Cloud-SPAN Handbook

Contents

Welcome! **5**

Handbook Contents 5

1 Introduction **7**

1.1 About Cloud-SPAN 7

1.2 About this handbook 7

1.3 Our Team 8

2 Code of Conduct **9**

2.1 Preamble 9

2.2 Our Code of Conduct 9

2.3 Incident Reporting Guidelines 11

2.4 Acknowledgements 12

3 Our Courses **13**

3.1 Prenomics Course 13

3.2 Genomics Course 14

3.3 Future Courses 14

4 The Cloud-SPAN commmunity **15**

4.1 Ways to contribute 15

4.2 Using GitHub to contribute 16

5 FAIR Principles **17**

5.1 What is FAIR data? 17

6 Cloud-SPAN Online Forum **19**

6.1 Signing up for Peerboard 19

Welcome!

Cloud-SPAN

Cloud-based High Performance Computing
SPecialised **AN**alyses for 'omics

Welcome to the Cloud-SPANhandbook! It's great to have you here!

Handbook Contents

Introduction
Code of Conduct Our Courses The Cloud-SPAN Community
FAIR Principles
Cloud-SPAN Online Forum

Chapter 1

Introduction

Welcome to the Cloud-SPAN Community Handbook! It's great to have you here.

1.1 About Cloud-SPAN

Cloud-SPAN deploys high quality learning resources that will train researchers to effectively generate and analyse a range of 'omics data using Cloud computing resources. Read more about our courses.

Cloud-SPAN is a collaboration between the Department of Biology at the University of York and the Software Sustainability Institute, and funded by the UKRI innovation scholars award under project reference MR/V038680/1.

1.2 About this handbook

This handbook is intended as a reference for both the core Cloud-SPAN team (see below) and for our wider community of learners. It's where you'll find our Code of Conduct, contributing guidelines and other practical information which will help you make the most of our resources in a friendly, understanding environment.

For example, did you know that we have an online forum for questions, suggestions and any concerns you might have about applying what you learn on our courses? We strongly encourage you to engage with the Cloud-SPAN community to enhance your learning and understanding. You can find the forum here or read more about using Peerboard later in this manual.

1.3 Our Team

1.3.1 Project team

Name	Role	% working on the project	Time period allocated to the project	Institution
Emma Rand	Project oversight	20%	2020-22	Uni of York
James Chong	Project oversight	10%	2020-22	Uni of York
Jorge Buenabad-Chavez	Content developer and Cloud deliverer	100%	2020-22	Uni of York
Evelyn Greeves	Content developer and FAIR/CoP deliverer	60%	2020-22	Uni of York
Sarah Dowsland	Project Manager	50%	2020-22	Uni of York
Annabel Cansdale	Content developer	20%	2020-22	Uni of York
Sarah Forrester	Content developer	20%	2020-22	Uni of York
Neil Chue Hong	Strategic advisor	10%	2020-22	Software Sustainability Institute
Emma Barnes	Project oversight	20%	2020-22	Uni of York
Dan Bishop	Strategic input	10%	2020-22	Uni of York

Chapter 2

Code of Conduct

2.1 Preamble

The CloudSPAN team are dedicated to providing a welcoming and supportive environment for all people, regardless of background or identity. As such, we do not tolerate behaviour that is disrespectful to our community members or that excludes, intimidates, or causes discomfort to others. We do not tolerate discrimination or harassment based on characteristics that include, but are not limited to: gender identity and expression, sexual orientation, disability, physical appearance, body size, citizenship, nationality, ethnic or social origin, pregnancy, familial status, veteran status, genetic information, religion or belief (or lack thereof), membership of a national minority, property, age, education, socio-economic status, technical choices, and experience level.

Everyone who participates in CloudSPAN project activities is required to conform to this Code of Conduct. This Code of Conduct applies to all spaces managed by the CloudSPAN project including, but not limited to, in person focus groups and workshops, and communications online via GitHub. By participating, contributors indicate their acceptance of the procedures by which the project core development team resolves any Code of Conduct incidents, which may include storage and processing of their personal information.

