



nextflow

Reproducible research with Nextflow & Building pipelines with nf-core

Oslo Bioinformatics Workshop Week 2023
Thursday 14 December 2023









nextflow

Anyway, where were we..

We're delighted
to be here

Your instructors



Phil Ewels, Ph.D.

PM for Open Source at [Seqera](#)

phil.ewels@seqera.io

@tallphil | Stockholm

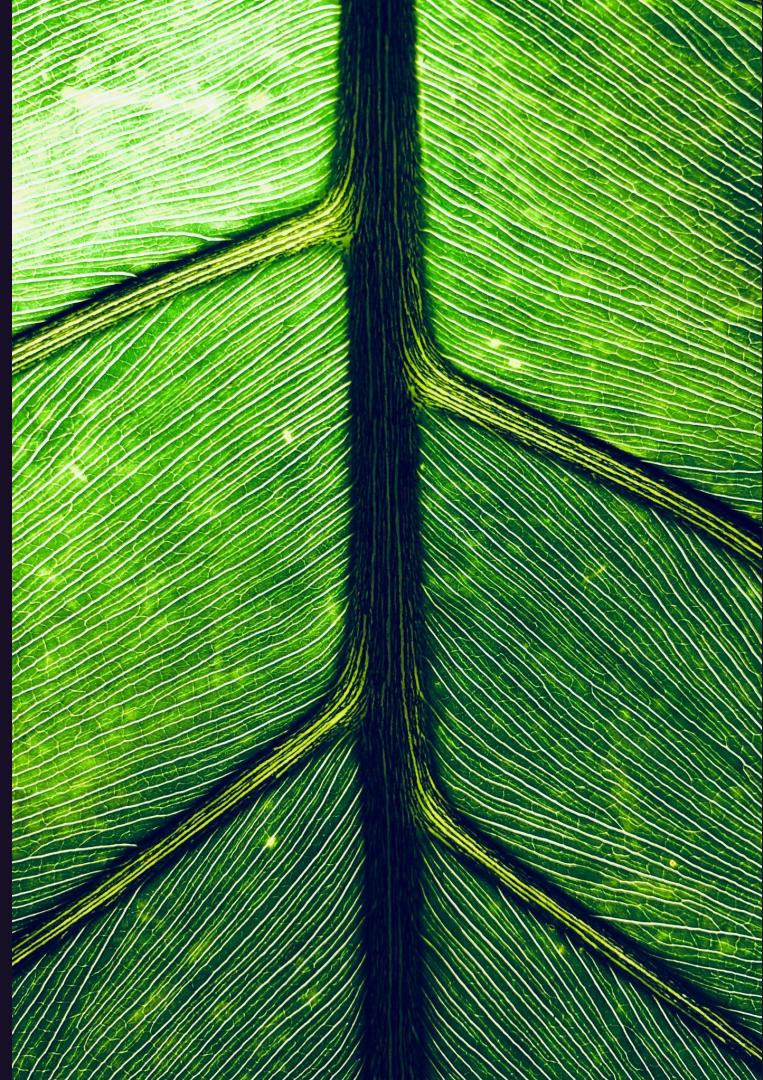


Maxime Garcia, Ph.D.

Bioinformatics Engineer at [Seqera](#)

maxime.garcia@seqera.io

@gau | Stockholm

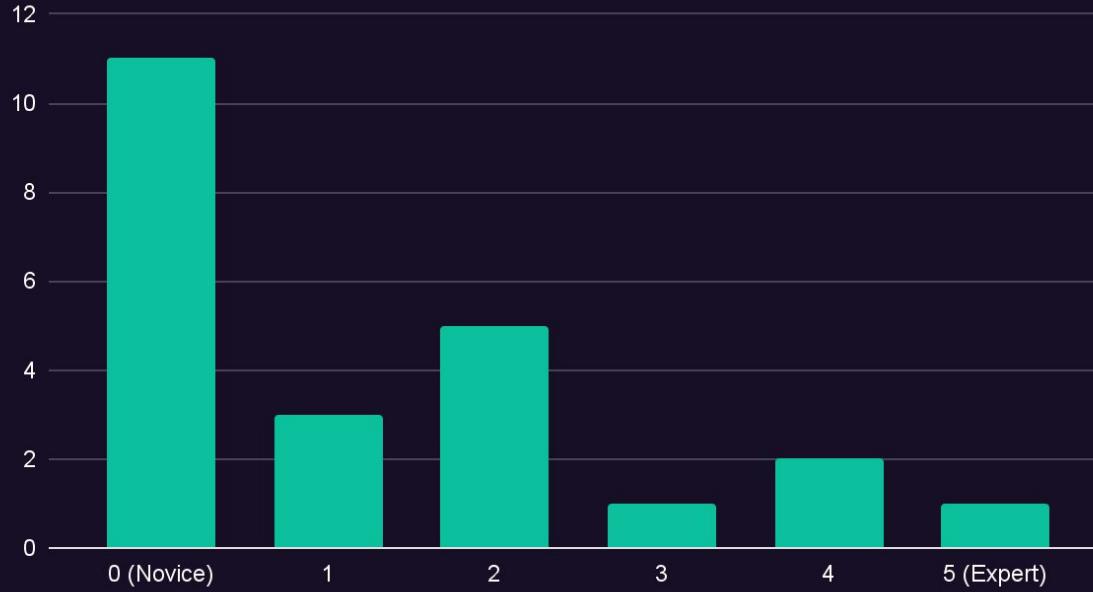


 nextflow multiqc wave fusion

A vertical decorative bar on the left side of the slide features a marbled pattern with organic, irregular shapes in shades of orange, yellow, blue, and white.

