

NEXTFLOW & NF-CORE

Phil Ewels

National Genomics Infrastructure
SciLifeLab Sweden

phil.ewels@scilifelab.se



NATIONAL
GENOMICS
INFRASTRUCTURE

SciLifeLab

<https://scilifelab.se>
<https://ngisweden.scilifelab.se>



SciLifeLab



Stockholm
University



Karolinska
Institutet



ROYAL INSTITUTE
OF TECHNOLOGY



UPPSALA
UNIVERSITET



SciLifeLab

163

Research
groups

8

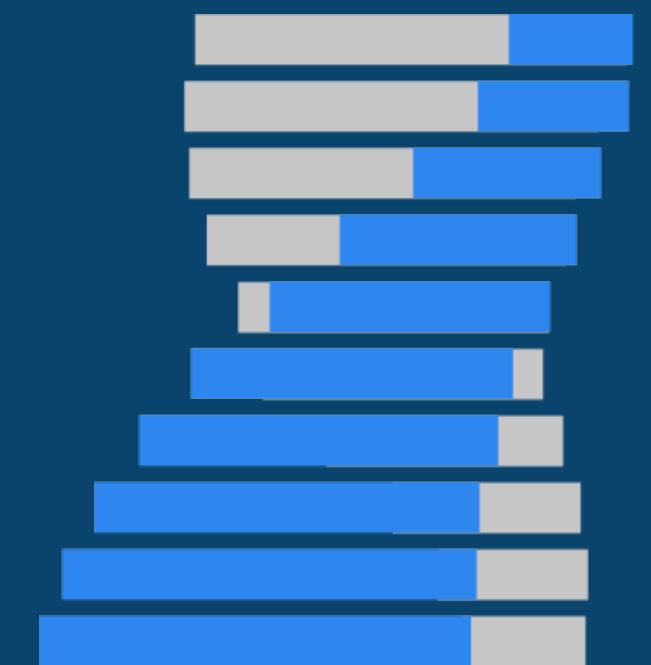
Platforms

37

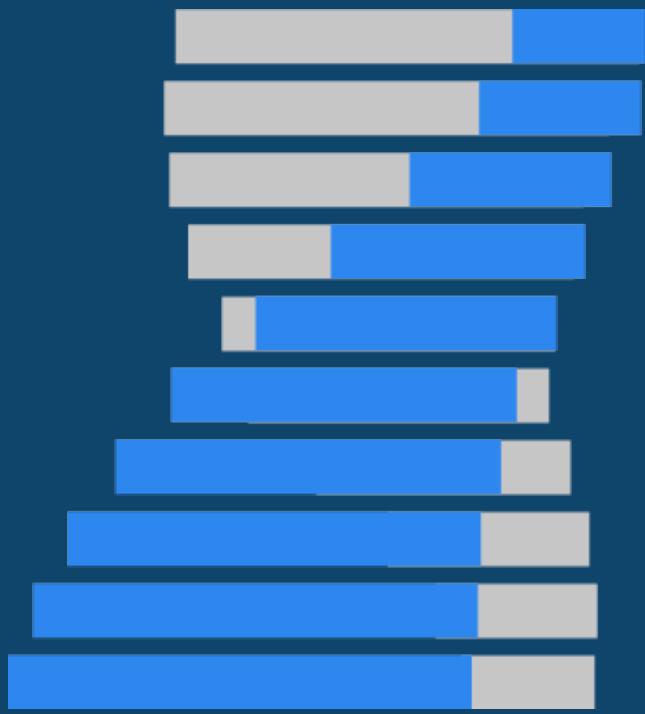
Facilities



SciLifeLab



NATIONAL
GENOMICS
INFRASTRUCTURE



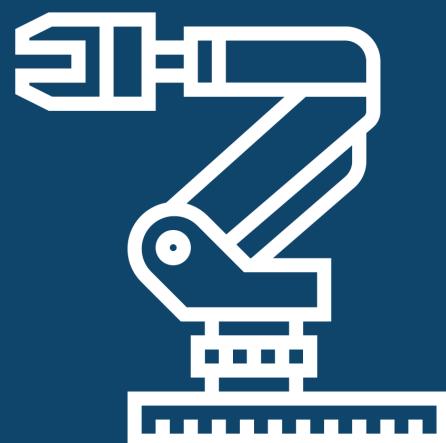
NATIONAL GENOMICS INFRASTRUCTURE



Guidelines
and support



Available to all
Swedish researchers



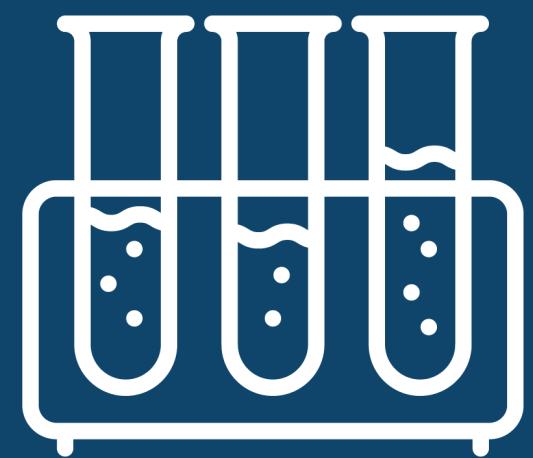
State-of-the-art
infrastructure



NATIONAL GENOMICS INFRASTRUCTURE



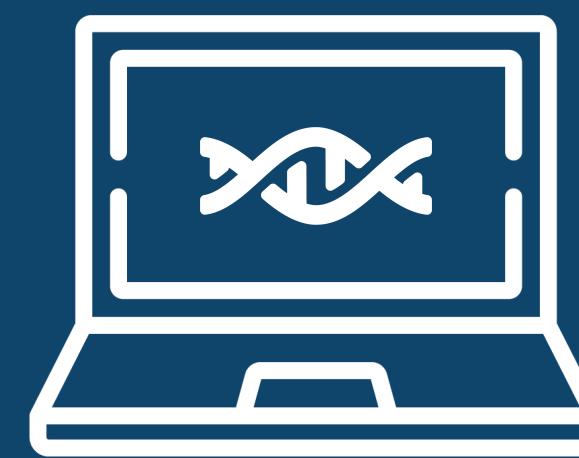
Meetings



QC & Library
Preparation



Sequencing
Genotyping



Data analysis
& delivery

- RNA-seq
- DNA-seq
- Methylation
- ATAC-seq
- Hi-C
- Modifications
- Enrichment



NATIONAL GENOMICS INFRASTRUCTURE



QC & Library
Preparation

- RNA-seq



PCR-free

- DNA-seq

- Methylation



Low input



Ideal for ChIP-seq!

