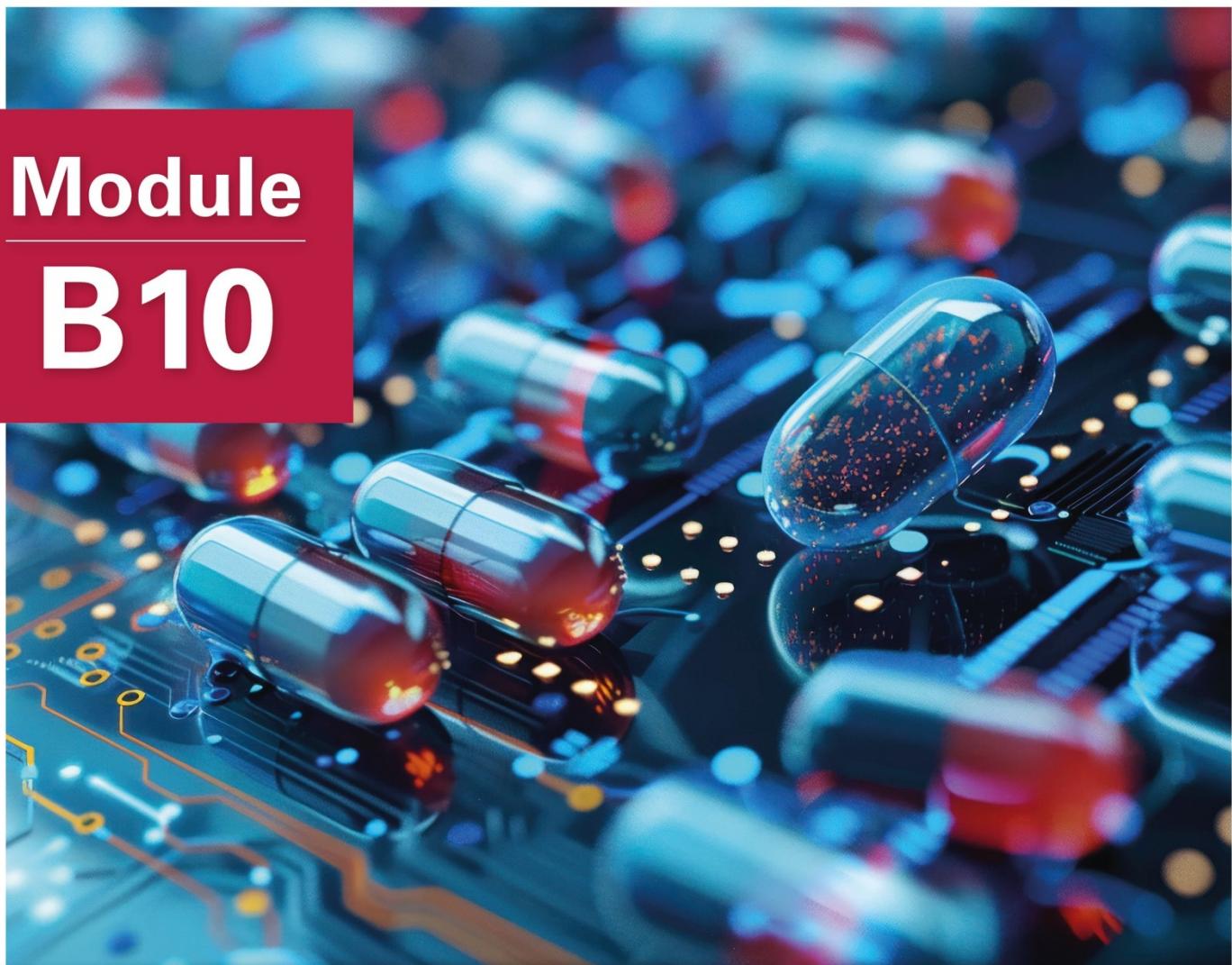


Module **B10**



Bioinformatics Foundational Course

**Ontologies, Standards, Data Curation
and Submission**

NGS Academy for the Africa CDC

Module B10

Ontologies, Standards, Data Curation and Submission

 [back to the table of modules](#)

Module last updated:

December 2024

Suggested or approximate number of sessions	1-2
Suggested or approximate total learning time	4-6 hours
Target audience	Bioinformaticians
Delivery format	Lectures, videos
Level of the module	Introductory



Contributors

George Githinji, Yiqun Li, Perceval Maturure and Nicola Mulder.



Module description

The data generated during pathogen surveillance is valuable to others. Submission of data to public repositories ensures that variants worldwide can be tracked more easily with data from around the globe. Having access to more open, public data for global researchers working on pandemic surveillance, response and therapeutics is very important to facilitate a successful response. Data should be FAIR (Findable, Accessible, Interoperable, Reusable) to facilitate re-use, which requires the data to be well curated using internationally recognized standards and ontologies, and to be submitted to public repositories. This module will cover topics around how to practically apply FAIR principles in terms of data curation and submission.

The module will introduce different pathogen databases (though specific details will be covered in pathogen-specific modules), nucleotide databases (NCBI, SRA & GISAID), Biosamples, etc. and how to retrieve data from them. It will describe how to curate and prepare data for submission to these databases. It will also cover the relevant ontologies used to standardize data and metadata, as well as pathogen-related standards developed by the WHO and PHA4GE. Finally, the module will discuss the need for data security and safeguarding digital information.



Module learning outcomes

On completion of this module, the participants will have knowledge of, or be able to:

- Describe the importance of ensuring that public health data can be accessed and integrated timeously
- Describe some of the challenges and considerations involved in data sharing
- Identify and apply appropriate standards for metadata
- Prepare data for submission to public repositories
- Differentiate between general and specific repositories, find a suitable repository for their data
- Describe what an ontology is and the need for standards and ontologies in databases and research
- Find and retrieve specific information from biological databases.



Module assessments

Module practical: Not applicable

Module quiz: Assessment questions available on the [ASLM platform](#)



Module resources

- [PHA4GE | Video - Contextual data curation](#)
- [SIB-SWISS | GitHub - Training collection](#)
- [SIB-SWISS | GitHub - FAIR data](#)
- [The Carpentry | GitHub - FAIR in \(biological\) practice](#)
- [The Carpentry | GitHub - FAIR for leaders](#)
- [WCS & COG-UK | Slides - Data tools and pipelines](#)
- <https://carpentries-incubator.github.io/fair-for-leaders/04-data-repositories/index.html>
- [H3ABioNet | Webpage - Biological Databases and Ontologies](#)
- [H3ABioNet | Webpage - Databases I](#)



Acknowledgements

We would like to thank the following individuals, in alphabetical order of last name, for their valuable time and effort spent in designing (i.e., drafting, reviewing, and refining) this module: **George Githinji, Yiqun Li, Perceval Maturure and Nicola Mulder.**

Furthermore, we would like to thank the following institutions, societies, journals and individuals from whom we sourced open-access resources, used in this module:

COVID-19 Genomics UK Consortium, H3ABioNet is a Pan African Bioinformatics Network for the Human Heredity and Health in Africa (H3Africa) consortium, Pathogen Health Alliance for Genomic Epidemiology, Wellcome Connecting Science; Amadou Diallo, Fatma Guerfali, Kirsty Lee Garson, Emma Griffiths, George Githinji, Leigh Jackson, Aquillah Kanzi, Benjamin Kumwenda, Stanford Kwenda, Alice Matimba, Nicola Mulder, SIB-SWISS GitHub contributors, The Carpentry GitHub contributors.