About you

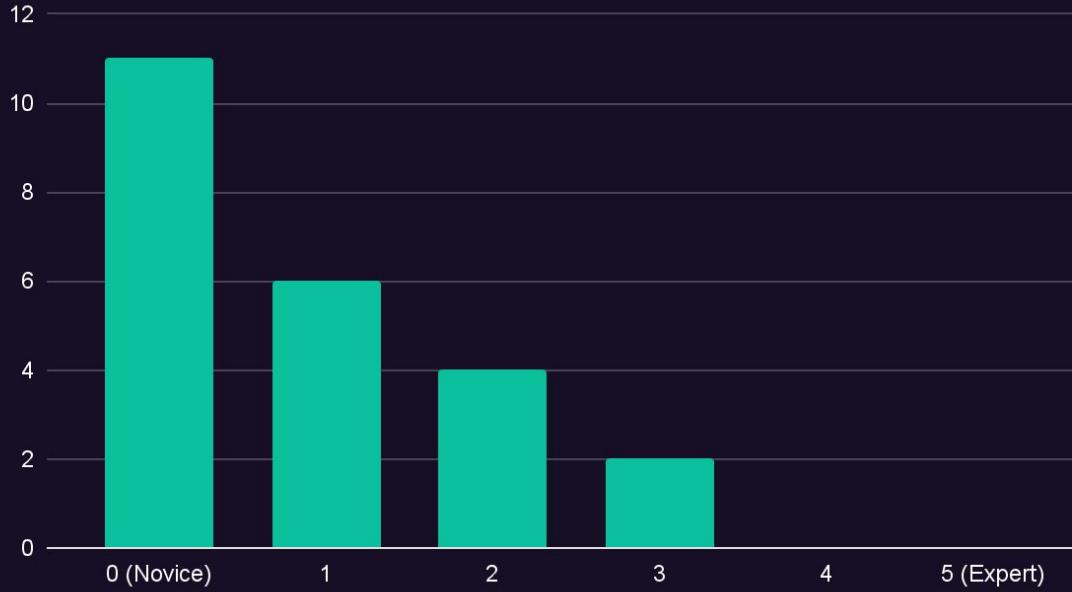
How experienced are you with Nextflow?



A vertical decorative bar on the left side of the slide features a marbled pattern with organic, irregular shapes in shades of orange, yellow, blue, and white.

About you

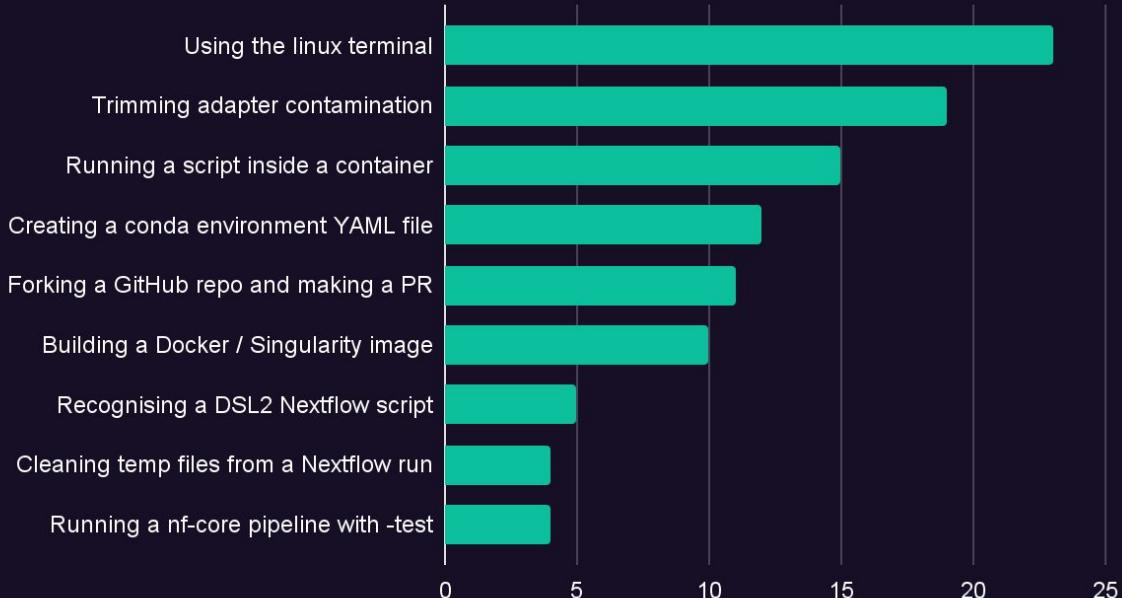
How experienced are you with nf-core?





About you

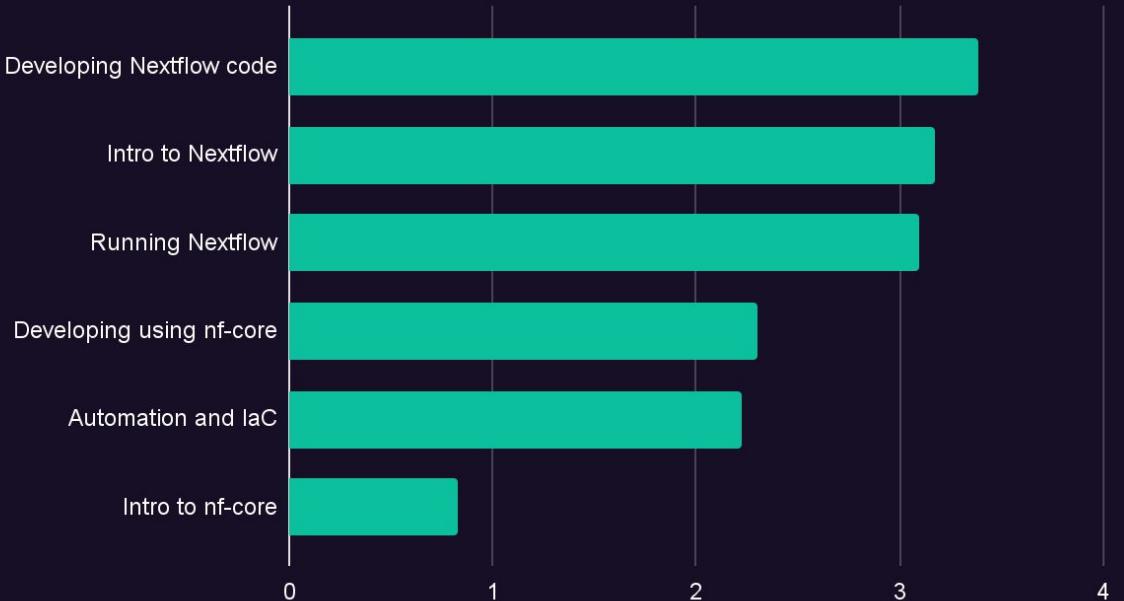
Which of the following would you be confident to do yourself?





