


# Global activities in bioinformatics training and education

Nicola Mulder

PI: H3ABioNet

University of Cape Town

# Outline

- ▶ Challenges in bioinformatics education & training
  - ▶ Addressing challenges -Education summits
  - ▶ H3ABioNet & NGS Academy -running a diverse bioinformatics training program
  - ▶ Other applications of trainer resources
- 

# General challenges in training

- **Qualified trainers** who are knowledgeable in the topic and are good trainers
- Developing and updating **training materials** –lectures, tools and practice datasets
- Teaching a **mixed audience**
- Keeping the **audience engaged**
- **Assessing competency** and skills gain
- **Going virtual**



# Challenges in bioinformatics training

- There is a **huge demand** for more people with bioinformatics skills
- Bioinformatics **topics are vast**, in flux and change rapidly
- **Audiences are broad** and need different levels of competency
- Competencies often require **additional foundational skills** (statistics, programming)
- Training requires **theory and hands-on practice**
- Few trainees have a standard bioinformatics education

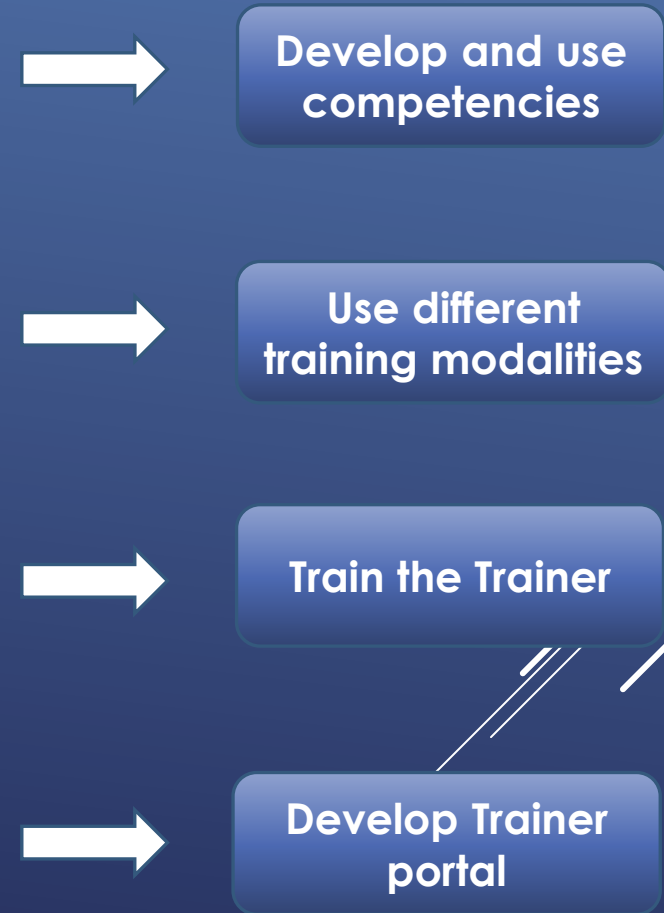


# Global efforts to address challenges

## Challenges:

- ▶ Audiences are broad and need different levels of competency
- ▶ Assessing competency and skills gain
- ▶ High demand, not enough trainers
- ▶ Qualified trainers who are knowledgeable in the topic and are good trainers
- ▶ Teaching to a mixed audience, keeping the audience engaged
- ▶ Developing and updating training materials
- ▶ Bioinformatics topics are vast and in flux

## Solutions:



# Education summits

**AIM:** Bring together Bioinformatics trainers and educators to drive the development of standards and guidelines for Bioinformatics training and education globally

- Summit I: 14-17<sup>th</sup> May 2019, Cape Town hosted by H3ABioNet
- Summit II: May 2020 virtual hosted by EBI
- Summit III: May 2021 virtual hosted by CABANA
- Summit IV: May 2022 virtual hosted by APBioNet

## Format:

- ▶ Minimal presentations
- ▶ Breakout working sessions
- ▶ Discussions



# Break out groups



# Education summit projects

**Competencies**

**Course endorsement**

**Trainer resources**

**Train-the-trainer**

**Going virtual**

**Training in LMICs**



# Competencies: what are they?

- ▶ <https://www.td.org/insights/what-is-a-competency>
- ▶ something you need to be able to do a specific job
- ▶ to demonstrate competence, workers must be able to perform certain tasks or skills with a required level of proficiency
- ▶ A competency is broken down into specific skills or tasks
- ▶ To achieve competence in a particular job, a person should be able to perform various tasks or skills at a target proficiency level