- ATAC-seq

- Hi-C

ngisweden.scilifelab.se

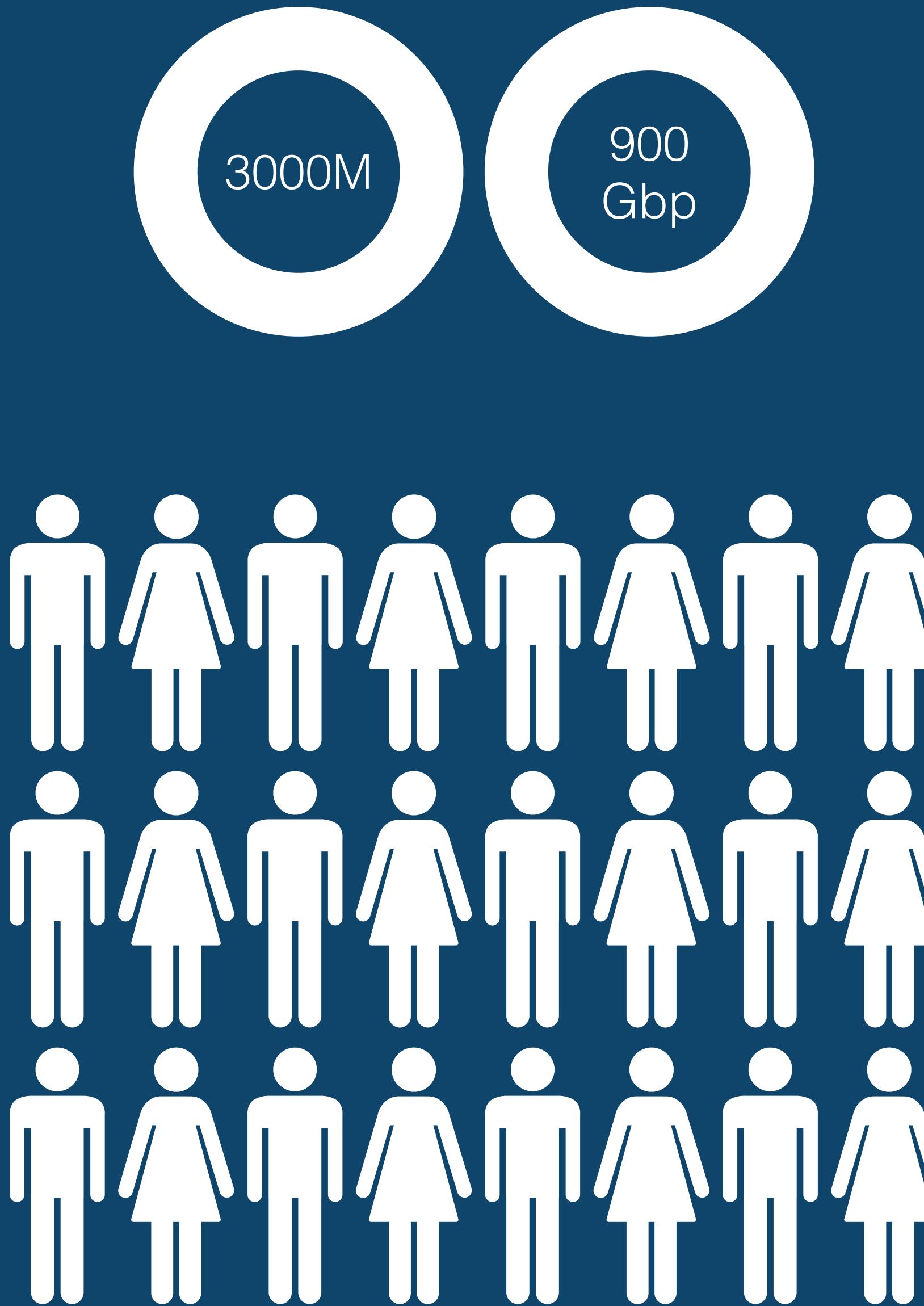
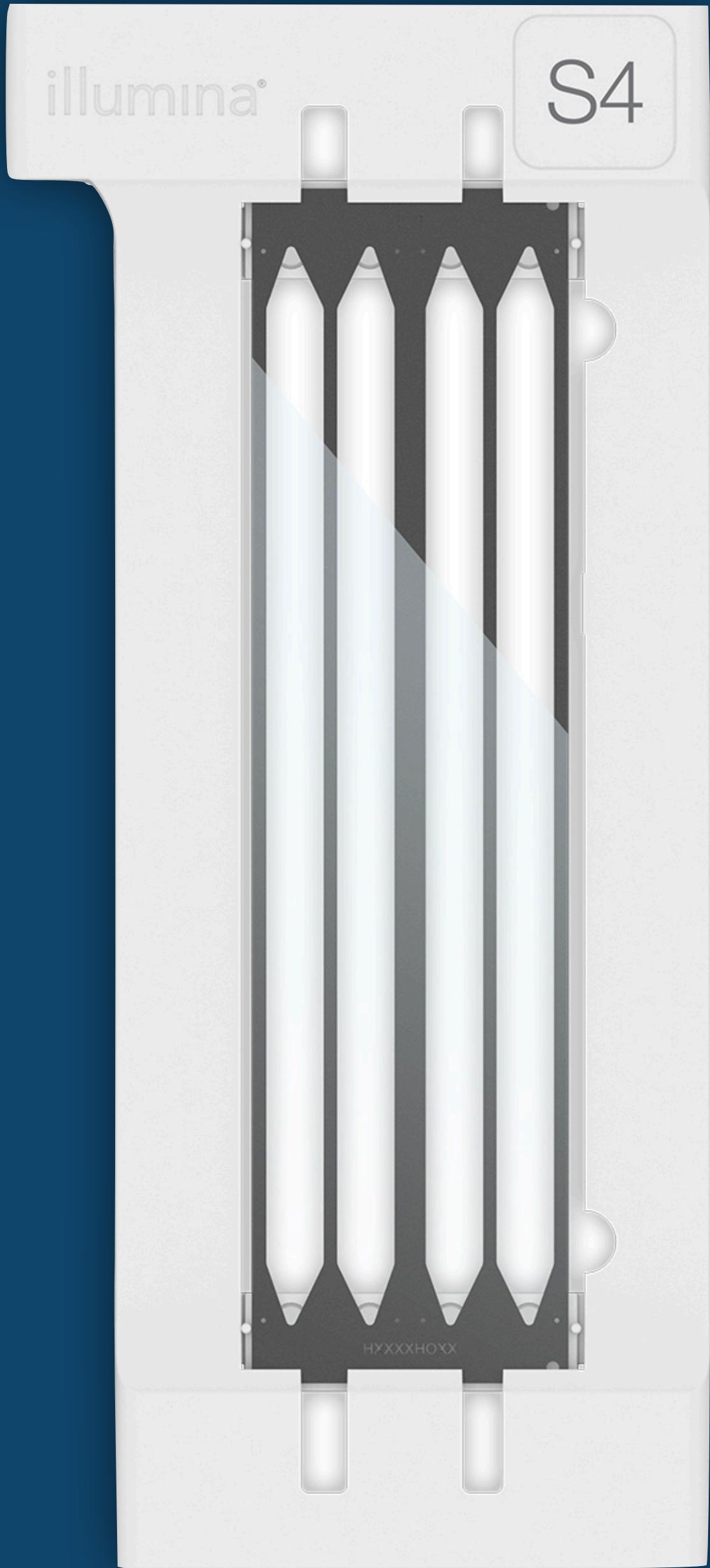
- Modifications