About you

Which of the following topics are most interesting for you?





About you

I taught myself nextflow by trial and error so I am keen to get a more formal introduction
Understand the logics behind a nextflow script
Make informed decision of when I should use NextFlow in my projects, and what I can do with it.
using and building softwares using nextflow
Know why, when and how to use nextflow, compared to using kale or snakemake
Understand the potential of using nf-core/nextflow
To better understand nextflow
How to use and run nf-flow using docker
Setting up nextflow

Deploying nf-core pipelines
Being capable of setting up a nextflow pipeline like I can now do with snakemake
Being able to install nextflow and run the RNAseq pipeline from nf-core in my laptop
Learn how to run nf-core/sarek for variant calling
Learning how to write a pipeline with nextflow
Learning to build a pipeline with Nextflow!
mounting to docker with nextflow
how to use the nf core tool to be more efficient working with large datasets
Learn to understand nextflow and how I can use it in my work. Be able to write my own code.
A better understanding of using already made nextflow modules in my own analysis
Gain some confidence in writing my own nextflow workflows

automating pipelines
Set up my own pipeline tuned to our AWS cluster



Reproducible research with Nextflow & building pipelines with nf-core

Oslo - 13 December, 2023

Session 1: Introduction to Nextflow

- Introduction to Nextflow/nf-core (30 min)
- Building our first pipeline (1 hour)
- Software containers (15 min)
- *Break (15 min) 10:45 – 11:00*
- Channels and processes (15 min)
- Configuring Nextflow (30 min)
- Q&A (15 min)

Session 2: Building with nf-core

- Finding and launching pipelines (15 min)
- Institutional configs (15 mins)
- Running with the test profile (30 min)
- Downloading pipelines (15 min)
- *Break (15 min) 14:15 – 14:30*
- Anatomy of an nf-core pipeline (30 min)
- nf-core tooling for developers (30 min)
- Rich interfaces for Nextflow (15 min)
- Q&A (15 min)



Reproducible research with Nextflow & building pipelines with nf-core

Oslo - 13 December, 2023

Session 1: Introduction to Nextflow

- Introduction to Nextflow/nf-core (30 min)
- Building our first pipeline (1 hour)
- Software containers (15 min)
- *Break (15 min) 10:45 – 11:00*
- Channels and processes (15 min)
- Configuring Nextflow (30 min)
- Q&A (15 min)

Session 2: Building with nf-core

- Finding and launching pipelines (15 min)
- Institutional configs (15 mins)
- Running with the test profile (30 min)
- Downloading pipelines (15 min)
- *Break (15 min) 14:15 – 14:30*
- Anatomy of an nf-core pipeline (30 min)
- nf-core tooling for developers (30 min)
- Rich interfaces for Nextflow (15 min)
- Q&A (15 min)



Workflows

Using computers to collect, store, analyze, and disseminate data and information



Large files

> 100 GB for one human genome...



Many languages

Bash, Python, R, Perl...



Complex interactions

Networks of software and their dependencies...



