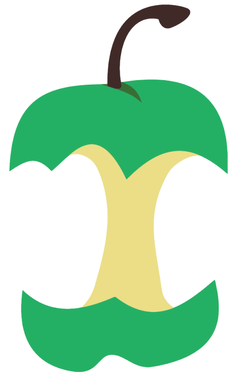
nf-core/bytesize



Nextflow & nf-core Survey 2022

Phil Ewels, Seqera Labs

Seqera Labs
@SeqeraLabs

State of the Workflow 2022 📋📋

Have your say and support Nextflow and nf-core in the annual community survey. Plus, win a Macbook Air M1 or one of ten merchandise packs. Enter now!

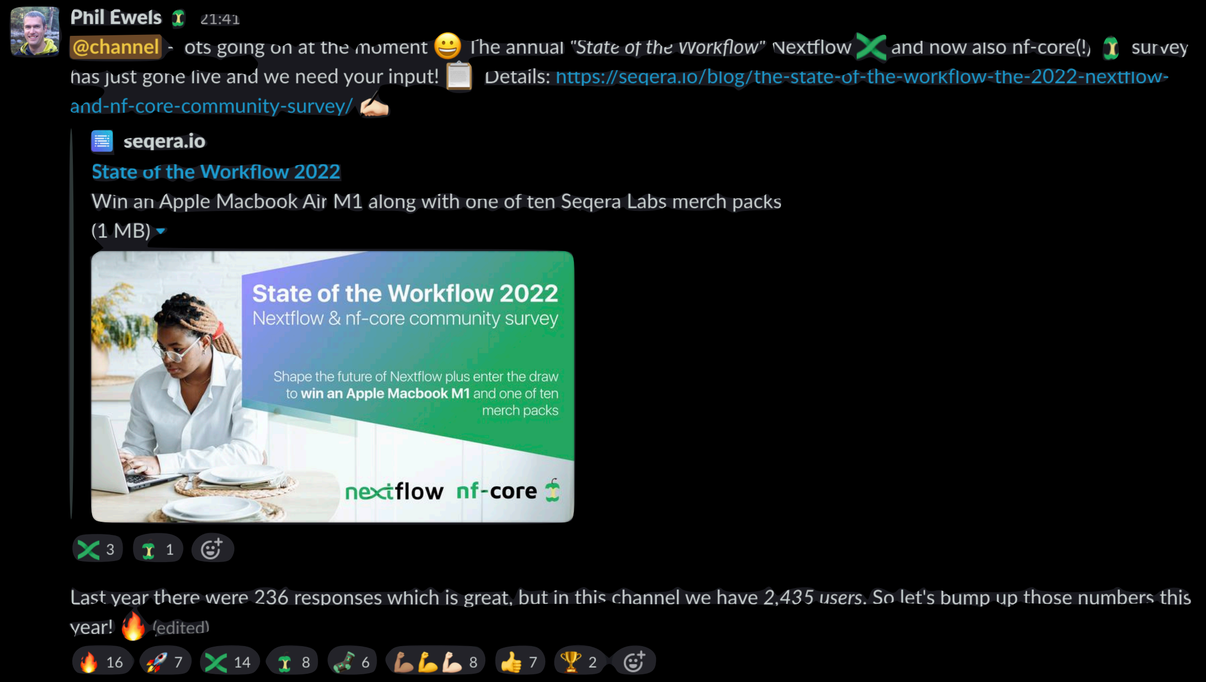
seqera.io
State of the Workflow 2022
Win an Apple Macbook Air M1 along with one of ten Seqera Labs merch packs

nf-core
@nf_core

These annual @nextflowio surveys are really helpful in shaping development efforts 📋

@nf_core joins this year, measuring impact for our @cziscience diversity grant 🌍🌎🌏

Please take a moment to fill this in, even if you don't use the tools a lot. It makes a big difference. 🙏



Evan Floden | Feb 08, 2022

State of the Workflow 2022: The Nextflow and nf-core community survey

NEXTFLOW NF-CORE COMMUNITY ANNOUNCEMENT

The Nextflow community has continued to grow over the past year with over 10,000 scientists, engineers and developers now reading our documentation each month! We are always eager to learn more about who makes up the community and their requirements.

Take the survey now

Nextflow and nf-core

PROJECT Nextflow, nf-core
PARTNER ORGANIZATION Seqera Labs
FUNDING CYCLE D&I

PROPOSAL SUMMARY
ASSOCIATED GRANT
PROJECT
PARTNER ORGANIZATION

Proposal Summary

To keep the current momentum of initiatives and push forward with new actions for accessibility, internationalization, mentorships and ambassadors.

