line script.
For more information about this resource, please see the GitHub readme at https://github.com/ewels/AWS-iGenomes.
Sync command builder
Use the dropdown boxes below to build an AWS S3 sync command for your reference.
Genome
Source
Build
Type

Open Data





Total storage: ~5.5TB

Human STAR index: ~30GB

Tb within AWS
Tb to internet

(stacked plot)

Open Contributions

Open Contributions

Open Contributions

Packaged (easy to install)

Simple user interface

Usage documentation

Sensible defaults

Open Contributions

Packaged (easy to install)

Simple user interface

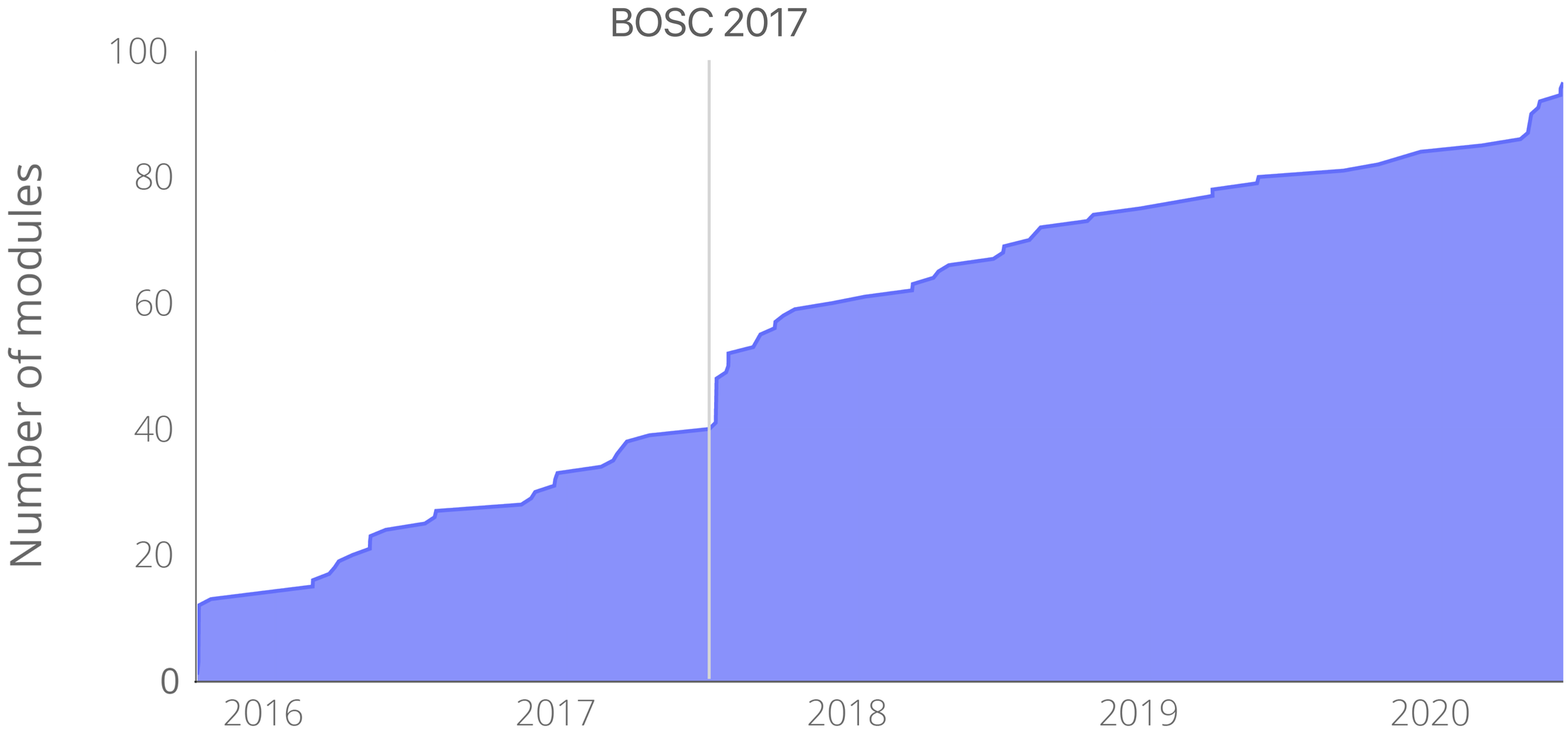