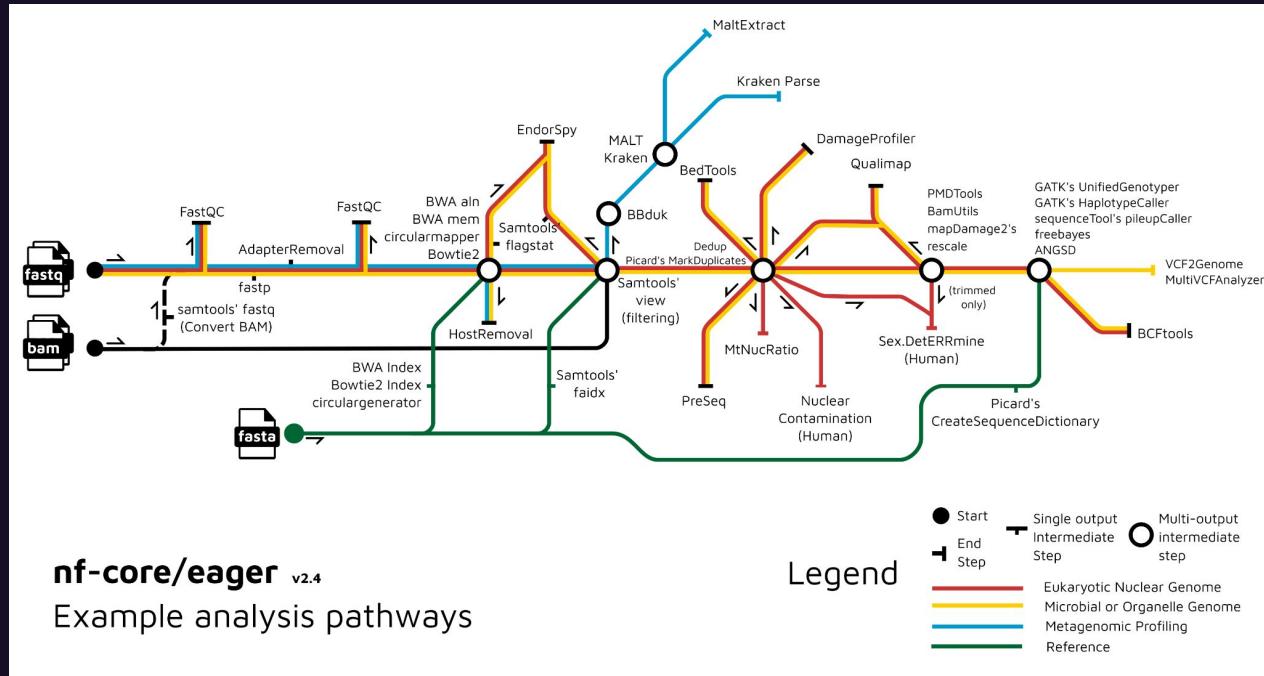
Reproducibility

Hidden reproducibility issues exist under the surface



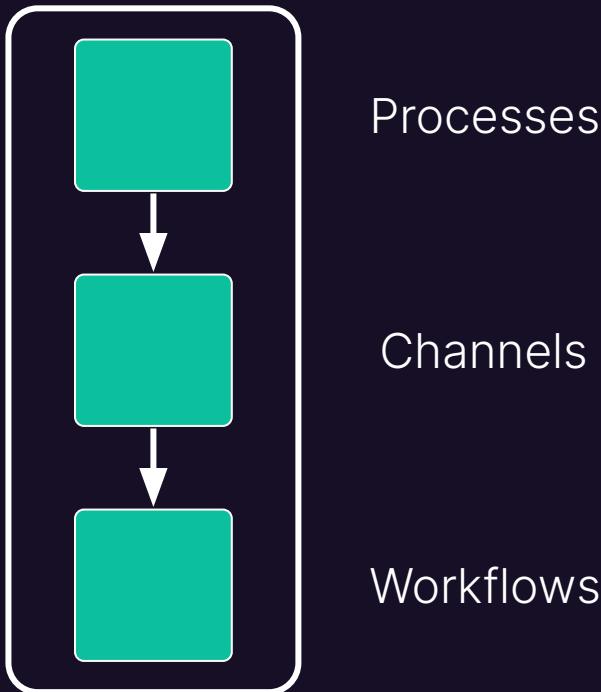
Nextflow

Managing modern workflows is complicated



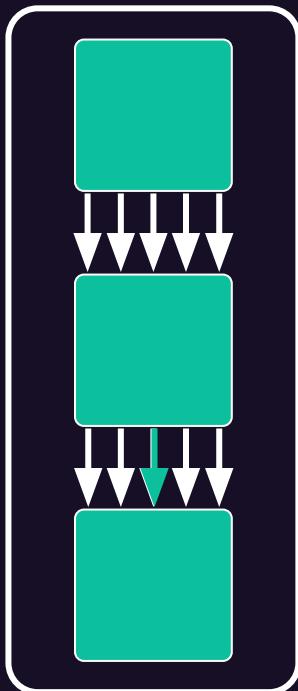
Nextflow

A reactive workflow framework and a programming DSL



Nextflow

A reactive workflow framework and a programming DSL



Parallelization

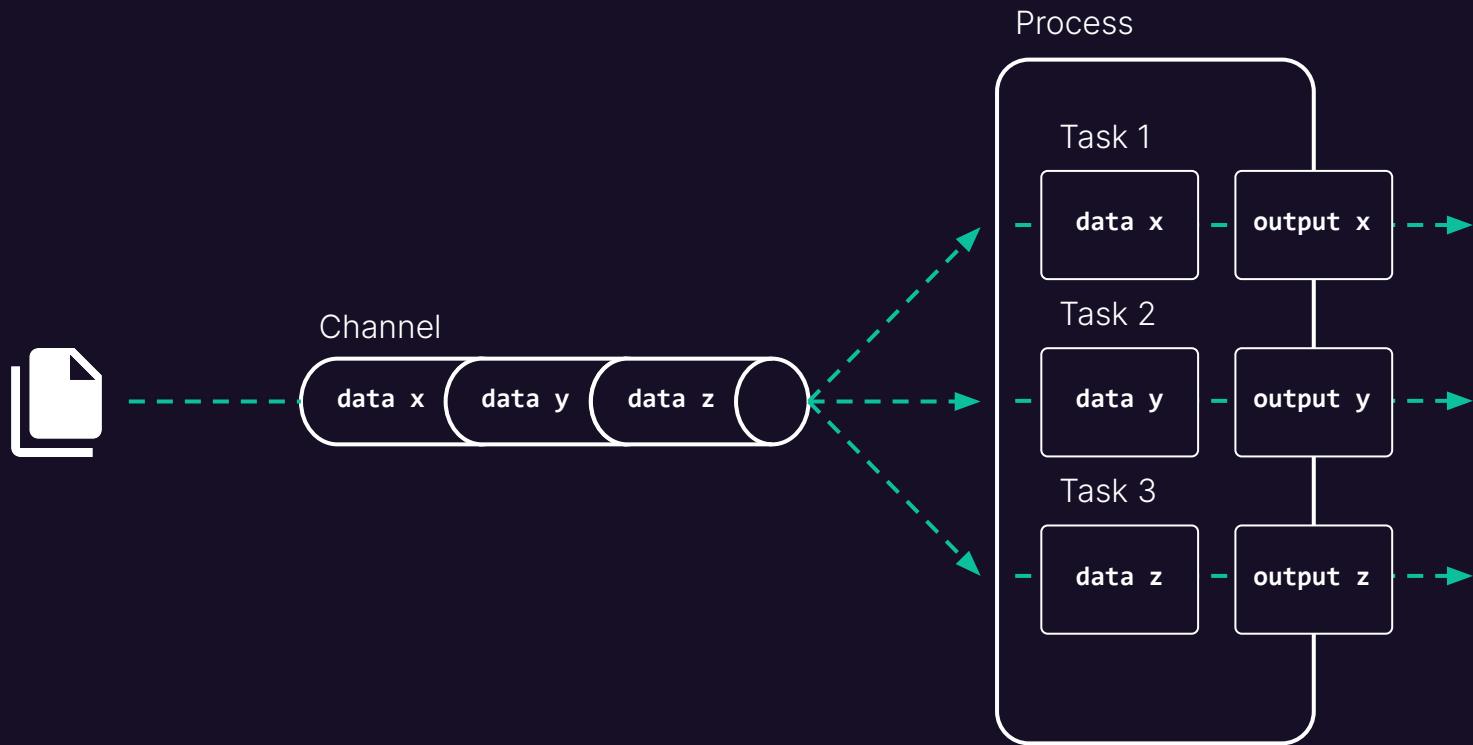
Reentrancy

Reusability



Nextflow

A reactive workflow framework and a programming DSL



Nextflow

Nextflow is a language, a runtime, and a community



Write code
in any language



Orchestrate tasks with
dataflow programming



Define software
dependencies via containers



Built-in version control
with Git



Task orchestration
and execution



Nextflow

Nextflow is a language, a runtime, and a community



Reproducible

Integration with code management tools, with versioned releases.