2.2 Our Code of Conduct

We are confident that our community members will together build a supportive and collaborative atmosphere at our events and during online communications. The following bullet points set out explicitly what we hope you will consider to be appropriate community guidelines:

- **Be respectful of different viewpoints and experiences.** Do not engage in homophobic, racist, transphobic, ageist, ableist, sexist, or otherwise exclusionary behaviour.
- **Use welcoming and inclusive language.** Exclusionary comments or jokes, threats or violent language are not acceptable. Do not address others in an angry, intimidating, or demeaning manner. Be considerate of the ways the words you choose may impact others. Be patient and respectful of the fact that English is a second (or third or fourth!) language for some participants.
- **Do not harass people.** Harassment includes unwanted physical contact, sexual attention, or repeated social contact (see below for an extended list of behaviours we consider to be harassment). Know that consent is explicit, conscious and continuous—not implied. If you are unsure whether your behaviour towards another person is welcome, ask them. If someone tells you to stop, do so.
- **Respect the privacy and safety of others.** Do not take photographs of others without their permission. Do not share other participant's personal experiences without their express permission. Note that posting (or threatening to post) personally identifying information of others without their consent ("doxing") is a form of harassment.
- **Be considerate of others' participation.** Everyone should have an opportunity to be heard. In update sessions, please keep comments succinct so as to allow maximum engagement by all participants. Do not interrupt others on the basis of disagreement; hold such comments until they have finished speaking.
- **Don't be a bystander.** If you see something inappropriate happening, speak up. If you don't feel comfortable intervening but feel someone should, please feel free to ask a member of the Code of Conduct response team for support.
- As an overriding general rule, **please be intentional in your actions and humble in your mistakes**.

All interactions should be professional regardless of platform: either online or in-person. See this explanation of the four social rules - no feigning surprise, no well-actually's, no back-seat driving, no subtle -isms - for further recommendations for inclusive behaviours.

2.2.1 Unacceptable Behaviour

Examples of unacceptable behaviour by community members at any project event or platform include:

- written or verbal comments which have the effect of excluding people on the basis of membership of any specific group
- causing someone to fear for their safety, such as through stalking, following, or intimidation
- violent threats or language directed against another person
- the display of sexual or violent images
- unwelcome sexual attention
- nonconsensual or unwelcome physical contact
- sustained disruption of talks, events or communications
- insults or put downs
- sexist, racist, homophobic, transphobic, ableist, or exclusionary jokes
- excessive swearing
- incitement to violence, suicide, or self-harm
- continuing to initiate interaction (including photography or recording) with someone after being asked to stop
- publication of private communication without consent

CloudSPAN prioritises marginalised people's safety over privileged people's comfort. We will not act on complaints regarding:

- 'Reverse' -isms, including 'reverse racism,' 'reverse sexism,' and 'cisphobia'.
- Reasonable communication of boundaries, such as "leave me alone," "go away," or "I'm not discussing this with you."
- Communicating in a 'tone' you don't find congenial.
- Criticism of racist, sexist, cissexist, or otherwise oppressive behavior or assumptions.

2.3 Incident Reporting Guidelines

2.3.1 Contact points

If you feel able to, please contact Emma Rand by email at emma.rand@york.ac.uk

2.3.2 Alternate contact points

If you do not feel comfortable contacting Emma Rand, please report an incident to Evelyn Greeves by email at evelyn.greeves@york.ac.uk

2.4 Acknowledgements

This Code was adapted from the Turing Way Code of Conduct, which itself draws from the Carpentries and Alan Turing Institute Data Study Group codes of conduct. Both are licensed for reuse under a CC BY 4.0 CA license.

Material was additionally drawn from the R Community Diversity, Equity, and Inclusion Working Group, also licensed under CC BY 4.0 CA.

Chapter 3

Our Courses

Our courses provide a thorough and well-paced introduction to the use of cloud computing resources (via Amazon Web Services) for genomic analysis.

We currently offer two foundational courses (Prenomics and Genomics) and are developing further courses on topics such as launching your own cloud instance and metagenomic analysis. A full learning path will be revealed soon.

All of our courses are offered **free of charge**.

3.1 Prenomics Course

'Prenomics' is an interactive online course on understanding file systems and using the command line which takes place over 2 half days (roughly six hours of content). We developed this course after finding that people taking the Genomics course vary in their experiencing of navigating file systems and the command line (shell).

Topics covered include file directory structure, logging onto a cloud instance, basic shell commands and using the shell to manipulate and search files.

NO prior experience is necessary.

This course will next run in mid-November. However, all training resources remain available online for self-study purposes.

Next course dates:
Tue 15th November, 10am-1pm
Weds 16th November, 10am-1pm

3.2 Genomics Course

'Genomics' is a practical, tutor-led course taking place over 4 half days or 2 full days (roughly twelve hours of content). It teaches data management and analytical skills using cloud resources for genomic research. Previously it has been held online due to Covid-19 but the next Genomics course will be held in-person.

Topics covered include project management for cloud genomics, writing shell scripts, automating workflows, assessing read quality, trimming/filtering reads and finding sequence variants.

A basic knowledge of the shell is required, which can be obtained by attending/self-studying the Prenomics course.

This course will next run in mid-November. However, all training resources remain available online for self-study purposes.