**Knowledge**

**Skills**

**Attitudes**

# Developing competencies

**Define  
audience/persona**

## **Bioinformatics User**

Physician

Lab technician

Ethicist

Biocurator

## **Bioinformatics Scientist**

Academic life  
science  
researcher

Molecular life  
science  
educator

Academic  
bioinformatics  
researcher

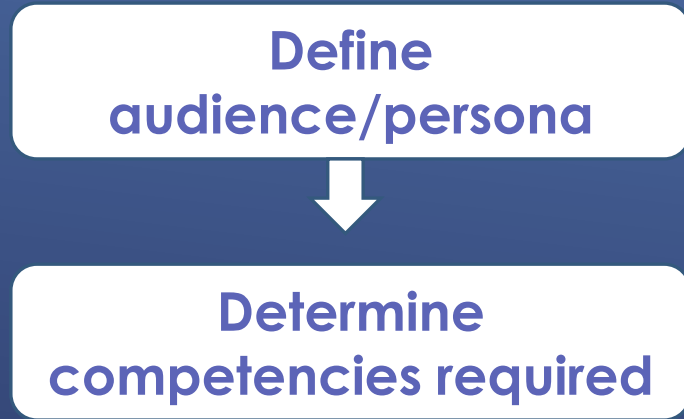
Core facility  
scientist

## **Bioinformatics Engineer**

Bioinformatician  
in academic or  
research  
infrastructure  
support role

Bioinformatics  
software  
developer/  
software  
engineer

# Developing competencies



| Relationship         | #  | Competency  |
|----------------------|----|---|
| Bioscience           | A3 | <i>Work at depth in at least one technical area aligned with the life sciences</i>  |
|                      | B3 | <i>Prepare life science data for computational analysis</i>   |
|                      | C3 | <i>Have a positive impact on scientific discovery through bioinformatics</i>  |
| Data Science         | D3 | <i>Use data science methods suitable for the size and complexity of the data</i>  |
|                      | E3 | <i>Manage own and others' data according to community standards and principles</i>  |
|                      | F3 | <i>Make appropriate use of bioinformatics tools and resources</i>   |
| Computer Science     | G3 | <i>Contribute effectively to the design and development of user-centric bioinformatics tools and resources</i>              |
|                      | H3 | <i>Make appropriate and efficient use of scripting and programming languages</i>  |
|                      | I3 | <i>Construct, manage and maintain bioinformatics computing infrastructure of varying complexity</i>                         |
| Professional Conduct | J3 | <i>Comply with professional, ethical, legal and social standards and codes of conduct relevant to computational biology</i> |
|                      | K3 | <i>Communicate meaningfully with a range of audiences - within and beyond your profession</i>                               |
|                      | L3 | <i>Work effectively in teams to accomplish a common goal</i>  |
|                      | M3 | <i>Engage in continuing professional development in bioinformatics</i>  |