Portable

Docker, Singularity, Conda, works with most compute environments.



Scalable

5 samples on your laptop, 5k on an HPC or 5 million in the cloud.



Nextflow plugins

Extensible components downloaded and installed at runtime



nf-validation

- Validate parameters
- Create channels from sample sheets



nf-prov

- Provenance reporting
- BioCompute Objects / RO-Crates



nf-co2footprint

- Report on energy consumption and CO2 emissions



nf-float

- Adds a new Nextflow executor for working with MemVerge





A simple testing framework for Nextflow pipelines

- Describes **expected behavior** using 'when' and 'then' blocks
- Abundance of functions for writing **elegant and readable** assertions
- Utilizes **snapshots** to write tests for complex data structures
- Provides commands for generating **boilerplate code**
- Includes a **test-runner** that executes these scripts
- Easy installation on **CI systems**

```
● ● ●  
nextflow_pipeline {  
  
    name "Test Hello World"  
    script "nextflow-io/hello"  
  
    test("hello world example should start 4 processes") {  
        expect {  
            with(workflow) {  
                assert success  
                assert trace.tasks().size() == 4  
                assert "Ciao world!" in stdout  
                assert "Bonjour world!" in stdout  
                assert "Hello world!" in stdout  
                assert "Hola world!" in stdout  
            }  
        }  
    }  
}
```