Associated Grant

Nextflow and nf-core: Reproducible Workflows for the Scientific Community

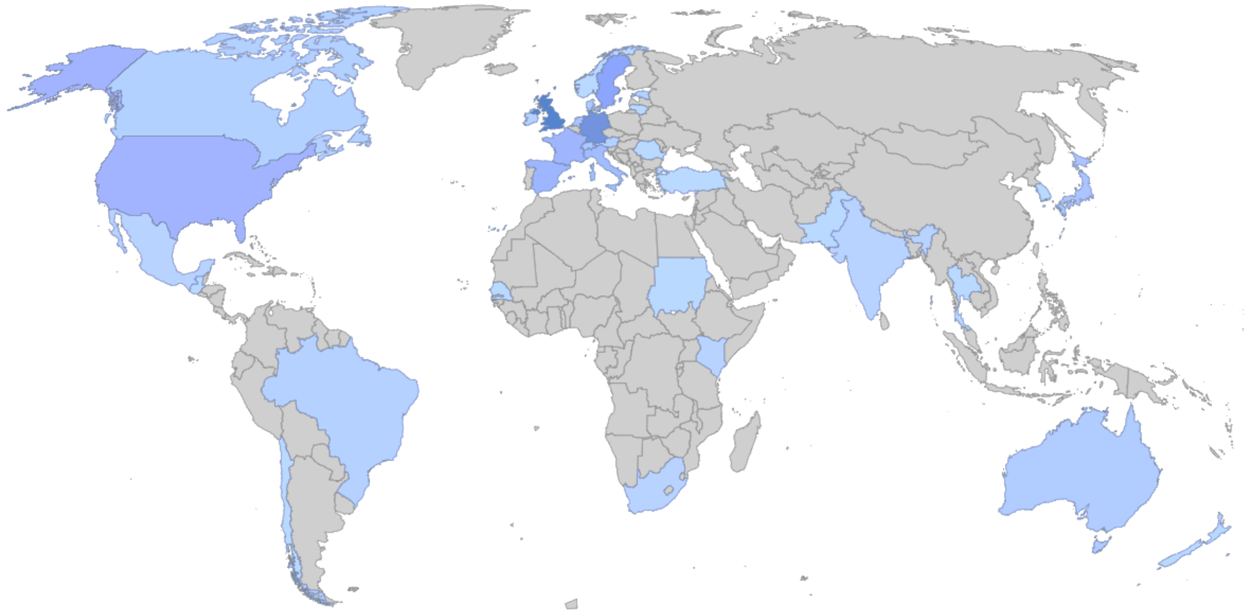
Project

Nextflow, nf-core

Proposed Work

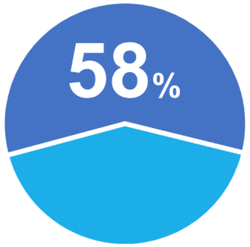
Chan Zuckerberg Initiative

Demographics



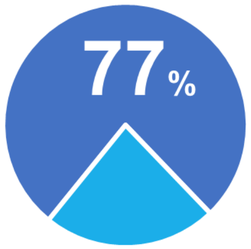
26.0%	4.4%	2.1%	1.0%
19.3%	4.2%	1.8%	1.0%
13.8%	3.4%	1.6%	0.8%
8.1%	2.1%	1.3%	0.8%

LANGUAGE



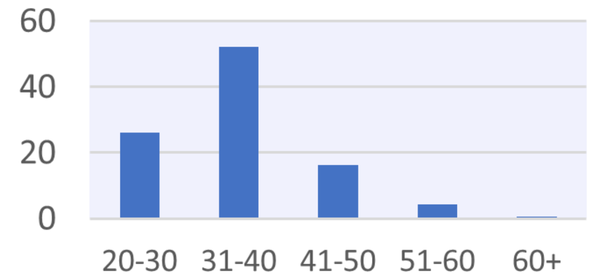
Most speak **English as** their first language

GENDER BREAKDOWN



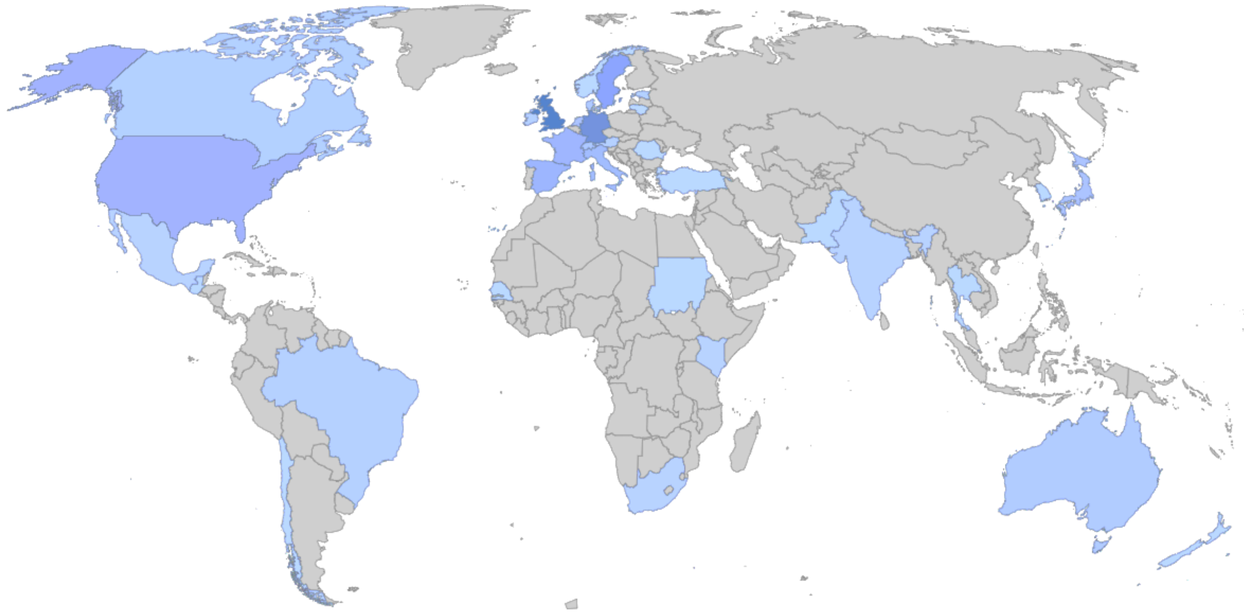
Most Nextflow users are male

AGE OF NEXTFLOW USERS



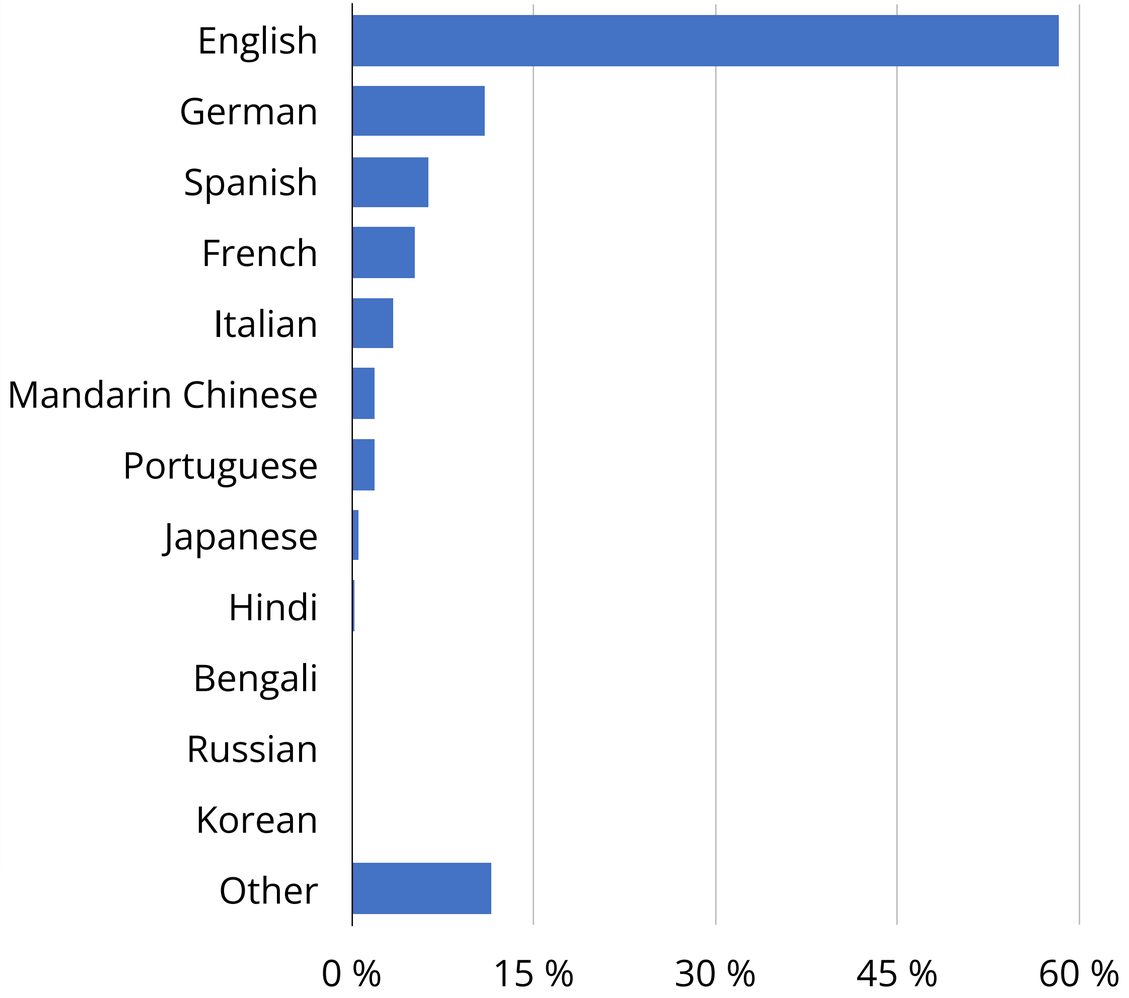
~50% of users are between 31 and 40 years old