support@ngisweden.se

- Enrichment



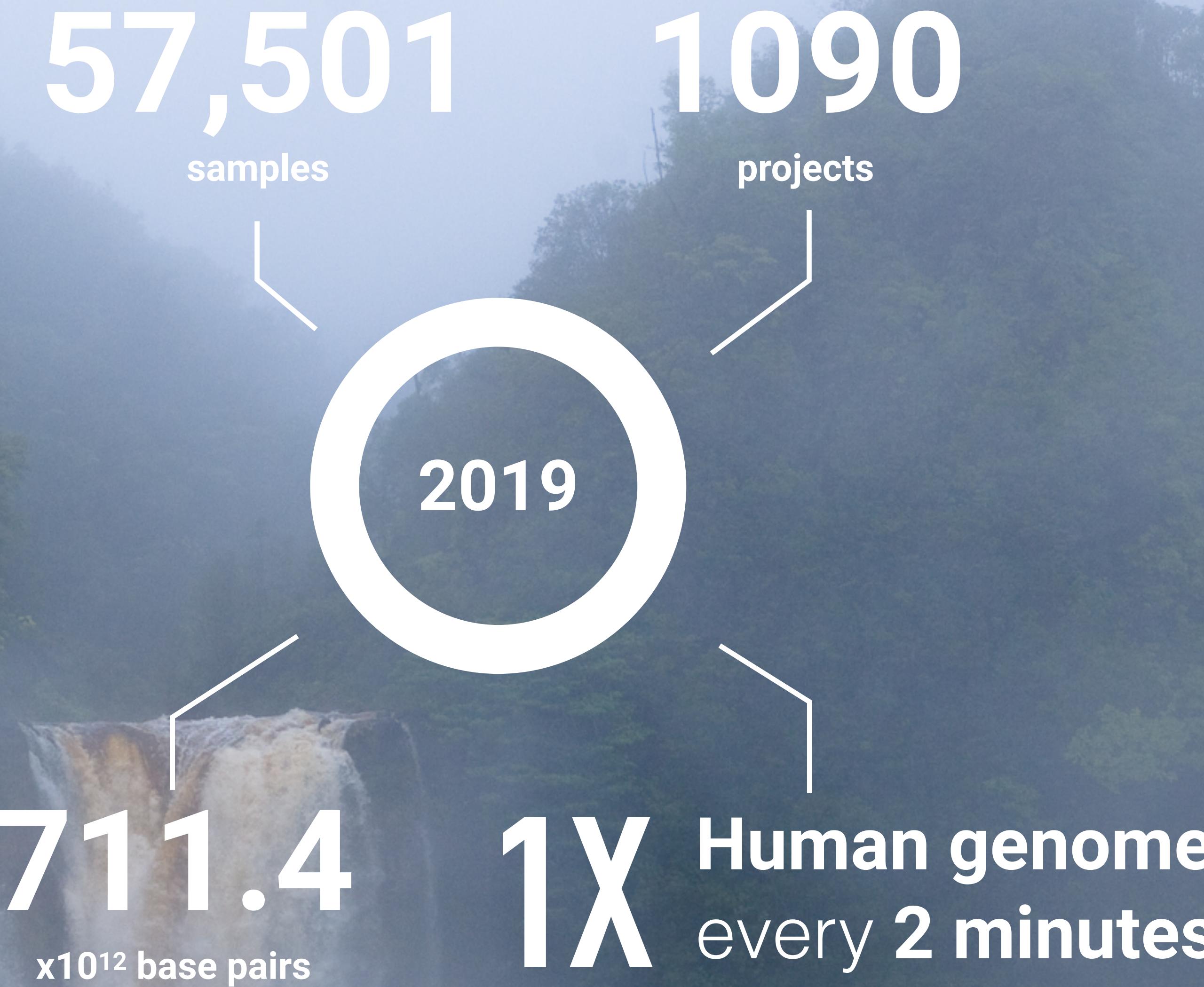




~24 genomes
30X



NATIONAL
GENOMICS
INFRASTRUCTURE



The background of the image is a wide-angle, blue-tinted photograph of a mountainous landscape. The foreground shows dark silhouettes of evergreen trees. In the middle ground, a dense forest covers the base of the mountains. The background features several layers of mountains, with the furthest ones appearing as a hazy blue. A single wind turbine is visible on one of the peaks in the distance.

nextflow

<https://www.nextflow.io/>

nextflow

```
#!/usr/bin/env nextflow
input = Channel.fromFilePairs(params.reads)

process fastqc {
    input:
        file reads from input

    output:
        file "*_fastqc.{zip,html}" into results

    script:
        """
        fastqc -q $reads
        """
}
```



SGE



CONDA



LSF

PBS

<https://www.nextflow.io/>

nextflow

```
#!/usr/bin/env nextflow
input = Channel.fromFilePairs(params.reads)

process fastqc {
    input:
        file reads from input

    output:
        file "*_fastqc.{zip,html}" into results

    script:
    """
        fastqc -q $reads
    """

}
```



Singularity



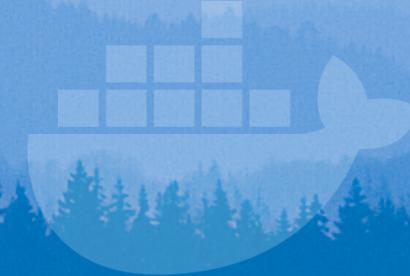
kubernetes



SGE



Google Cloud



docker



CONDA



slurm
workload manager



LSF

PBS

<https://www.nextflow.io/>

nextflow

```
#!/usr/bin/env nextflow
input = Channel.fromFilePairs(params.reads)

process fastqc {
    input:
        file reads from input

    output:
        file "*_fastqc.{zip,html}" into results

    script:
    """
        fastqc -q $reads
    """

}
```



LSF

PBS

<https://www.nextflow.io/>



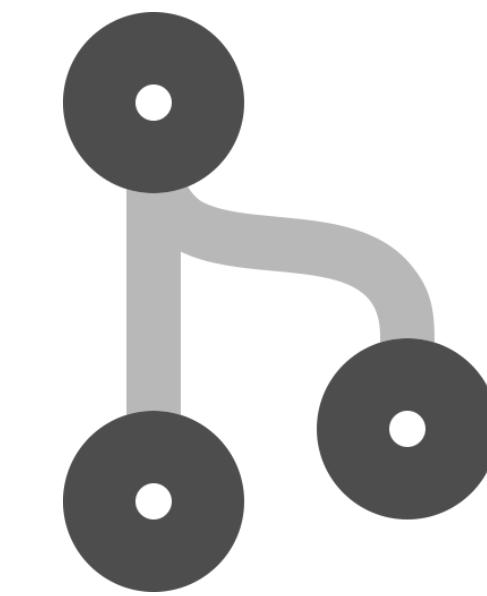
A community effort to collect a curated set of
analysis pipelines built using Nextflow.