Next course dates:
Tues 22nd November, 9:30am-4:30pm
Weds 23rd November, 9:30am-4:30pm

3.3 Future Courses

3.3.1 Create your own AWS instance

This mostly online course will be on [creating and managing an Amazon Web Service instance(https://cloud-span.github.io/create-aws-instance-0-overview/), intended to guide participants through the process of setting launching their own cloud instance in order to:

- study course materials further
- perform more complex analyses on their own data.

3.3.2 Metagenomics

This mostly online course will focus on analysis methods and tools used in metagenomics.

Chapter 4

The Cloud-SPAN commmunity

Under construction

Our aim is to build a friendly and involved community of people who have used our resources, are interested in our resources, or who have expertise in the areas we cover.

That means that whether you are…

an expert in 'omics analyses
a complete newbie at 'omics
not entirely sure what an "omic" is
an experienced Cloud user
a little bewildered by the Cloud

…then the Cloud-SPAN community is for you!

There are lots of ways to contribute and we welcome all of them! Contributing, and joining our community, doesn't have to mean writing technical code.

Here are some ideas of ways you can contribute:

4.1 Ways to contribute

Learn

- Attend or work through our Foundations in Genomics course.
- Ask questions on our community forum.

Connect

- Join our *Community of Practice*.

Help

- Answer questions on our community forum.
- Tell us about bugs or problems you encounter in the course.

Expand

- Suggest new/different software tools for analysis.
- Contribute new examples.
- Attend one of our 'Train the Trainer" courses so you can take parts of the course back to teach at your home institution.

4.2 Using GitHub to contribute

We use GitHub as a tool for managing version control (AKA keeping a record of the project's development). This helps us stay accountable and transparent. It's also one of the ways we are making steps towards adhering to the FAIR Principles.

If you want to contribute any content such as an update to the course or a new example then via GitHub is the best place to get in contact. For lots more guidance about how to contribute via GitHub, read our GitHub Contribution Guide.

Git (the programming language underlying Github) and Github can be a little intimidating at first but don't worry, the team are here to hold your hand!

Chapter 5

FAIR Principles

5.1 What is FAIR data?

FAIR data is **Findable**, **Accessible**, **Interoperable** and **Reusable**.

These principles are designed to help both humans and machines find and reuse data as easily as possible. They are aspirational but tangible steps can be made towards realising them.

You can read about the ethical values underlying the FAIR principles via the FAIR Cookbook here.

Findable

Findable is all about making sure data/resources are as easy to find as possible.

How we're making resources **findable** at Cloud-SPAN:

- We will be assigning persistent identifiers to our teaching materials to prevent "link rot", or broken links.
- We will register our teaching materials with an appropriate registry (e.g. Carpentries Incubator), so they are easier to find.
- We will be describing all our resources with rich metadata so they can be aggregated by the right registries.

Accessible

Accessible means it is easy to find out how to access the data/resources.

How we're making resources **accessible** at Cloud-SPAN:

- Our resources will be openly available, with no caveats, for use by those who cannot attend our workshops or who prefer self-led study.

Interoperable

Interoperable means data/resources can be easily integrated with other data/resources, and be viewable in different programs, applications or workflows.

How we're making resources **interoperable** at Cloud-SPAN:

- We will supply data which are readable across different programs
- We are providing resources in Markdown, meaning they should display in most browsers.

Reusable

Reusable is about making sure that data/resources are suitable for re-use in different settings.

How we're making resources **reusable** at Cloud-SPAN:

- We will be applying Creative Commons licenses to our resources so they can be reused and remixed by others.
- We welcome (and encourage!) outside contributions of explanations and examples - see the Ways to contribute for more information.

Chapter 6

Cloud-SPAN Online Forum

There are two main places where you'll find the Cloud-SPAN community: GitHub and Peerboard.

To read more about how you can interact with us on GitHub, take a look at our Ways to Contribute and our GitHub Contribution Guide.

For most questions and conversations, however, you'll probably want to use our Peerboard forum. Here you can ask questions, pick people's brains and share any insights you've gained during or after one of our courses. But the forum isn't just for problems. We'd also love to hear about your achievements and successful applications of what you've learned!

6.1 Signing up for Peerboard

Signing up for Peerboard is free, easy and only needs an email address. You should be able to view posts on the forum without making an account, but for the best experience possible we recommend signing up. This will allow you to create posts, join in with discussions and stay up-to-date with conversations.

1. Visit cloudspan.peerboard.com and click 'Sign in' in the top right-hand corner.
2. If you would like to sign up with a Google, Facebook or LinkedIn account then choose one of these options and follow the instructions provided.
3. If you would prefer to make a new account then toggle over from 'Log in' to 'Sign up' and enter your email address and password.

- You will be asked to confirm your email address via a confirmation email.

4. After signing up you will be asked to enter your profile information. Don't forget to decide whether you want a daily digest of new posts delivered to your email inbox.

That's it! You're all signed up!