Demographics



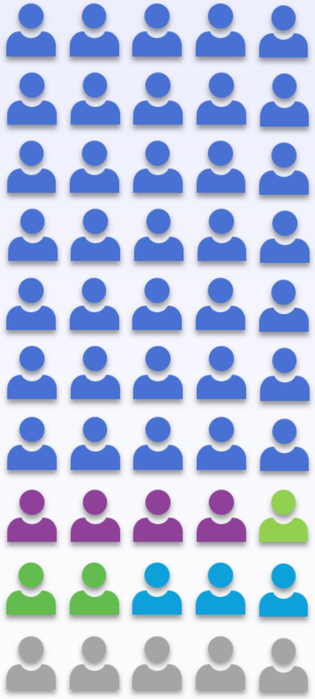
26.0% 4.4% 2.1% 1.0%
19.3% 4.2% 1.8% 1.0%
13.8% 3.4% 1.6% 0.8%
8.1% 2.1% 1.3% 0.8%

Most proficient written or spoken language



Demographics

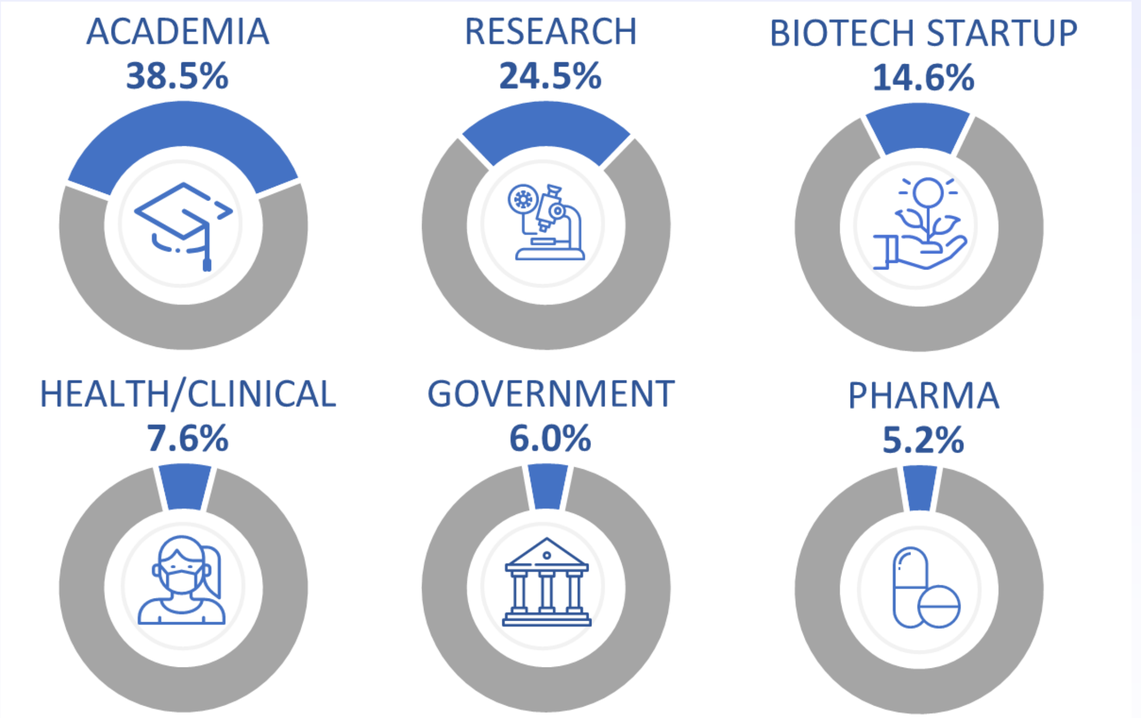
WHO IS USING NEXTFLOW?



Bioinformaticians	70.6%
Principal Investigators	7.3%
Software Engineers	6.5%
Data Scientists	6.3%
Others	9.3%

Most Nextflow users are
Bioinformaticians

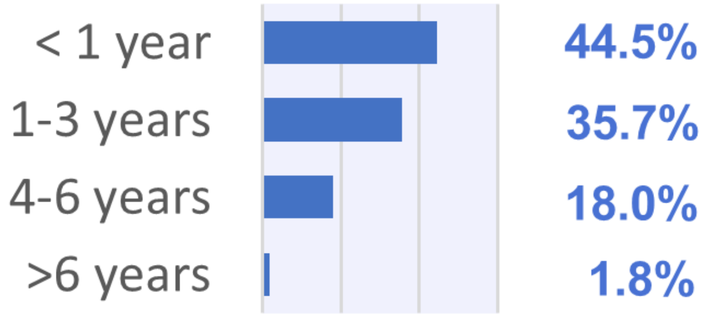
USERS BY INDUSTRY



* multiple options possible

Nextflow Usage

RAPID ADOPTION



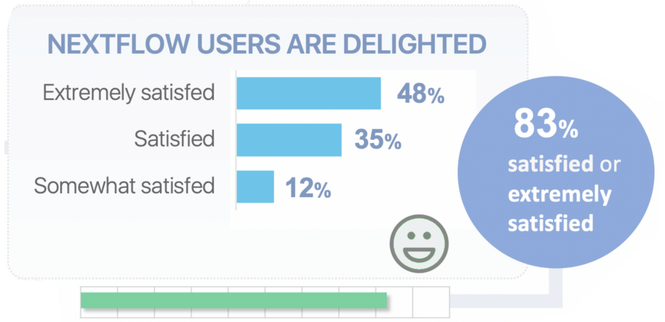
Almost half of Nextflow users are new in the last year