# Developing competencies

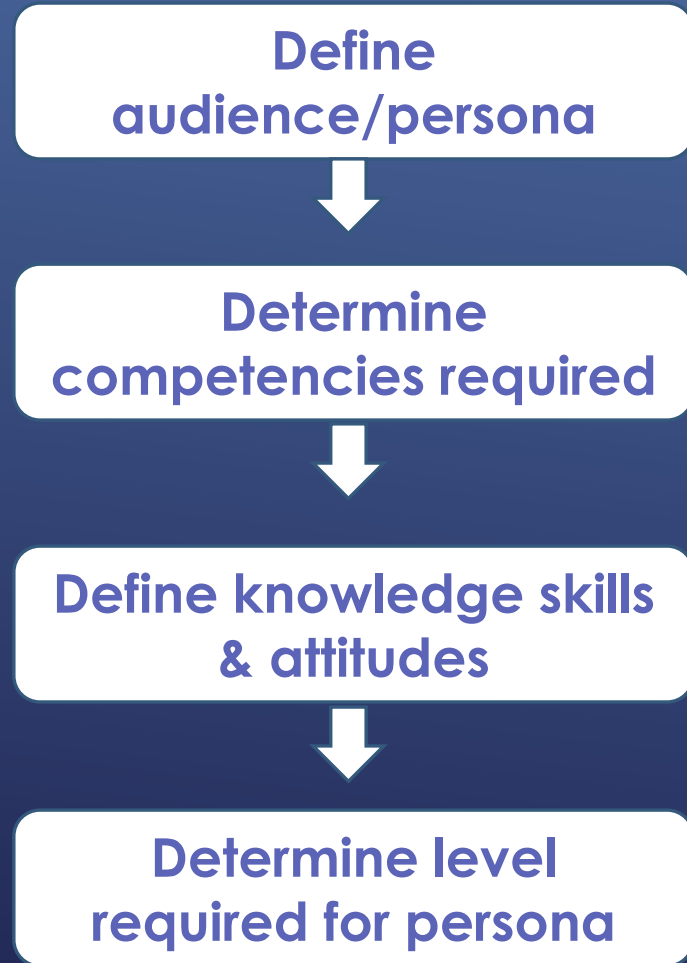


## E3 - Manage own and others' data according to community standards and principles (UKNOS: COGBIO1; COGBIO2)

| What do you need to know to exhibit competency in this area?   | What skills do you need to exhibit competency in this area?  |
|--|--|
| KE3-1 (UA K15). Database design and management, including information security considerations, big-data technologies, and database languages and systems.<br>KE3-2 (UA K6). Current approaches for modelling and warehousing of life science data.<br>KE3-3 (UA K20). The role of governance, curation and information architecture in data management.<br>KE3-4 (EA3.4c). Common document identification, tracking and control procedures.<br>KE3-5. Broader implications of using/storing sensitive data.<br>KE3-6 (UA K8). Knowledge representation including file formats, ontologies and other controlled vocabularies.<br>KE3-7 (UA K10). Data storage and format requirements of downstream techniques to integrate, interpret, analyse and visualise biological data sets. | SE3-1 (UA S27). Designs and implements appropriate data storage formats and associated database structure.<br>SE3-2 (UA S28). Chooses appropriate computational infrastructure and database solutions - including internal or external/cloud resources.<br>SE3-3 (UA S29). Stores and analyses data in accordance with ethical, legal and commercial standards, including checking who has access.<br>SE3-4 (UA S30). Curates biological data using suitable metadata, ontologies and/or controlled vocabularies.<br>SE3-5 (UA S31). Makes use of suitable programming languages and/or workflow tools to automate data handling and curation tasks.<br>SE3-6. Drafts and files an appropriate Data Management Plan.<br>SE3-7 (UA S33). Prepares data for submission to appropriate public data repositories as required, being aware of ethical and legal considerations. |
| How does a person with this competence behave?   | How does a person with this competence avoid behaving?   |
| AE3-1 (UA A51). Acts with awareness of the wider context in which scientific research operates, recognising the implications for professional practice.  | NE3-1. Uses proprietary systems for data management.<br>NE3-2. Publishes scientific results without making adequate data available for reproducibility.<br>NE3-3. Uses third-party data without appropriate legal or ethical approval and/or without citation or acknowledgement appropriate to the discipline.  |