Guidelines



Tools



Pipelines

nf-core



Available Pipelines

Can you think of another pipeline that would fit in well? [Let us know!](#)

Search keywords

Filter: Released 25 Under development 14 Archived 3

Sort: Last Release Alphabetical Stars

Display:

nf-core/viralrecon ✓

☆ 18

amplicon assembly covid-19 covid19 illumina metagenomics
sars-cov-2 variant-calling viral virus

Assembly and intrahost/low-frequency variant calling for viral samples

Version 1.1.0

Published 7 days ago

nf-core/sarek ✓

☆ 49

cancer germline somatic variant-calling

Analysis pipeline to detect germline or somatic variants from WGS / targeted sequencing

Version 2.6.1

Published 1 week ago



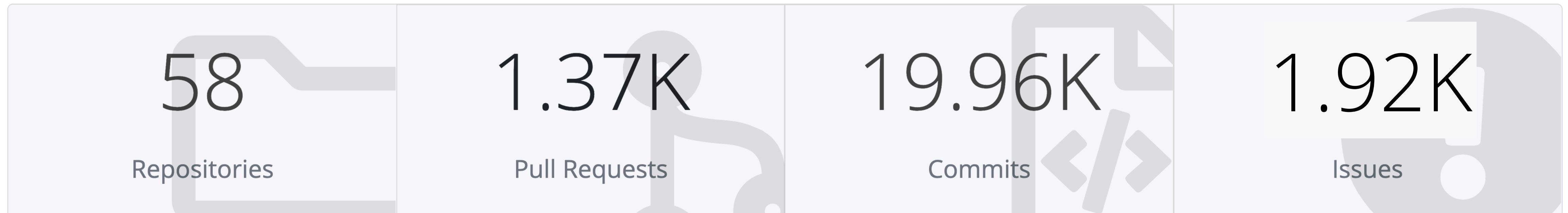
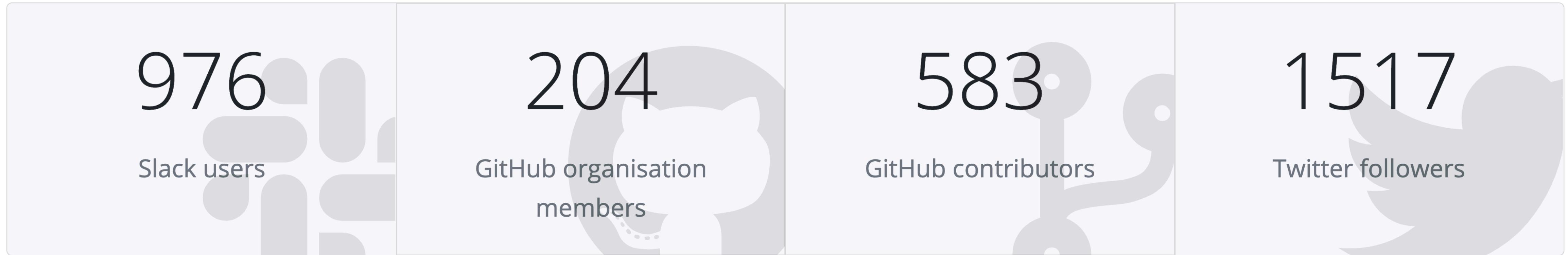


nf-core 

The main title "nf-core" is in large, bold, black font. To the right of the "e" is a stylized green apple logo with a bite taken out of it, featuring a brown stem and a small leaf.