HAPPY USERS

98% user satisfaction rate

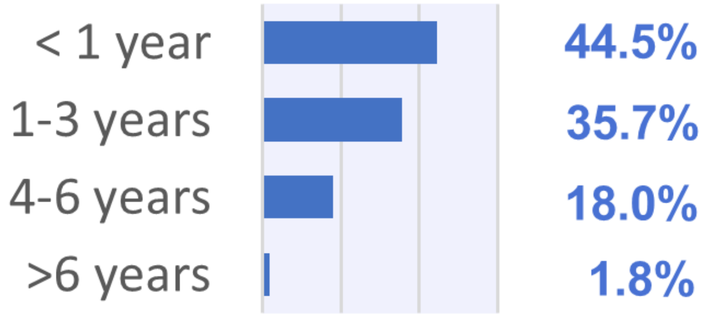
96% would recommend Nextflow to a colleague

2021



Nextflow Usage

RAPID ADOPTION



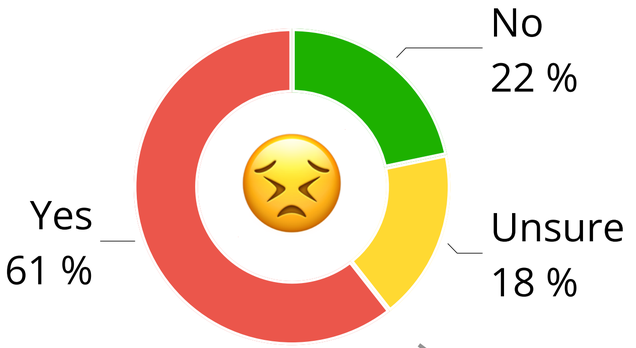
Almost half of Nextflow users are new in the last year

HAPPY USERS

98% user satisfaction rate

96% would recommend Nextflow to a colleague

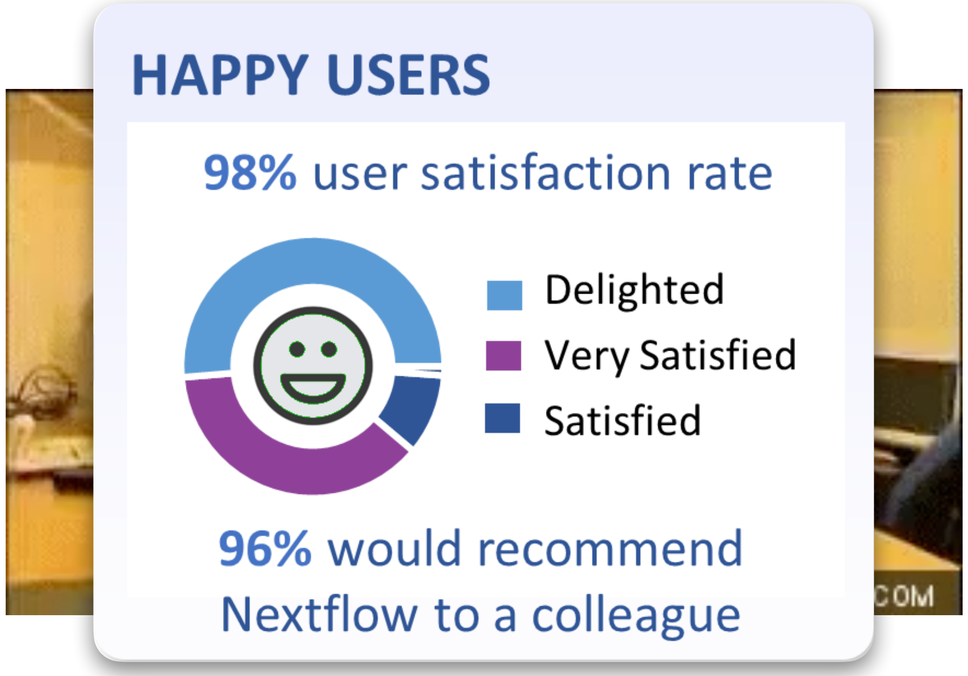
Have you felt frustrated with Nextflow?



Writing Groovy
Difficulty interpreting error messages
Steep learning curve
Speed of change