|    |   |   |                       |
|----|---|---|-----------------------|
| 48 | D3: Use data science methods suitable for the size and complexity of the data |   |                       |
| 49 | D3: Use data science methods suitable for the size and complexity of the data | KD3-1: Appropriate statistics in the context of bioinformatics and life science data analysis               | Knowledgeledge        |
| 50 | D3: Use data science methods suitable for the size and complexity of the data | KD3-2: Statistical and mathematical modelling methods, and key scientific and statistical analysis          | Knowledgeledge        |
| 51 | D3: Use data science methods suitable for the size and complexity of the data | KD3-3: General data science approaches to life science problems   | Knowledgeledge        |
| 52 | D3: Use data science methods suitable for the size and complexity of the data | KD3-4: Experimental design to ensure the statistical validity of high-throughput experiments                | Knowledgeledge        |
| 53 | D3: Use data science methods suitable for the size and complexity of the data | KD3-5: Comprehensions the importance of statistics in experimental design, data analysis and interpretation | Knowledgeledge        |
| 54 | D3: Use data science methods suitable for the size and complexity of the data | SD3-1: Determines the best methods for data analysis, including the selection of statistical tests          | Skill                 |
| 55 | D3: Use data science methods suitable for the size and complexity of the data | SD3-2: Identifies and defines appropriate computing infrastructure requirements for the analysis            | Skill                 |
| 56 | D3: Use data science methods suitable for the size and complexity of the data | SD3-3: Applies statistical methodologies appropriate to the analysis of data in the context of the project  | Skill                 |
| 57 | D3: Use data science methods suitable for the size and complexity of the data | SD4-4: Can process data into formats suitable for analysis, whilst maintaining integrity of the data        | Skill                 |
| 58 | D3: Use data science methods suitable for the size and complexity of the data | AD3-1: Approaches problems with a systems-based, data-driven approach to scientific discovery               | Effective attitudes   |
| 59 | D3: Use data science methods suitable for the size and complexity of the data | AD3-2: Recognises own limitations and consults experts when required  | Effective attitudes   |
| 60 | D3: Use data science methods suitable for the size and complexity of the data | AD3-3: Is conscious of the risks of overfitting and of appropriate methods for validation and comparison    | Effective attitudes   |
| 61 | D3: Use data science methods suitable for the size and complexity of the data | AD3-4: Critically reviews results before interpretation and communication                                   | Effective attitudes   |
| 62 | D3: Use data science methods suitable for the size and complexity of the data | AD3-5: Reports on statistical methods applied in a project, the rationale, limitations and results          | Effective attitudes   |
| 63 | D3: Use data science methods suitable for the size and complexity of the data | ND3-1: Does not engage with statistical components of project   | Ineffective attitudes |

# Developing competencies



|            |  | Scientist roles  |                                 |                                    |                         | Engineer roles  |   |  |
|------------|--|--|---------------------------------|------------------------------------|-------------------------|---|---|--|
| Competency |  | Discovery biologist (e.g. pharma or agri-food industry); Academic molecular life science researcher; | Molecular life science educator | Academic bioinformatics researcher | Core facility scientist | Bioinformatician supporting an academic lab or department | Bioinformatician in a research infrastructure setting | Software developer/ software engineer in a bioinformatics role |
|            | General biology  | evaluation   | comprehension                   | synthesis                          | knowledge               | application   | application   | application  |
| A          | Depth in at least one area of biology (e.g., evolutionary biology, genetics, molecular biology, biochemistry, anatomy, physiology) | create   | analyze                         | create                             | evaluate                | comprehension   | comprehension   | comprehension  |
| B          | Biological data generation technologies.   | evaluation   | understand                      | evaluation                         | evaluation              | comprehension   | comprehension   | comprehension  |
| C          | Details of the scientific discovery process and of the role of bioinformatics in it.   | application  | evaluation                      | synthesize to create               | application             | application   | application   | application  |
| D          | Statistical (research) methods in the context of molecular biology, genomics, medical, and population genetics research.           | application  | evaluation                      | synthesize to create               | application             | application   | application   | synthesis to application                                       |

# Competency hub example

The image shows a screenshot of the Competency Hub website. The top navigation bar includes the URL 'competency.ebi.ac.uk', a progress indicator '50%', a star icon, and a search bar. The main header features the 'Competency Hub' logo and the tagline 'Supporting competency-based training and professional development.' Below this, an 'Overview' section states that the hub is a web-based tool for creating and managing competency frameworks, with a 'read more' link. A grid of partner logos is displayed, including BioExcel, CORBEL, RITrain, ISCB, NHS Health Education England, and CINECA. On the right, a user profile for 'Martha - Bioinformatics researcher' is shown, featuring an avatar, an 'Admin login' button, and a 'Create your profile' button. The profile details include a list of qualifications and background information, followed by a paragraph about her academic career.

competency.ebi.ac.uk 50% ☆ Search

**Competency Hub**  
Supporting competency-based training and professional development.

https://competency.ebi.ac.uk/framework/iscb/3.0/profile/view/8193/martha-0 67% ☆ Search

Home > ISCB 3.0 > View career profile

**Competency Hub**  
Supporting competency-based training and professional development.

**Overview**  
Competency Hub is a web-based tool to support the creation and management of competency frameworks [read more](#)