Nextflow plugins

nf-validation: a Nextflow plugin to natively handle scheme files



Parameter validation

Functions to validate pipeline parameters and print CLI help text



Sample sheet validation

Validate input sample sheet - formatting and contents



Sample sheet channels

`fromsamplesheet()`
channel factor,
including meta maps



nextflow-io/nf-validation



<https://github.com/nextflow-io/nf-validation>



nf-core

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

7k+

Slack
users

2k+

GitHub
contributors

120+

GitHub
repositories

40k

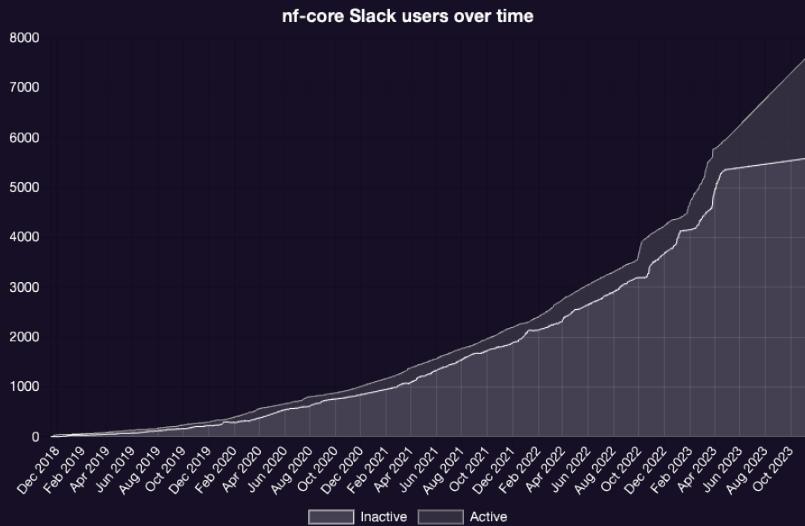
GitHub
commits

16k+

Pull
requests

7k+

GitHub
issues



nf-core principles

Principles that guide the nf-core community



Cooperation

Develop with the community



Standards

Use a common template



Collaboration

No duplicate pipelines within nf-core



Helper Tools

Tools built for everyone



Compatibility

Tools work for any Nextflow pipeline



Components

Collaborate on component-level



nf-core resources

Pick and choose which components you need



Pipelines

>90 pipelines and a base template



Subworkflows

>50 subworkflows



Modules

>1050 modules



Linting

Choose conventions to test for consistency



Schema

Validation, channels and user interface



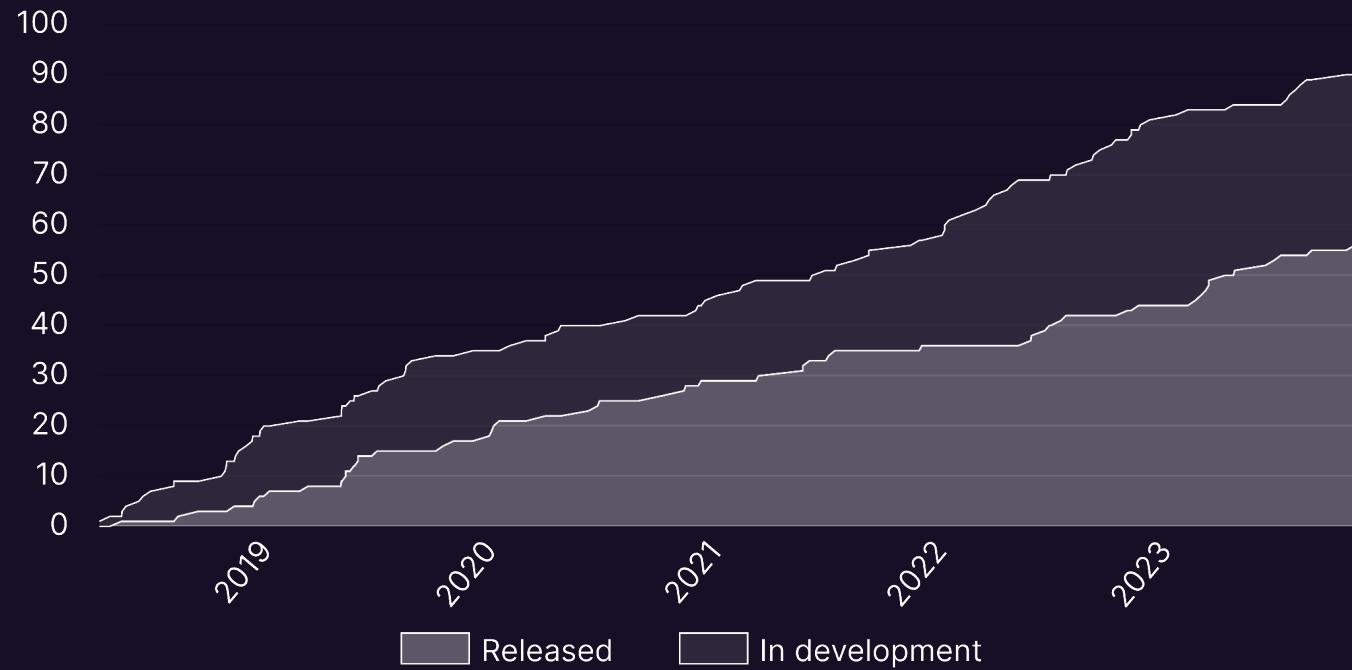
Tooling

Development and deployment



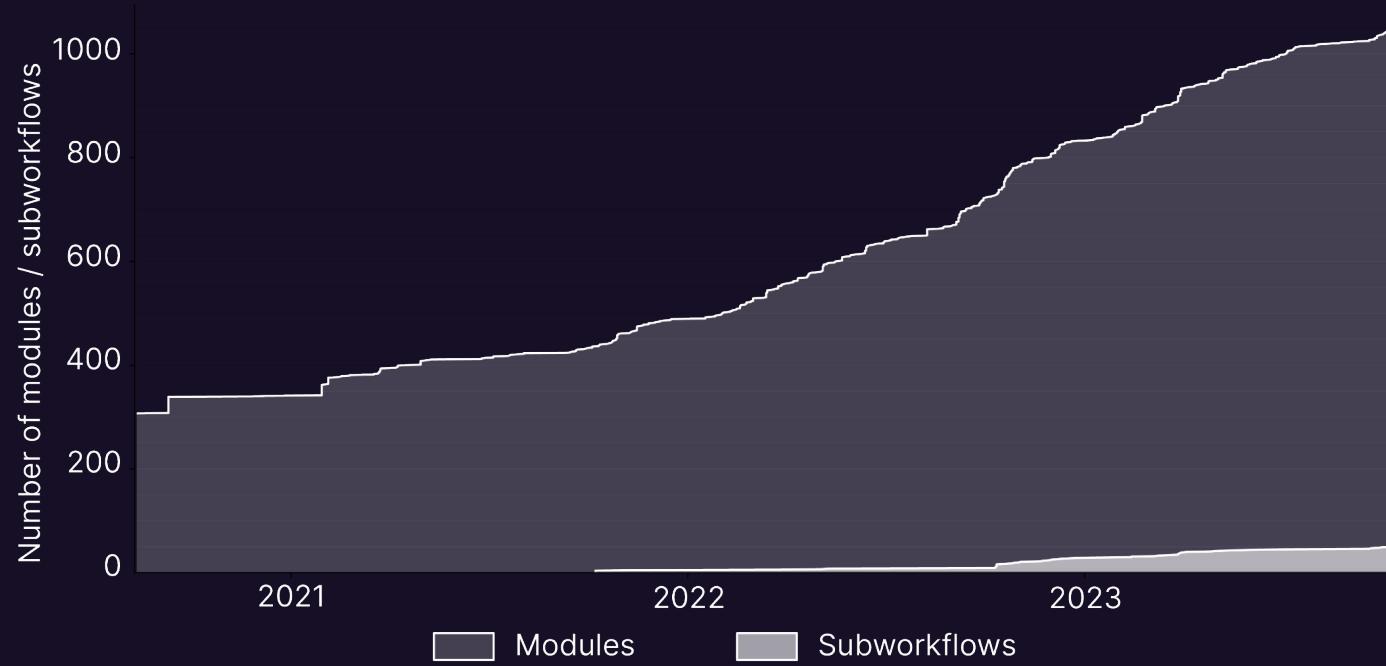
nf-core pipelines

Accelerating



nf-core modules

Number of nf-core modules and subworkflows over time



nf-core components

Building from lego bricks

Custom

nf-core

Pipeline

- Pipeline parameters
- Resource requests
- Tool arguments
- Input channels

Subworkflow

- Channels
- Pipeline logic

Module

Module

Module



nf-core components

Building from lego bricks

Custom

nf-core

Pipeline

- Pipeline parameters
- Resource requests
- Tool arguments
- Input channels

Subworkflow

- Channels
- Pipeline logic

Module

Module

Module



nf-core components

Building from lego bricks

Custom

nf-core

Pipeline

- Pipeline parameters
- Resource requests
- Tool arguments
- Input channels

Subworkflow

- Channels
- Pipeline logic

Module

Module

Module



nf-core components

Building from lego bricks

Custom

nf-core

Pipeline

- Pipeline parameters
- Resource requests
- Tool arguments
- Input channels

Subworkflow

- Channels
- Pipeline logic

Module

Module

Module



nf-core/tools

Command line tools to help you build your pipeline with ease



Pipelines

Create from template,
sync to get updates



Subworkflows

Create, install and
update



Modules

Create, install, update,
patch, test



Schema

Build your pipeline
schema with a GUI



Linting

Test nf-core standards
and best practices



Download

Fetch with singularity
images for offline use



nf-core/tools

Command line tools to help you build your pipeline with ease



Helper tools

- Improved download capability
- More fine-grained control of pipeline creation from the nf-core template
- Helper tools for subworkflows & nf-test support



Coming soon...