Nextflow Usage

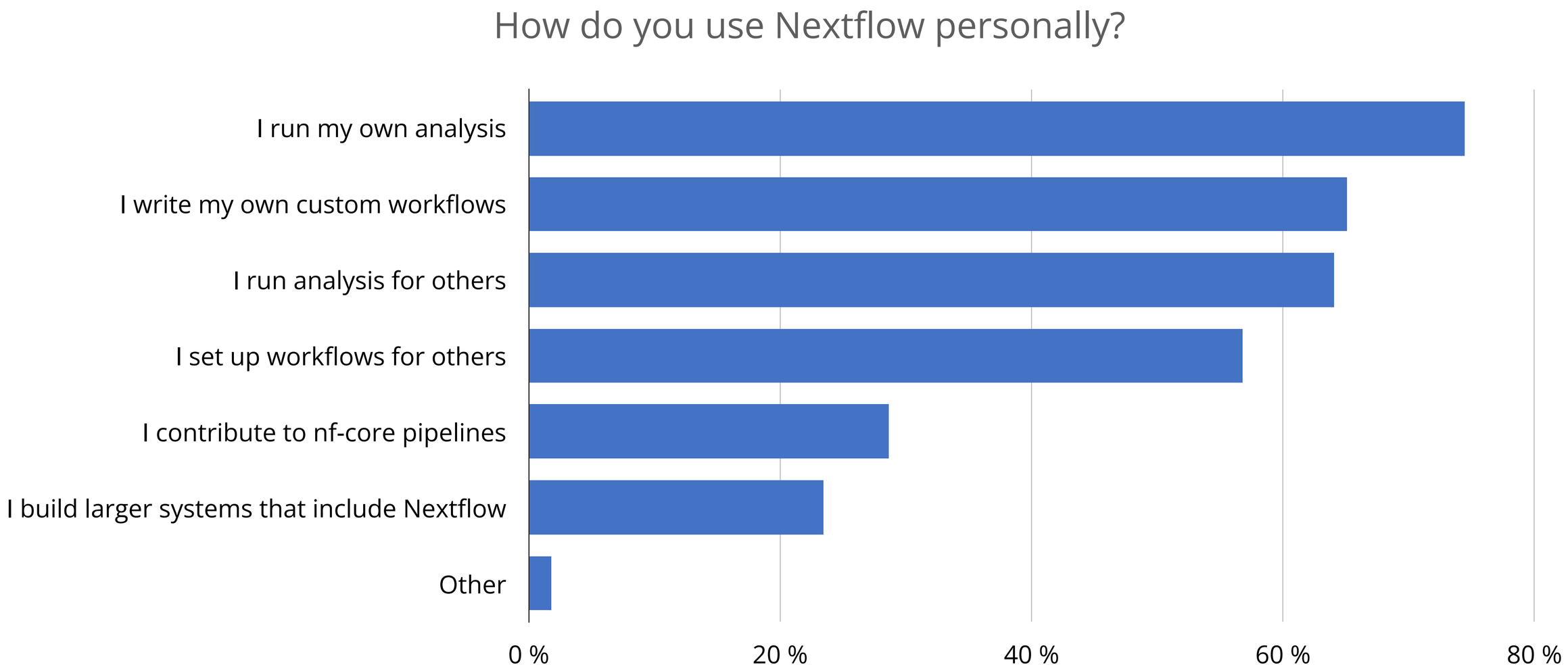
nf-core/ frustration



nextflow
errors
difficult
learning
documentation
syntax
dsl2
pipelines
messages
nf-core
hard
groovy
language
debugging
curve
examples
channels
time
code
understanding
workflows
frustrating
pipeline
lack
steep
learn
resume
issues
dsl
process
processes
tower
channel
parameters
debug
tool
finding
file
inability
tools
seems
number
experience
parameter
within
another
keep
cryptic
things
get
work
docs
using
found
beginners
without
reason
certain
tricky
just
it's
outdated
troubleshoot
compatibility
different
run
workflow
even
trying
concepts
management
environments
resources
misleading
confusing
data
feel
understand
implementing
name
unit
dsl1
example
mix
getting
extremely
changes
gets
complex
small
find
setting
multiple
helpful
might
right
system
challenging
real
logic
outside
access
good
helped
cloud
convert
much
script
little
support
troubleshooting
crashes
perhaps
group
github
w/o
intuitive
feels
files
around
copy
annoying
compared
clear
reports
testing
written
gitter
adopted
message
functionality
capabilities
though
always
point
easy
long
validation
flow
specific
make
cases
tutorials
operators
feature
version
know
basics
going
documented
input
able
seem
samples
needs
python
second
structures
dataflow-datatypes
nextflow-console
harness
commercial
command.run
overwhelming
user-defined
adaptation
done
confuse
parallelism
correct
metadata
across
manually
reference
someone
amount
hidden
forget
necessary
programming
structured
near
competent
almost
handle
bash
features
every
unhelpful
running
large
like
makes
lacking
snakemake
help
nice
failed
frequently
java
easily
far
straight
fail
operations
directories
original
problems
people
missing
implementation
https://github.com/nextflow-io/nextflow/issues/1702
late
undocumented
mastering
processed
wait/wait
architecture
beginning
file-system
contents
changing
languages
hackable
jump
generic
madness
bioinformatics
groovy...can't
demonstration
programmed
handling
sufficient
bully
traceback
spamming
double
insist
wrong
course
providing
uniform
groupings
mentioned
combining
scarcity
production
workflow-on
additional
migration
majority
somebody
transitioning
causing
unpredictable
workflow-for
unpredictably
unstably
fast
instead
script/process/channel
create
happens
inconsistent
looking
mostly
custom
cluster
took
useful
obvious
flexibility
folder
times
tests
properly
error-causing
download/run
wrong
updated
bioinformatic
integrate
question
resolving
overlooked
launcher
displayed
overnight
explained
perfect
developer
execution
groovy/java
problematic
light
coding/scripting
propagated
troubleshoot
download/run
wrong
error-causing
compatibility
different
success
per
using
outdated
test
it's
familiar
something
frustrating
workflows
just
found
understanding
code
certain
reason
without
beginners
huge
container
abstract
sample
tricky
coding/steps
ec2
professional
overlooked

Nextflow Usage

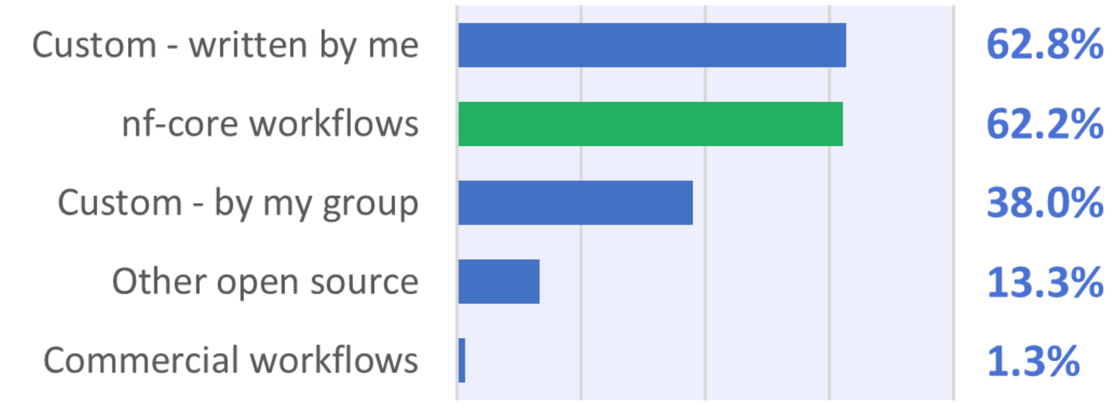
* multiple options possible



Nextflow Usage

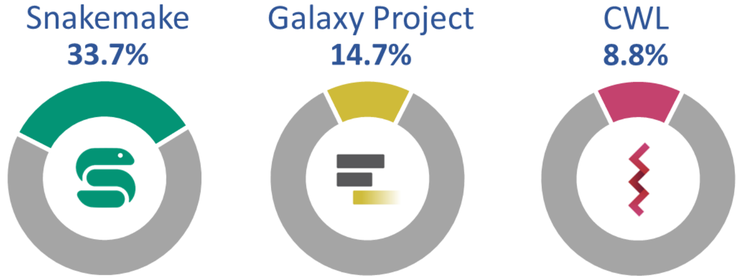
* multiple options possible

WHAT TYPES OF WORKFLOWS DO YOU USE?



nf-core pipelines are used by **62.2%** of respondents

NEXTFLOW USERS ALSO RUN:

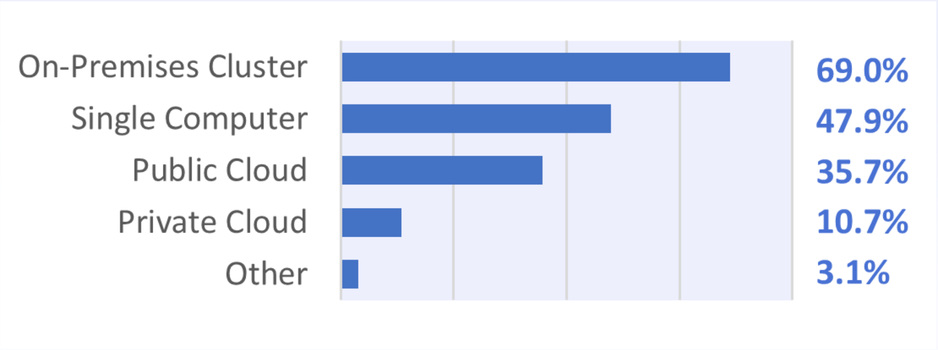


Half of users use more than one tool.
Others include Cromwell, Spark, and Airflow

Nextflow Usage

* multiple options possible

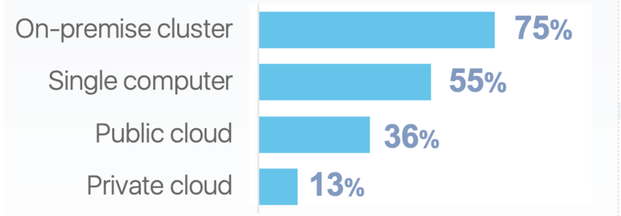
WHERE DO YOU RUN NEXTFLOW?



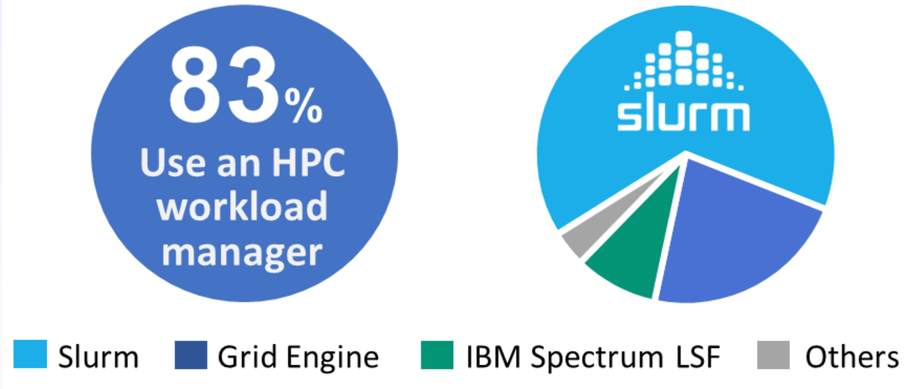
On-premises HPC clusters remain the dominant runtime platform for Nextflow

2021

WHERE DO USERS RUN NEXTFLOW?



SLURM RULES THE *HPC ROOST*

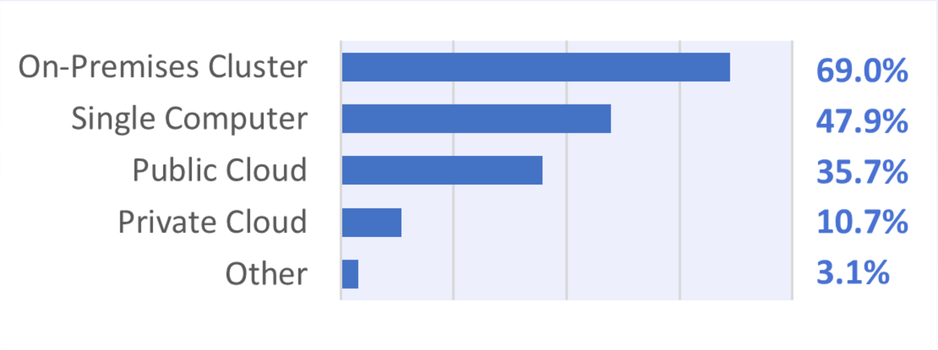


Slurm is used to manage **53%** of all Nextflow Clusters

Nextflow Usage

* multiple options possible

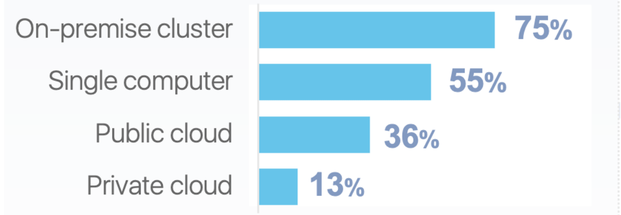
WHERE DO YOU RUN NEXTFLOW?



On-premises HPC clusters remain the dominant runtime platform for Nextflow

2021

WHERE DO USERS RUN NEXTFLOW?

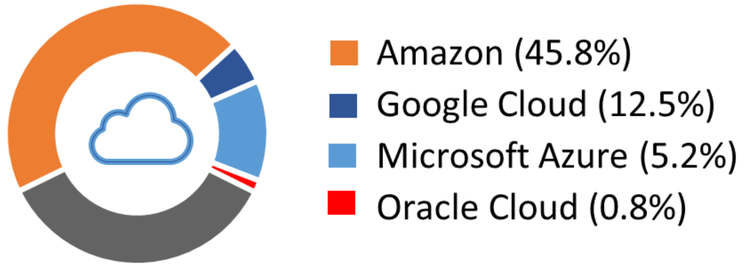


WORKLOADS SHIFT TO THE CLOUD

13.4% of Nextflow users use more than one public cloud [1]

PREFERRED CLOUDS

Among users planning a move to the cloud within 2 years, most plan to use AWS

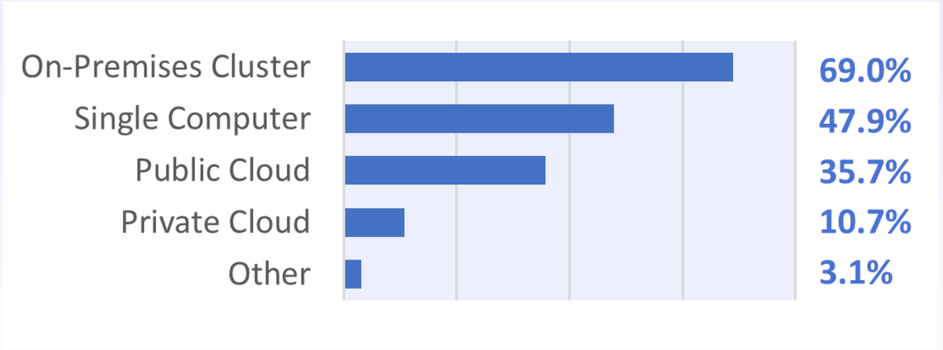


Public cloud adoption is higher among private sector users – **77% use public cloud** [2]