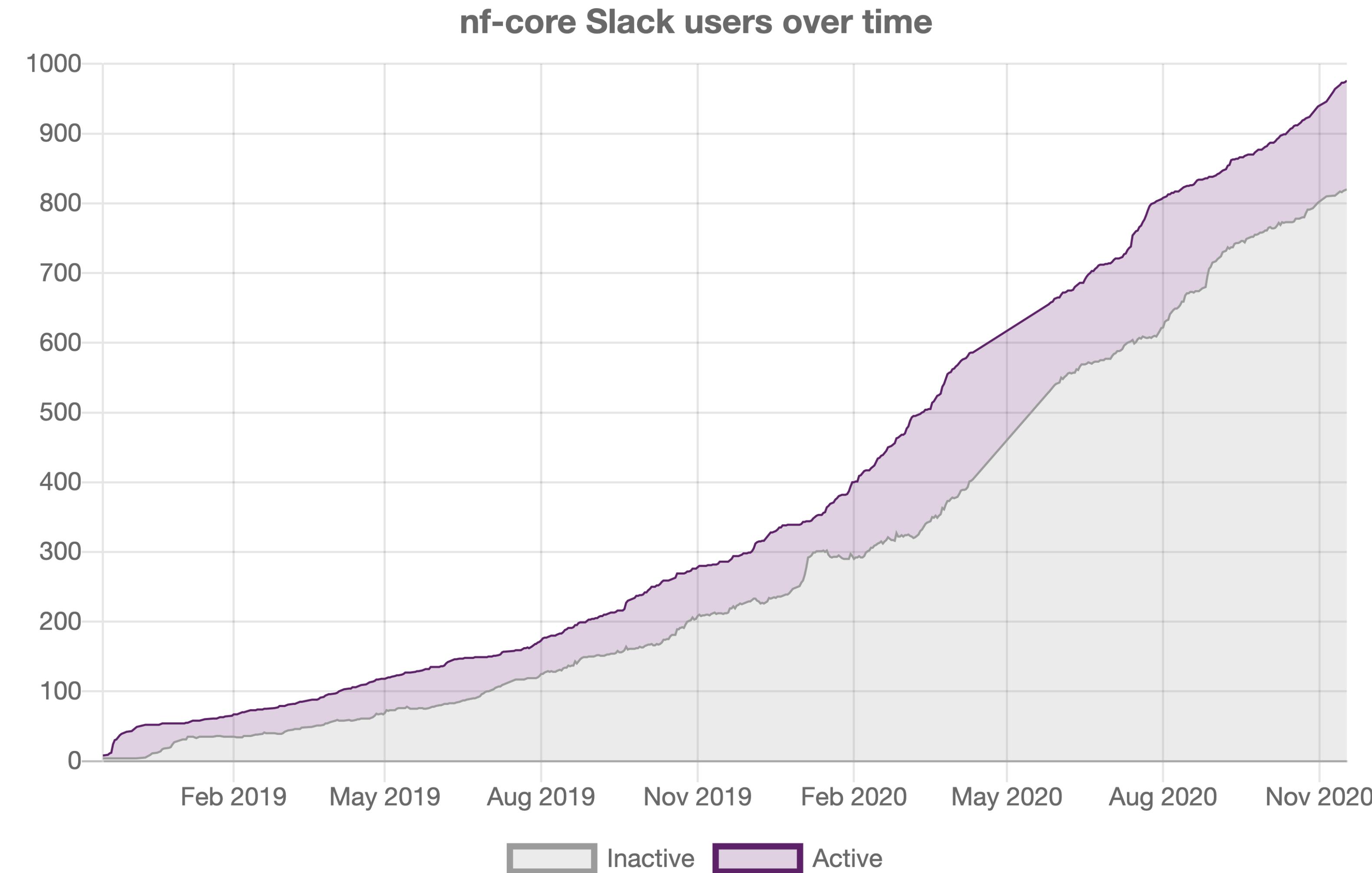
Community

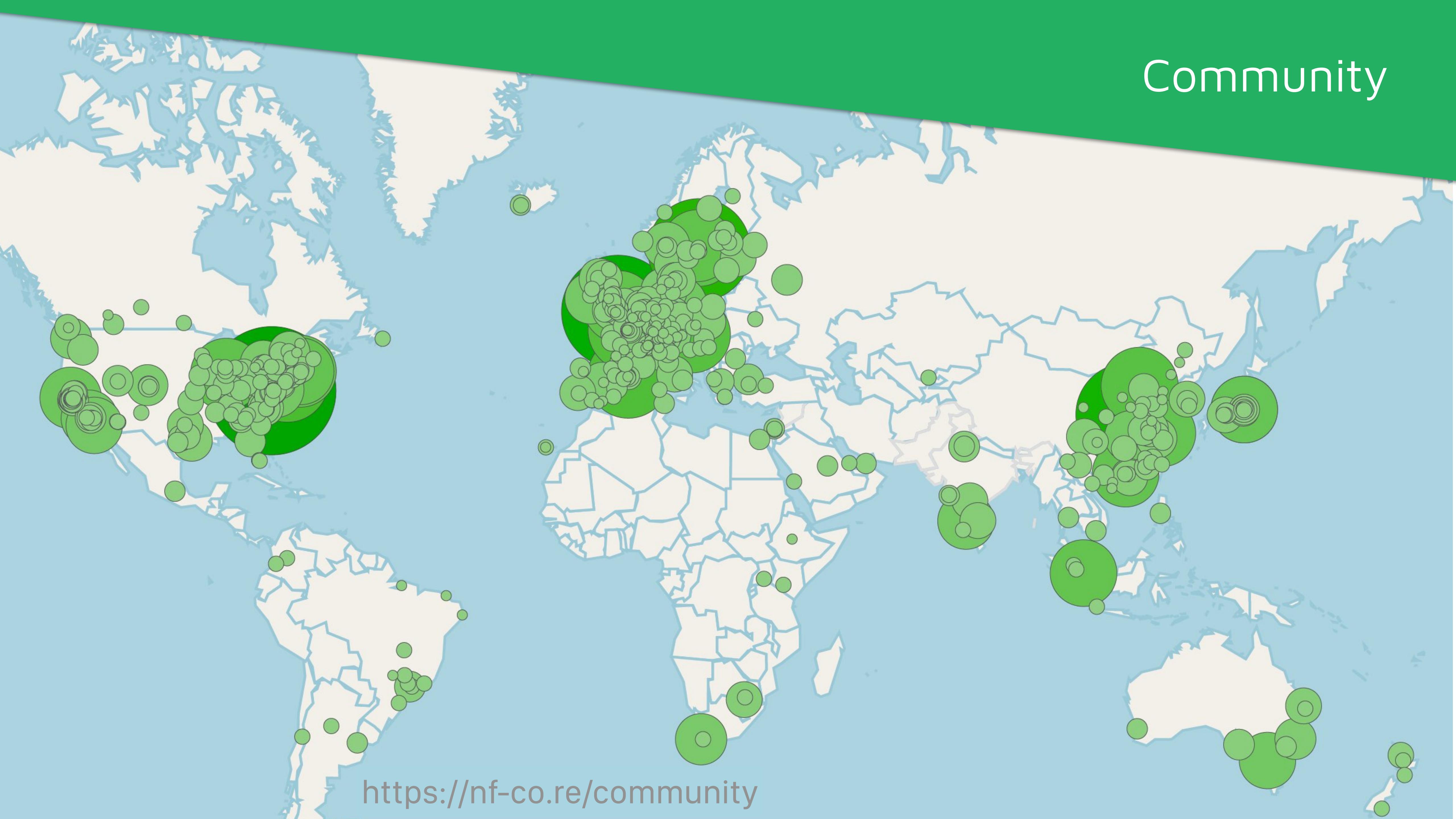
<https://nf-co.re/stats>



Community

<https://nf-co.re/stats>





Community

<https://nf-co.re/community>



<https://nf-co.re/join>

Correspondence | Published: 13 February 2020

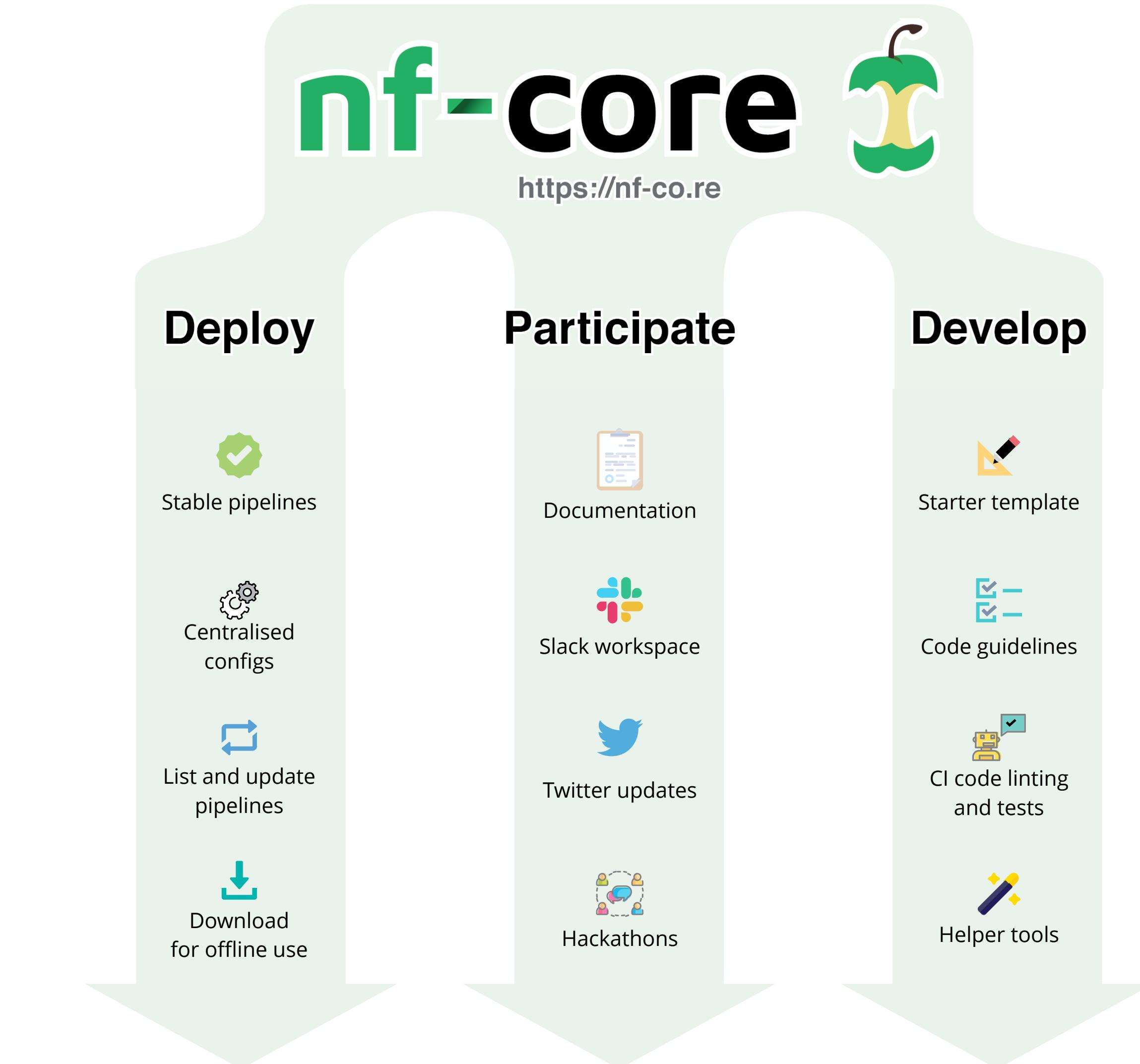
The nf-core framework for community-curated bioinformatics pipelines