**BioExcel**  
Professionals in computational biomolecular research

**CORBEL**  
Technical operators of biomedical research infrastructures

**RITrain**  
Managers of research i

**ISCB**  
INTERNATIONAL SOCIETY FOR COMPUTATIONAL BIOLOGY

**NHS**  
Health Education England

**CINECA**

**Martha - Bioinformatics researcher**

Admin login

Create your profile

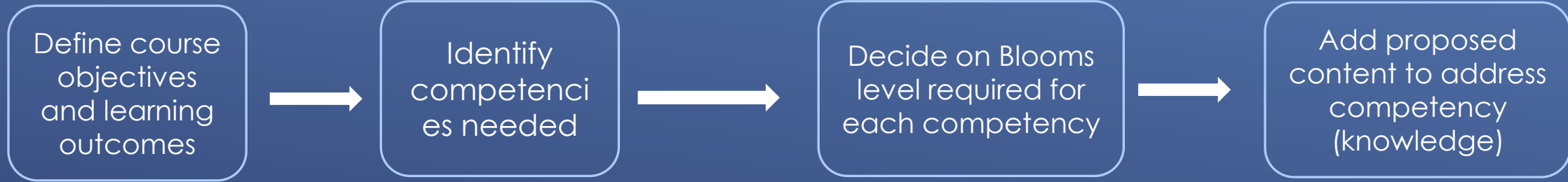
**Qualification and background**

- Math major, Cornell University, USA
- PhD in Physics, Princeton University, USA
- 1st postdoc, University of Saskatchewan, Canada
- 2nd postdoc, Laboratory of Molecular Biology, Cambridge, UK
- Tenure track position, University of Toronto, Canada

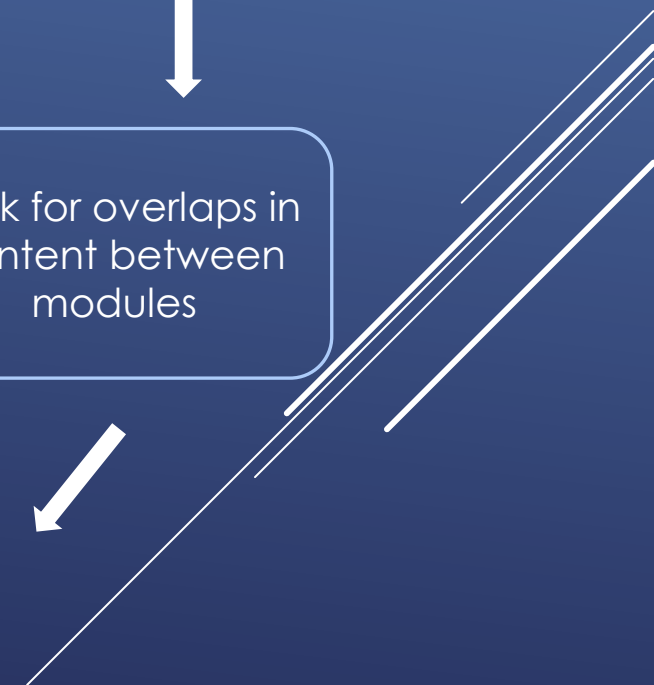
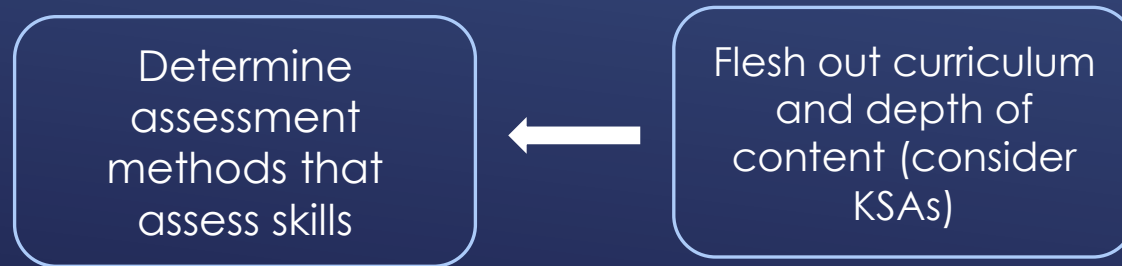
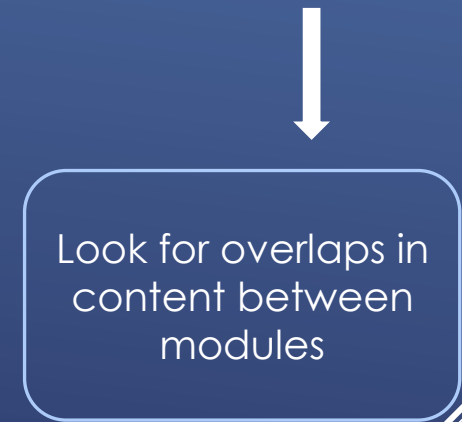
Martha's academic career began with an interest in mathematics and statistics, as well as a childhood interest in computers. Her PhD involved developing statistical analysis methods of X-ray diffraction data. This research led to post-docs in Canada and the UK working on structural biology where Martha developed expertise at the interface between physics, computing and biology. More recently Martha has established her own research group in Toronto, gaining a tenure track position and several high-profile grants from the pharmaceutical industry and the NIH.