Nextflow Usage

* multiple options possible

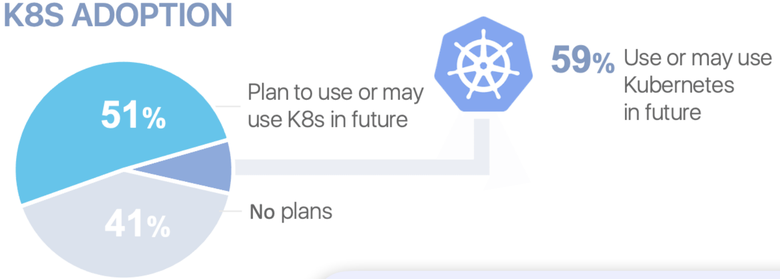
WHERE DO YOU RUN NEXTFLOW?



On-premises HPC clusters remain the dominant runtime platform for Nextflow

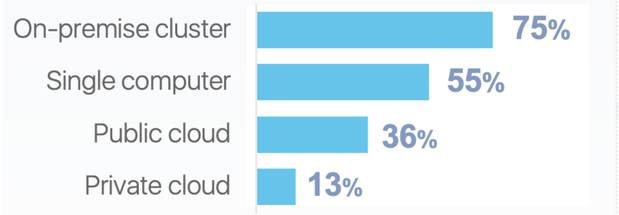
2021

K8S ADOPTION



2021

WHERE DO USERS RUN NEXTFLOW?



LIMITED KUBERNETES ADOPTION

Kubernetes
7.2%

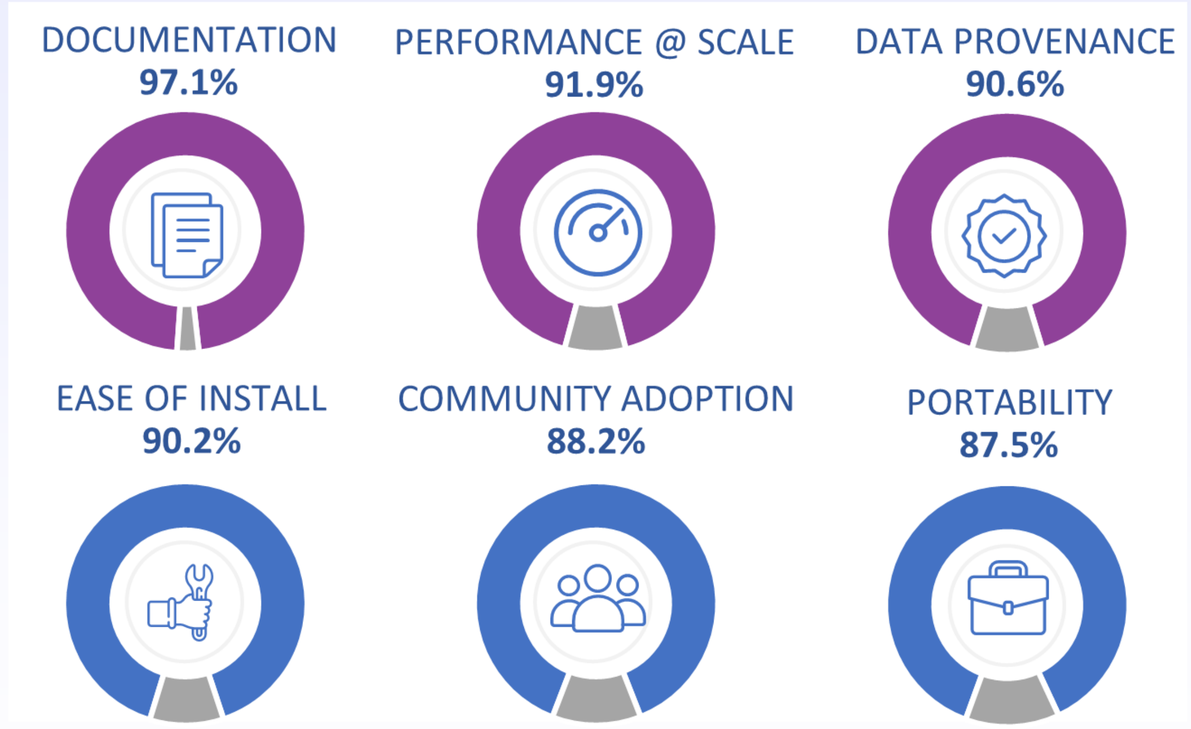
Kubernetes adoption was unchanged from 2021

Amazon EKS, Google GKE, Azure AKS, OpenShift, and Rancher

Nextflow Usage

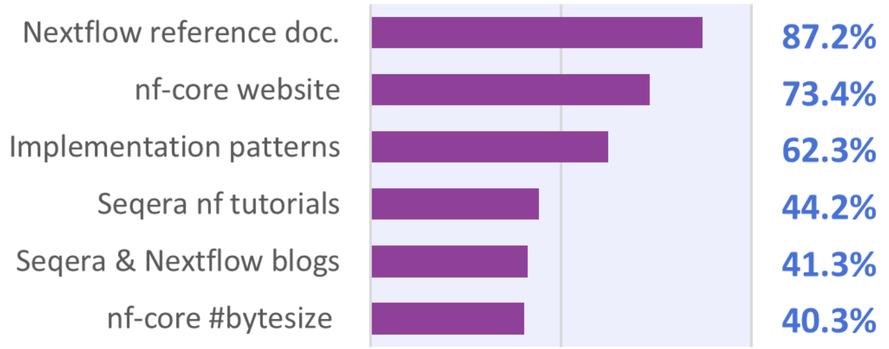
* multiple options possible

WHAT IS IMPORTANT IN A WORKFLOW MGR?



Percentage of users that view this as **important** or **very important**

WHAT MATERIALS ARE MOST USEFUL?