Philip A. Ewels, Alexander Peltzer, Sven Fillinger, Harshil Patel, Johannes Alneberg,
Andreas Wilm, Maxime Ulysse Garcia, Paolo Di Tommaso & Sven Nahnsen 

Nature Biotechnology 38, 276–278(2020) | Cite this article

3253 Accesses | 3 Citations | 172 Altmetric | Metrics

To the Editor — The standardization, portability and reproducibility of analysis pipelines are key issues within the bioinformatics community. Most bioinformatics pipelines are designed for use on-premises; as a result, the associated software dependencies and execution logic are likely to be tightly coupled with proprietary computing environments. This can make it difficult or even impossible for others to reproduce the ensuing results, which is a fundamental requirement for the validation of scientific findings. Here, we introduce the nf-core framework as a means for the development of collaborative, peer-reviewed, best-practice analysis pipelines (Fig. 1). All nf-core pipelines are written in Nextflow and so inherit the ability to be executed on most computational infrastructures, as well as having native support for container technologies such as Docker and Singularity. The nf-core community (Supplementary Fig. 1) has developed a suite of tools that automate pipeline creation, testing, deployment and synchronization. Our goal is to provide a framework for high-quality bioinformatics pipelines that can be used across all institutions and research facilities.



TLDR pages *(too long, didn't read)*

Simplified and community-driven man pages

Super-fast and easy reference for the command line

<https://tldr.sh>

pip install tldr

brew install tldr

<https://github.com/raylee/tldr-sh-client>

- List currently running docker containers:
`docker ps`
- List all docker containers (running and stopped):
`docker ps -a`
- Start a container from an image, with a custom name:
`docker run --name container_name image`
- Start or stop an existing container:
`docker start|stop container_name`
- Pull an image from a docker registry:
`docker pull image`
- Open a shell inside of an already running container:
`docker exec -it container_name sh`
- Remove a stopped container:
`docker rm container_name`
- Fetch and follow the logs of a container:
`docker logs -f container_name`

Nextflow

nextflow

```
curl -s https://get.nextflow.io | bash
```

```
module load nextflow
```

```
conda install nextflow BIOCONDA
```

Nextflow



~/.bashrc

```
# Nextflow singularity image cachedir
export NXF_SINGULARITY_CACHEDIR=/proj/MYPROJ/nobackup/phil/singularity-images

# Nextflow environment variables
export NXF_OPTS=' -Xms1g -Xmx4g'
export NXF_HOME=/proj/MYPROJ/nobackup/phil/nxf-home
export NXF_TEMP=${SNIC_TMP:-$HOME/glob/nxftmp}
```

nf-core/

tools

```
pip install nf-core
```



```
conda install nf-core
```

BIOCONDA[®]

Get Miniconda (Linux - Python 2.7, 64 bit)

```
wget https://repo.anaconda.com/miniconda/Miniconda2-latest-Linux-x86_64.sh  
bash Miniconda2-latest-Linux-x86_64.sh
```

Set up bioconda channels (<https://bioconda.github.io/>)

```
conda config --add channels defaults  
conda config --add channels bioconda  
conda config --add channels conda-forge
```

conda install nf-core

BIOCONDA®

Set up bioconda channels (<https://bioconda.github.io/>)

```
conda config --add channels defaults  
conda config --add channels bioconda  
conda config --add channels conda-forge
```

Make an environment

```
conda create --name py3.8 python=3.8  
conda activate Mpy3.8
```

conda install nf-core

BIOCONDA®

nf-core/ tools

[live demo]

Fetching a pipeline

Let Nextflow fetch the pipeline for you!

```
nextflow pull nf-core/<pipeline> -r <version>
```

```
nextflow run nf-core/<pipeline> -r <version>
```

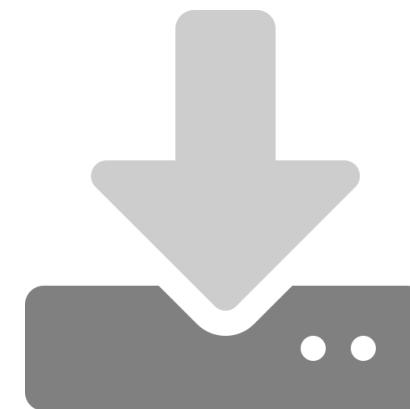
GitHub pipeline name
Clones to ~/.nextflow

Pipeline release tag
Repo branch

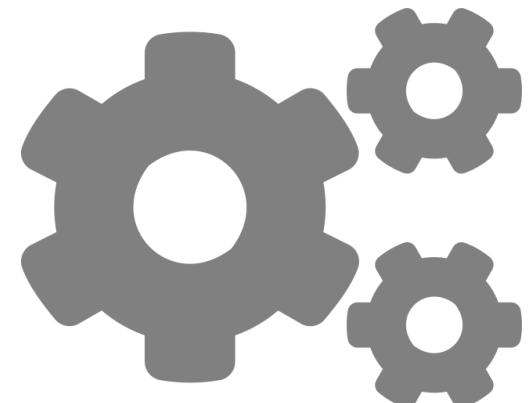
Fetching a pipeline

Offline? Fetch using nf-core helper tool

```
nf-core download nf-core/<pipeline> -r <version>
```



Downloads pipeline



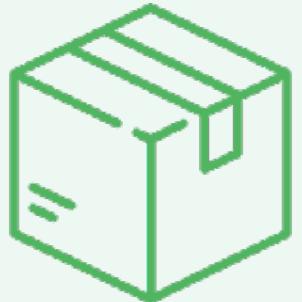
Downloads configs



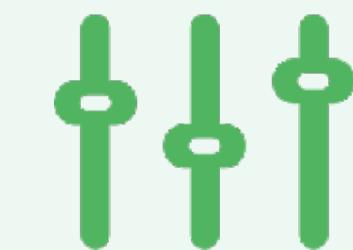
Builds Singularity container



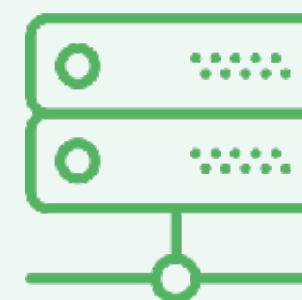
Creates archive



Default 'base' config (always loaded)