- Rewritten launch and schema tooling
- User friendly terminal app, using Textual



Channels: The Nextflow Podcast

Talking about news in the Nextflow ecosystem and speaking with pioneers in the field

 nextflow

Features Quick start Learn Nextflow Examples ▾ Community ▾ Blog Podcast About

 **CHANNELS**
the *nextflow* podcast

Talking about news in the Nextflow ecosystem and speaking with pioneers in the field.

Episode 26 [Code & community at the 2023 nf-core hackathon](#)

Hear from participants of the nf-core hackathon discussing their work and the value of community

1 Interview 21 November 2023

 Ep. 26 | November 21, 2023
Interviews
Code & community at the 2023 nf-core hackathon
Rajpal Manzana | Edmund Miller
Steve Hansen | Francesco Lescal

Stay Tuned Where to listen

 Listen on Spotify

 Listen on Apple Podcasts

 Listen on Google Podcasts

 Watch on YouTube



Nextflow Ambassadors

An initiative designed to foster collaboration, knowledge sharing, and community growth



Features Quick start Learn Nextflow Examples ▾ Community ▾ Blog Podcast About

Become an ambassador!

The Nextflow community is a dynamic ecosystem that brings together thousands of researchers, developers, and enthusiasts who are passionate about data-driven workflows and scientific computing. It thrives on the principles of open-source software, open science, and knowledge sharing. Here, you'll have the opportunity to connect with like-minded individuals from around the world, exchange ideas, and contribute to the advancement of cutting-edge bioinformatics, genomics, and data analysis pipelines. Whether you're a seasoned expert or just beginning your journey, the Nextflow community offers a supportive environment that values your insights and encourages your growth.

[Check our ambassadors here!](#)



Why become a Nextflow Ambassador?

The Nextflow ambassador program is a unique and exciting initiative designed to empower passionate individuals within the Nextflow community to play a more active role in fostering collaboration, knowledge sharing, and engagement. As a Nextflow Ambassador, you will have the opportunity to amplify the reach and impact of Nextflow by promoting its adoption, organizing community events, sharing your expertise through presentations and tutorials, and assisting fellow community members with their questions and challenges. This program is a way to recognise the dedicated volunteers that contribute their time, enthusiasm, and expertise to help grow and strengthen the Nextflow ecosystem. By becoming a Nextflow Ambassador, you'll not only enhance your own skills and network but also contribute to the advancement of scientific research and data analysis on a global scale.



Nextflow Ambassadors

An initiative designed to foster collaboration, knowledge sharing, and community growth



Features Quick start Learn Nextflow Examples ▾ Community ▾ Blog Podcast About

Nextflow Ambassadors



Abhinav Sharma



Adam Talbot



Alex Peltzer



Anabella Trigila



Anders Jemt



Bhargava Morampalli



Chris Hakkaart



Christopher Mohr



<https://community.seqera.io/>

The new home of the Nextflow community

The screenshot shows the homepage of the Seqera community website at <https://community.seqera.io/>. The page features a dark header with the Seqera logo and navigation links. A large, bold welcome message "Welcome to the community!" is centered above a search bar. Below the search bar, a sub-header encourages users to search before posting. On the left, a sidebar provides links to Topics, My Posts, Review, Admin, More, Categories, and specific sections for Seqera Platform (Tower) and Nextflow. The main content area displays a list of topics under the "Categories" tab, with filters for all categories, all tags, and latest topics. Two specific posts are highlighted: one about the Seqera Platform (Tower) and another about Nextflow configuration.

seqera

Welcome to the community!

We're happy to have you here. If you need help, please search before you post.

Search

Topics •

My Posts

Review

Admin

More

Categories

Seqera Platform (Tower)

Collaborative data analysis at scale, on-premises or in any cloud.

Nextflow

8 / month

1 / month

all categories ▾

all tags ▾

Categories

Latest

New (2)

Unread (1)

Top

Filter

+ New Topic

Category

Topics

Latest

Nextflow configuration for sequential and parallel execution in multipod •

Help

0 3h

<https://community.seqera.io/>

The new home of the Nextflow community

The screenshot shows the homepage of the Nextflow community website. At the top, there is a navigation bar with links for Home, Pipelines, Resources, Docs, Events, About, a search bar, and social media icons for GitHub, Twitter, YouTube, Mastodon, and Slack. A prominent green header section features the text "Join nf-core" and a subtext "Read about the different ways you can get involved with nf-core". Below this, a white content area contains a message about community organization tools like GitHub, Twitter, Mastodon, YouTube, and Slack. It also includes a warning about the code of conduct and a note about where to ask Nextflow-related questions. On the right side, there is a sidebar titled "On this page" with links to various community resources.

Home Pipelines Resources ▾ Docs Events ▾ About ▾

Search

Join nf-core

Read about the different ways you can get involved with nf-core

We use a few different tools to organise the nf-core community - you are welcome to join us at any or all!

GitHub twitter mastodon YouTube Slack

All nf-core community members are expected to adhere to the nf-core [code of conduct](#)

If your question is about Nextflow and not directly related to nf-core, please use the [Slack community chat](#) or the [discussion forum](#) on GitHub.