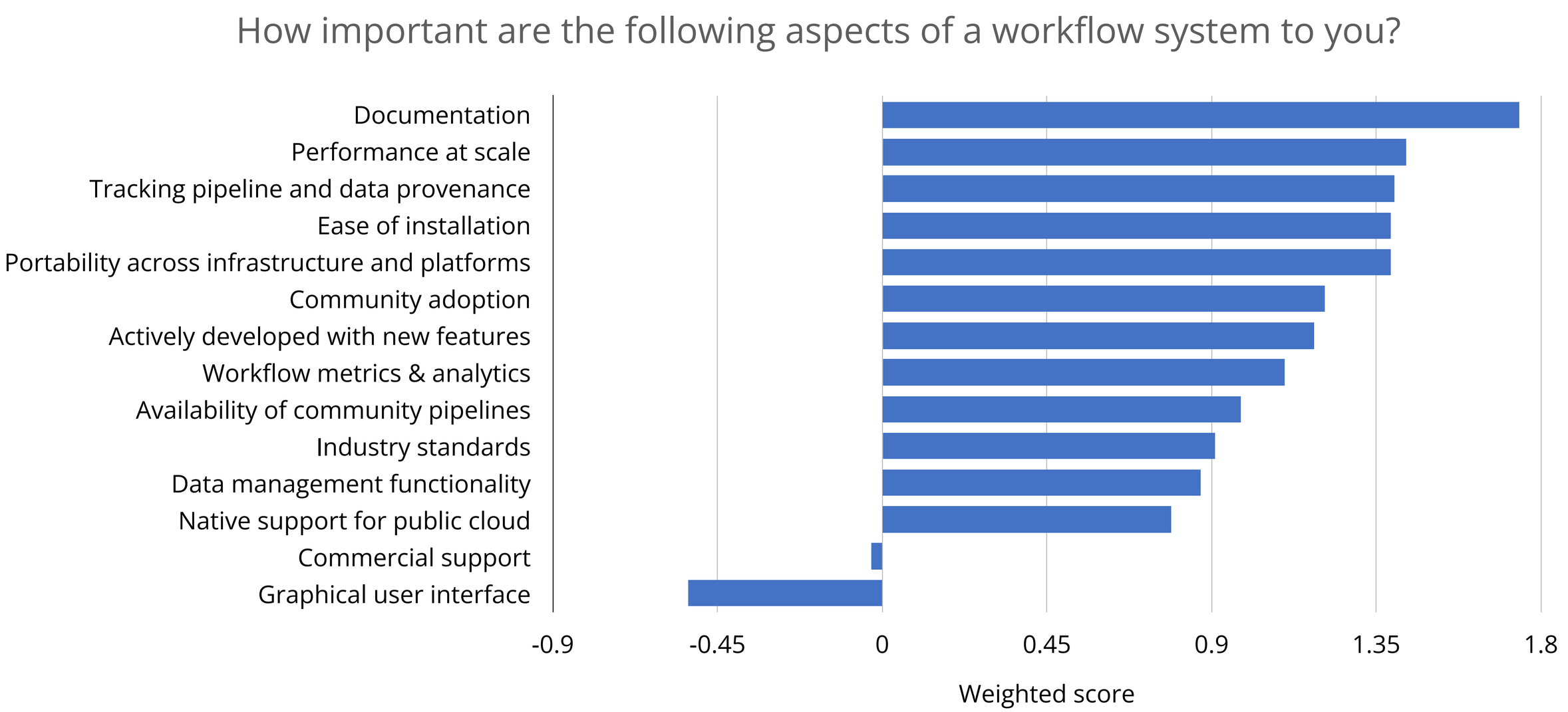
Percentage of respondents viewing materials as ***important*** or ***very important***

67% Use the nf-core Slack Channel for support

29% Contribute to community pipelines

Nextflow Usage

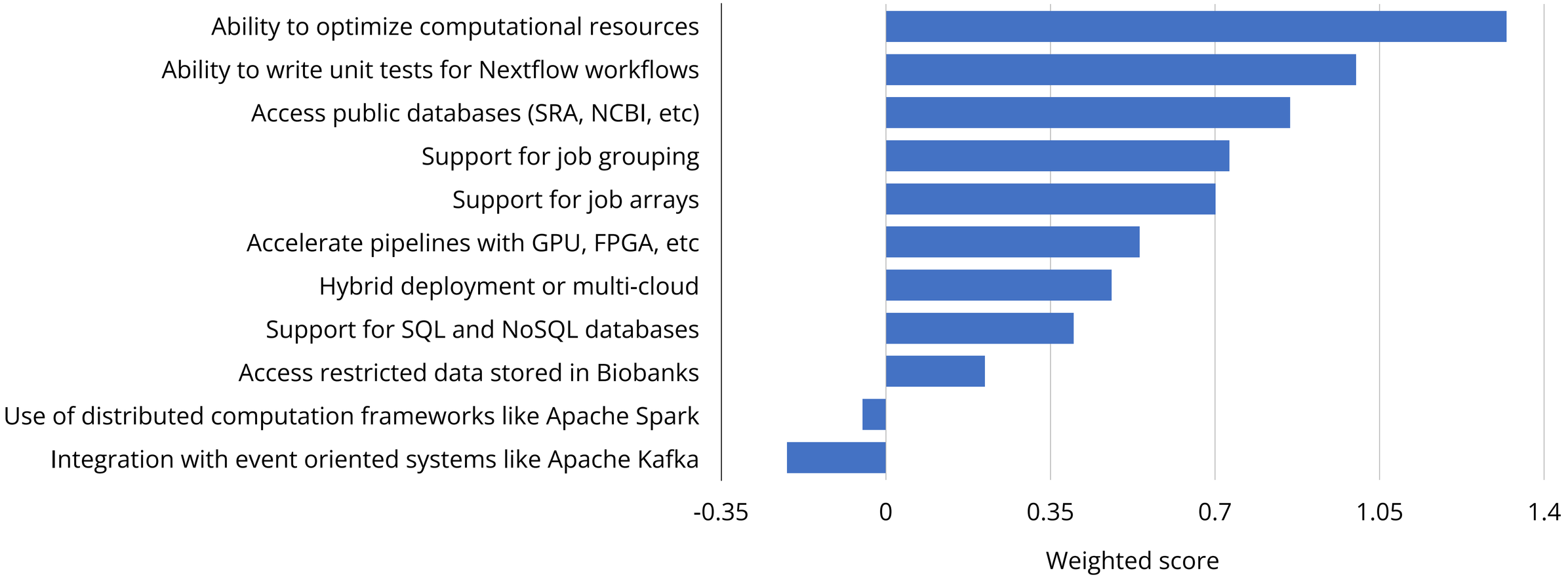
* multiple options possible



Nextflow Usage

* multiple options possible

How important do you consider the following integrations or technologies for your current or future pipelines?



Read the blog post →

https://seqera.io/blog/state-of-nextflow-2022-results/

Nextflow & nf-core Survey 2022

Phil Ewels, Seqera Labs

Prizes!

1 x Apple Macbook Air M1

10 x Sustainability-certified Seqera Labs merchandise packs

SurveyMonkey