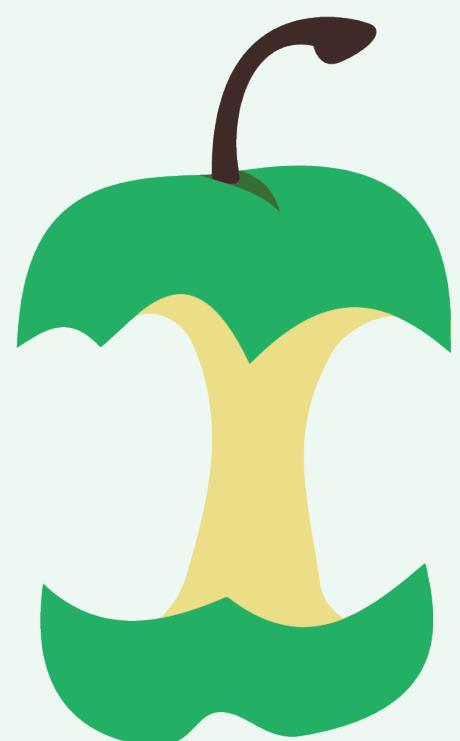
Core profiles (eg. docker, conda, test)



Institutional profiles (nf-core/configs)

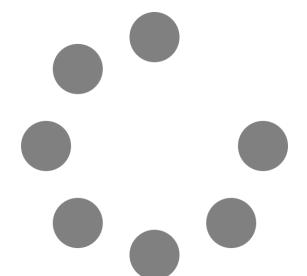


Your local config files (eg. the -c flag)





Default 'base' config



Automatically loaded



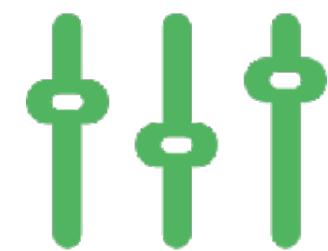
Sensible default resource requests



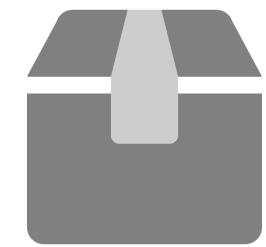
No software packaging



No job submission (runs locally)



Core profiles



Specify software packaging

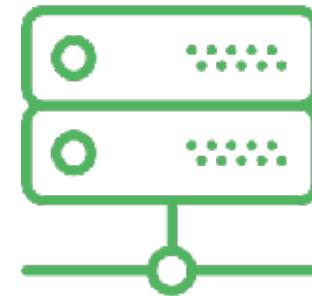


Specify common presets

Multiple profiles

No space after comma!

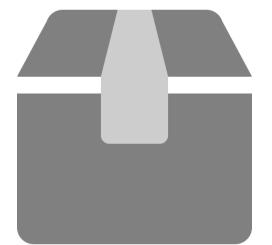
```
nextflow run <pipeline> -profile test,singularity
```



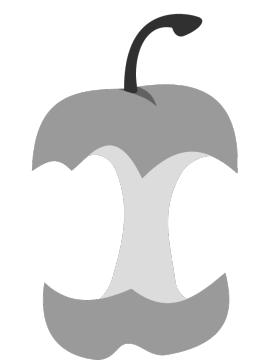
Institutional profiles



Specify job submission for your cluster



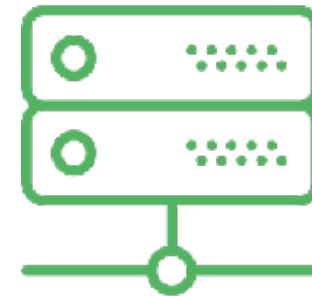
Specify software packaging and other settings



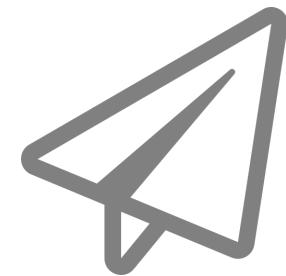
For all nf-core pipelines



For all cluster users



Institutional profiles - UPPMAX



SLURM, with parameters for UPPMAX projects



Singularity, automatic settings for resource limits

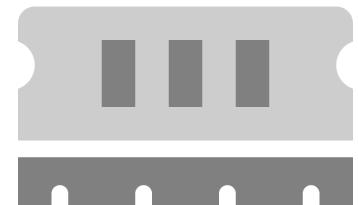


Local copies of iGenomes references

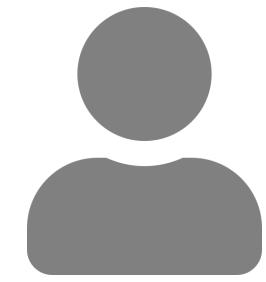
```
nextflow run <pipeline> -profile uppmax --project XXX
```



Your local config files



Custom resource requirements



User specific parameters

```
~/.nextflow/config
```

```
nextflow run -c myconf.config
```