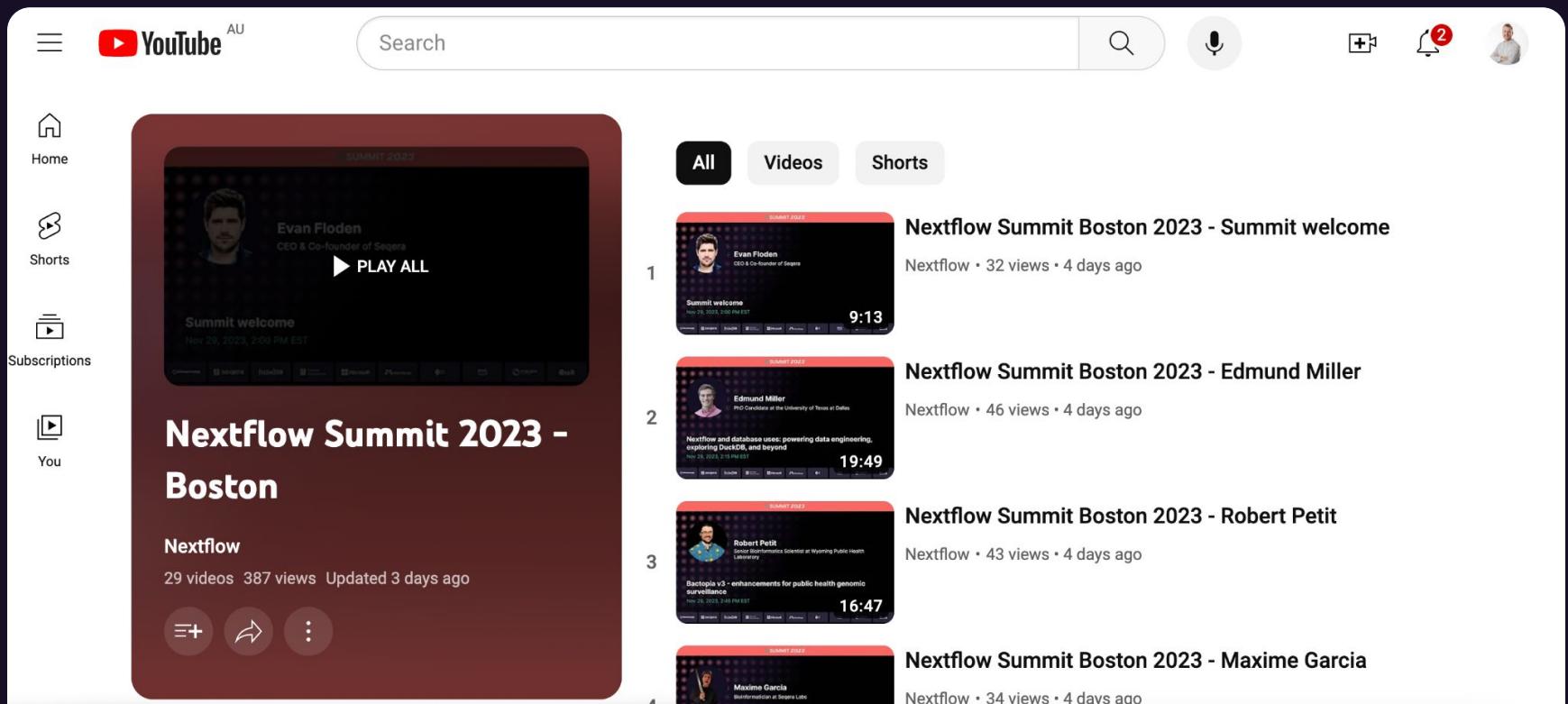
On this page

- Slack
- GitHub organisation
- Twitter
- Mastodon
- YouTube
- Gather Town

[Back to top](#)

Nextflow Summit - Barcelona and Boston 2023

In-person and online



The image shows a screenshot of a YouTube search results page. The search query "Nextflow Summit" has been entered into the search bar. The results are filtered by "All" media types, showing a mix of Videos and Shorts. The top result is a channel page for "Nextflow" featuring a video thumbnail for "Summit welcome" by Evan Floden. Below the channel page are four video thumbnails for individual sessions: "Nextflow Summit Boston 2023 - Summit welcome" by Edmund Miller, "Nextflow Summit Boston 2023 - Edmund Miller" by Robert Petit, "Nextflow Summit Boston 2023 - Robert Petit" by Maxime Garcia, and "Nextflow Summit Boston 2023 - Maxime Garcia" by Nextflow.

YouTube AU

Search

All Videos Shorts

1 9:13

2 19:49

3 16:47

4

Nextflow Summit Boston 2023 - Summit welcome

Nextflow • 32 views • 4 days ago

Nextflow Summit Boston 2023 - Edmund Miller

Nextflow • 46 views • 4 days ago

Nextflow Summit Boston 2023 - Robert Petit

Nextflow • 43 views • 4 days ago

Nextflow Summit Boston 2023 - Maxime Garcia

Nextflow • 34 views • 4 days ago

Home

Shorts

Subscriptions

You

Nextflow

29 videos 387 views Updated 3 days ago

Summit welcome

Nov 29, 2023, 2:00 PM EST

Evan Floden

CEO & Co-founder of Seqera

PLAY ALL

Summit welcome

Nov 29, 2023, 2:00 PM EST

Edmund Miller

PhD Candidate at the University of Texas at Dallas

Nextflow and database uses: powering data engineering, exploring DuckDB, and beyond

Nov 29, 2023, 1:15 PM EST

Robert Petit

Senior Bioinformatics Scientist at Wyoming Public Health Laboratory

Bactopia v3 - enhancements for public health genomic surveillance

Nov 29, 2023, 1:45 PM EST

Maxime Garcia

bioinformatician at Seqera Labs

Nextflow Summit 2023 - Boston

Participate

Seminars, training, hackathons, and more

- Bytesize seminars
- Training
- Hackathons
- Social media
- Blogs
- Community Slack
- Documentation
- Swag
- Mentorships





 nextflow

Sign up for updates to be the first to know
when tickets go on sale!

<https://summit.nextflow.io>

Online: March (TBD)
Boston: May 14-17
Barcelona: October 21-25



Phil Ewels, Ph.D.

PM for Open Source at [Seqera](#)

phil.ewels@seqera.io
@tallphil | Stockholm



Maxime Garcia, Ph.D.

Bioinformatics Engineer at [Seqera](#)

maxime.garcia@seqera.io
@gau | Stockholm

Thank you