Usage documentation

Sensible defaults

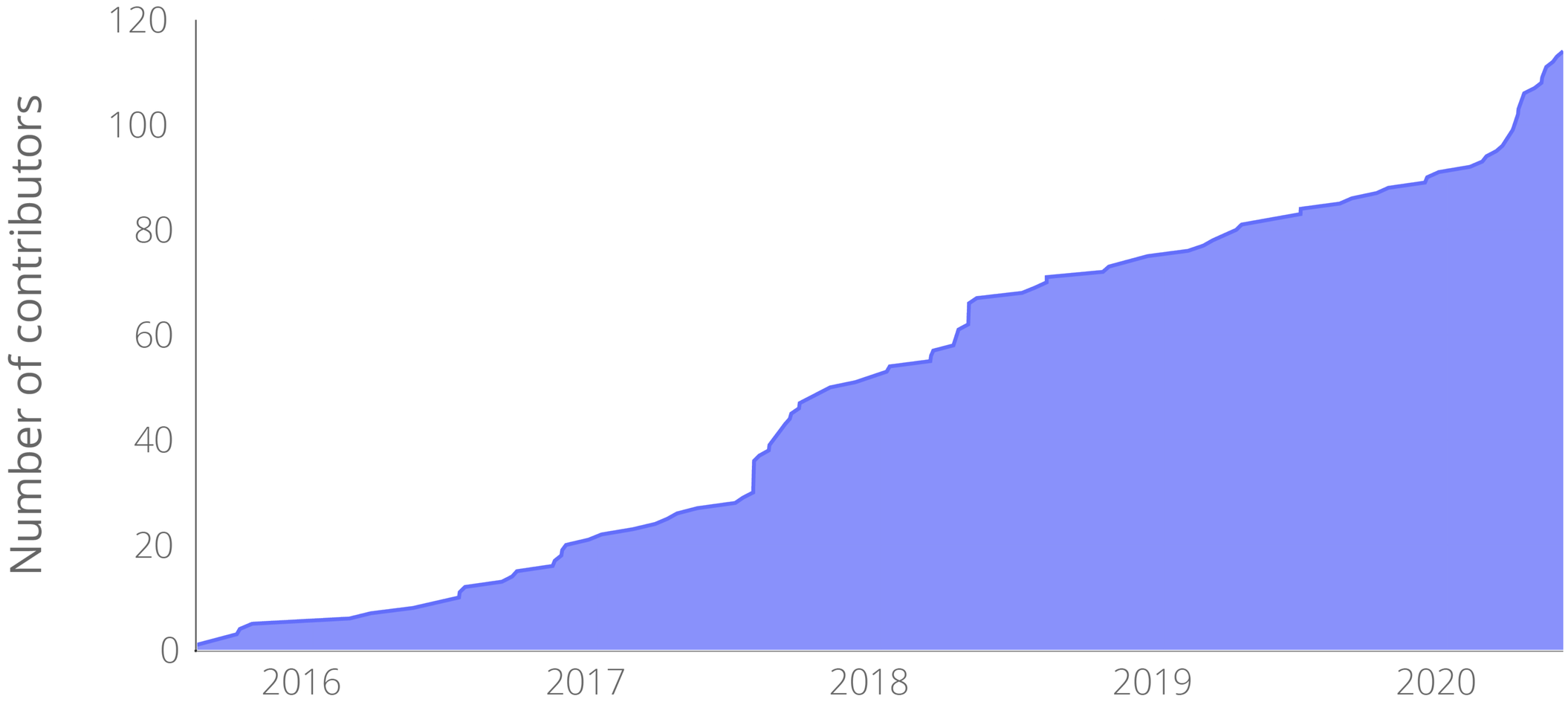
Extensibility

Developer documentation

Open Contributions



Open Contributions

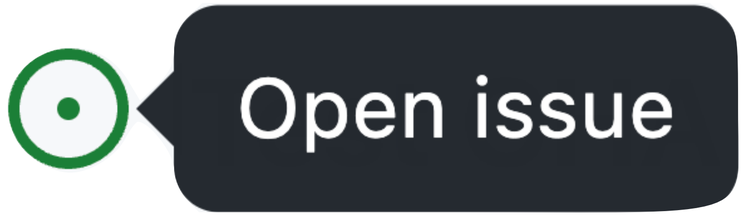


Open Contributions

MultiQC

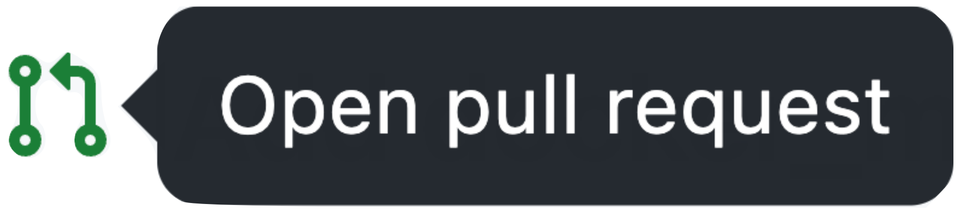
0.45

per day



0.23

per day



Open Community

Open Community

nf-core

nf-core

71

PIPELINES

https://nf-co.re

nf-core

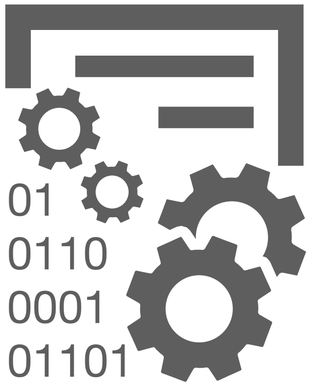
Running pipelines

Writing pipelines

Testing / automation

https://nf-co.re

nf-core

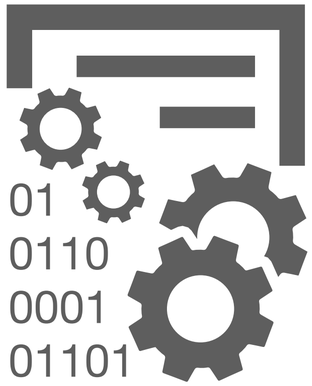


692

MODULES

https://nf-co.re

nf-core



692

MODULES

21

SUB-WORKFLOWS

https://nf-co.re

Open Community

Removal of institutional branding

Defined scope and guidelines

Generous access rights

Focus on developer tooling

Open Community

Removal of institutional branding

Defined scope and guidelines

Generous access rights

Focus on developer tooling

Automated testing

Enforced code review

nf-core

Open Community

nf-core

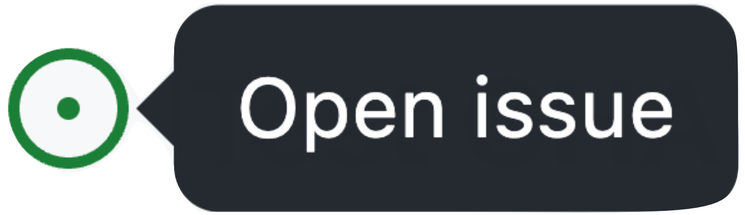
> 4000

Members on Slack

Open Community

nf-core

320



nf-core hackathon
(3 days)

298

Open Core

Open Core

Nextflow and nf-core are supported by:

Chan Zuckerberg Initiative

aws

Microsoft

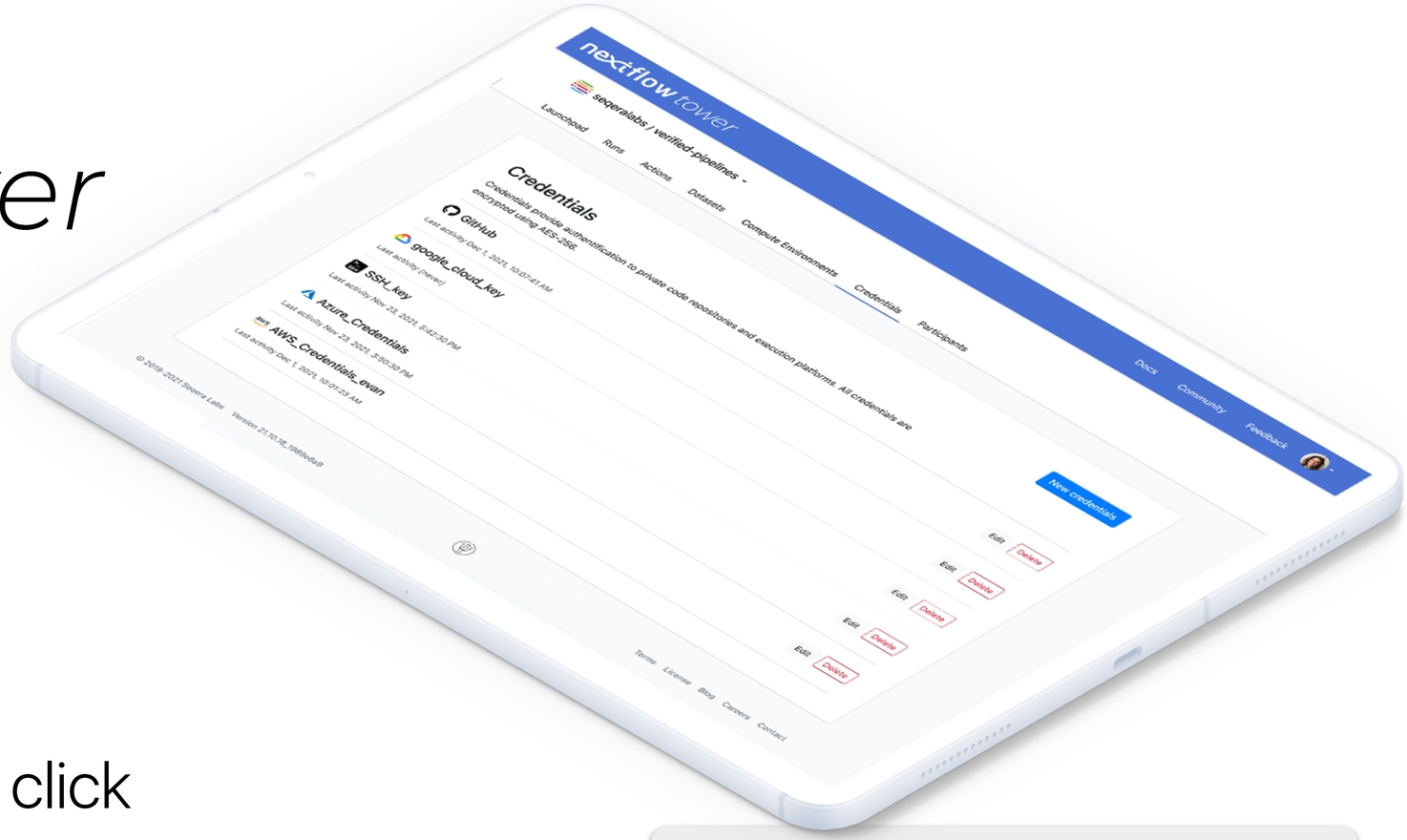
SciLifeLab

seqeralabs

Open Core

seqeralabs

nextflow *tower*



Intuitive launchpad interface

Launch, manage, and monitor

Share runs and work in teams

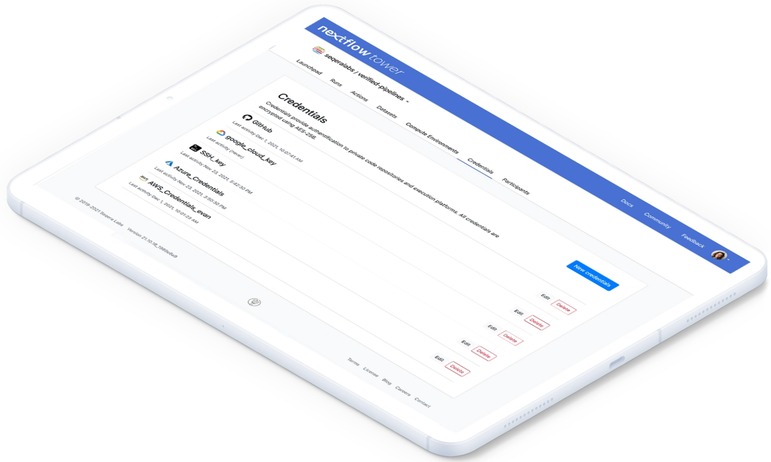
Create cloud infrastructure with a click

https://tower.nf

Open Core

seqeralabs

nextflow tower



Community: Open source

Cloud: Free & paid tiers

Enterprise: Commercial

`https://tower.nf`

Open Core

seqeralabs

Open training materials

Open source code

Open communities

Phil Ewels

https://phil.ewels.co.uk
phil@seqera.io
tallphil
ewels

seqeralabs
https://seqera.io

Chan Zuckerberg Initiative

nextflow
SUMMIT 2022

https://summit.nextflow.io

Mentorship applications

Close November 1st 2022

Nextflow / nf-core training

6-10 March 2023

nf-core hackathon

20-24 March 2023

```
https://nf-co.re/join
```