Running a pipeline

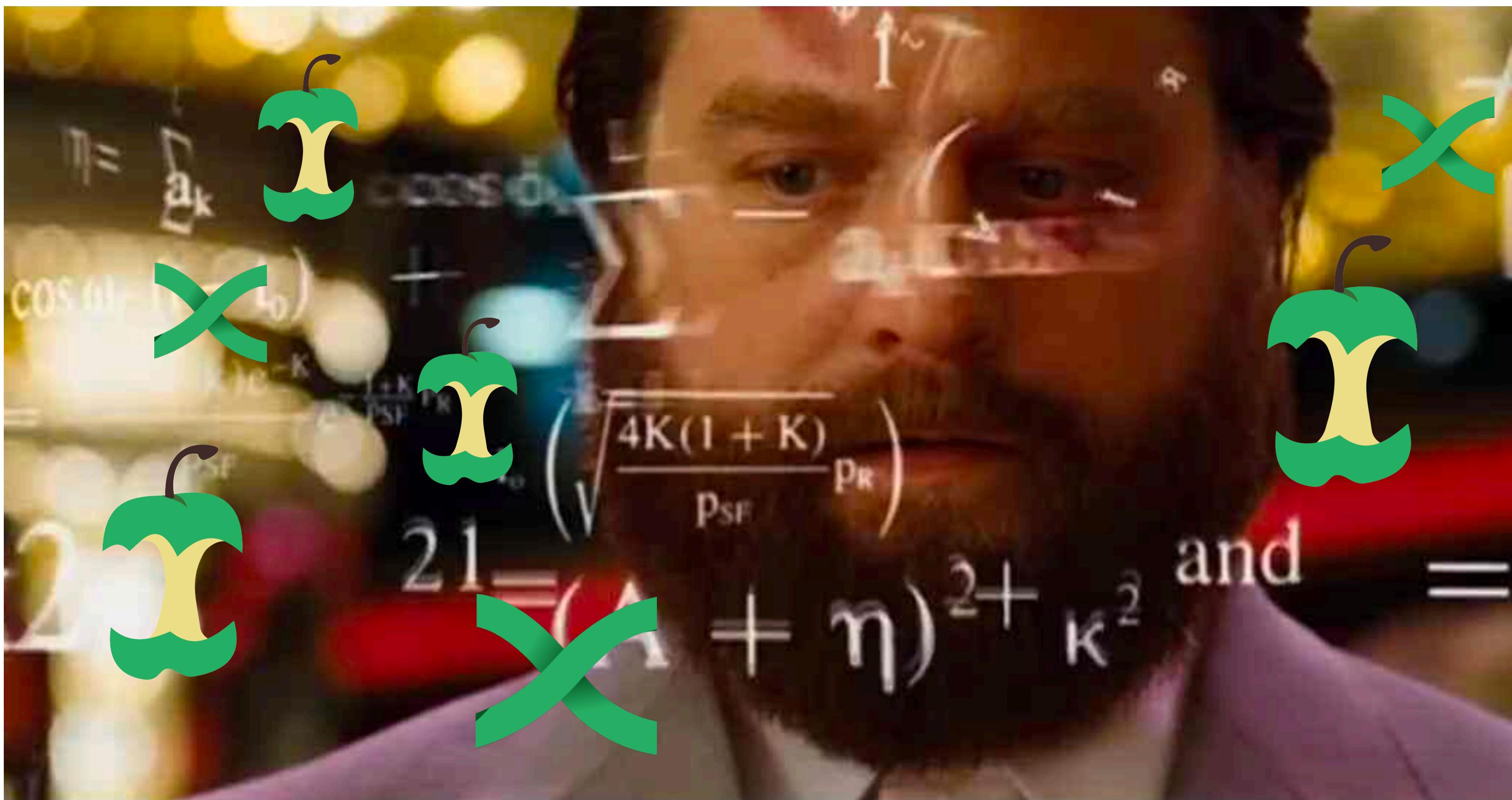
params.option_one

```
nextflow run <pipeline> --option_one
```

```
nextflow run <pipeline> -c myrun.config
```

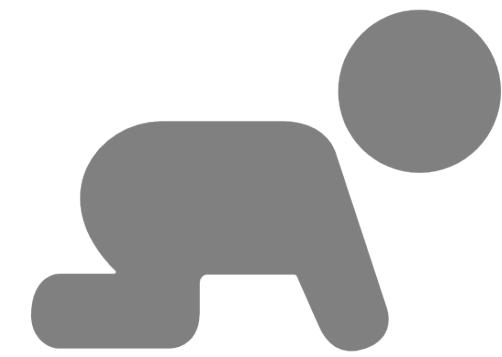
```
nextflow run <pipeline> -params-file myrun.json
```

Configuration



Running a pipeline

```
nextflow run nf-core/<pipeline> -profile test
```



- Runs a very tiny test dataset
- Confirms that your system is set up correctly

Running a pipeline

```
nextflow run nf-core/<pipeline> [params]
```

Remember to detach from your terminal!



-bg flag

screen or tmux

Submit Nextflow as a job

Running a pipeline

```
nf-core launch nf-core/<pipeline>
```



Help / validation for pipeline parameters
Use online form or offline command-line wizard

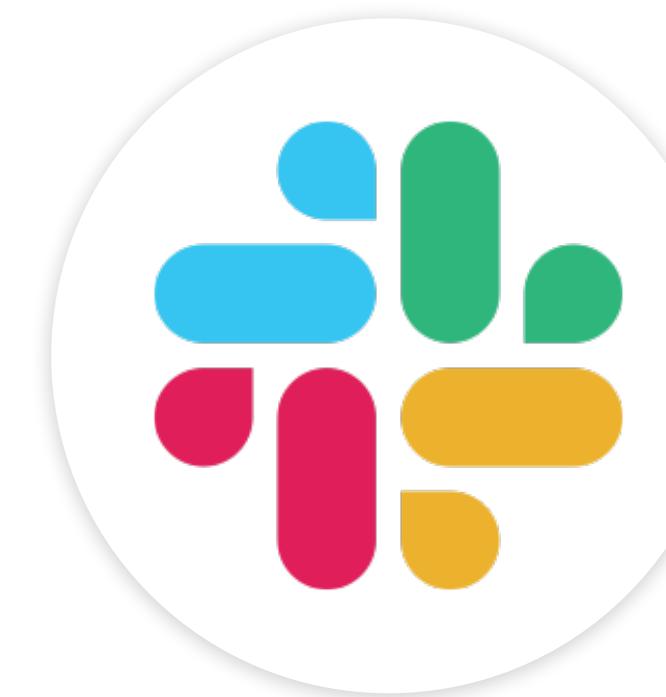
Running a pipeline

```
nf-core launch nf-core/<pipeline>
```



[live demo]

nf-core



<https://nf-co.re/join>

nf-core

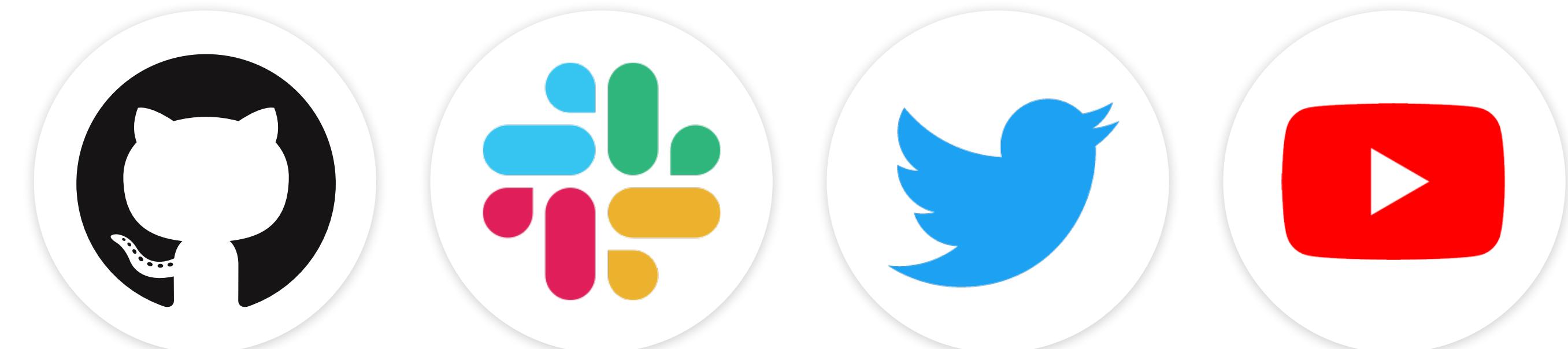


Phil Ewels

<https://phil.ewels.co.uk>



<https://ngisweden.scilifelab.se>



<https://nf-co.re/join>