How do we use the competencies?

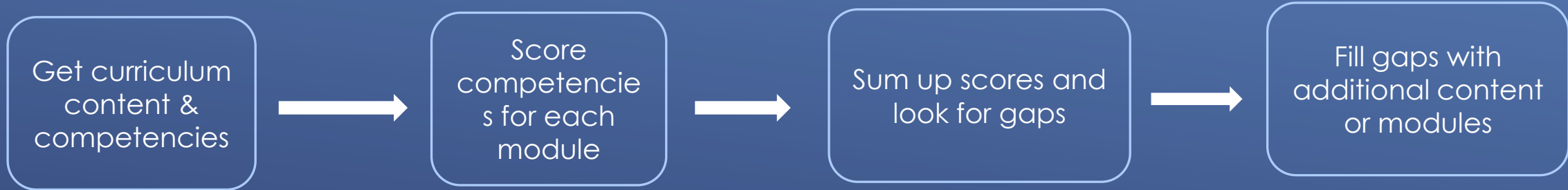
# Designing a new course



| Core competency   | Desired target level (Bloom's) | Starting student level (Bloom's) | Proposed courses/modules/activities |
|---|--------------------------------|----------------------------------|-------------------------------------|
| (a) General biology   |                                |                                  |                                     |
| (b) Depth in at least one area of biology (e.g., evolutionary biology, genetics, molecular biology, biochemistry, anatomy, physiology). |                                |                                  |                                     |
| (c) Biological data generation technologies.  |                                |                                  |                                     |
| (d) Details of the scientific discovery process and of the role of bioinformatics in it.  |                                |                                  |                                     |
| (e) Statistical research methods in the context of molecular biology, genomics, medical, and population genetics research.              |                                |                                  |                                     |
| (f) Bioinformatics tools and their usage.   |                                |                                  |                                     |



# Mapping competencies to a course



## DEGREES

| Core competency   | Intro to biology | Biostatistics | Programming |
|---|------------------|---------------|-------------|
| (a) General biology   | 2                |               |             |
| (b) Depth in at least one area of biology (e.g., evolutionary biology, genetics, molecular biology, biochemistry, anatomy, physiology). | 2                |               |             |
| (c) Biological data generation technologies.  | 1                |               |             |
| (d) Details of the scientific discovery process and of the role of bioinformatics in it.  | 1                |               |             |
| (e) Statistical research methods in the context of molecular biology, genomics, medical, and population genetics research.              | 0                |               |             |
| (f) Bioinformatics tools and their usage.   | 0                |               |             |

## COURSES

Decide on Blooms level targeted for each competency

Map to specific KSAs?

| Core competency   | Bloom | Knowledge | Skills | Attitude |
|---|-------|-----------|--------|----------|
| (a) General biology   |       |           |        |          |
| (b) Depth in at least one area of biology (e.g., evolutionary biology, genetics, molecular biology, biochemistry, anatomy, physiology). |       |           |        |          |
| (c) Biological data generation technologies.  |       |           |        |          |
| (d) Details of the scientific discovery process and of the role of bioinformatics in it.  |       |           |        |          |
| (e) Statistical research methods in the context of molecular biology, genomics, medical, and population genetics research.              |       |           |        |          |
| (f) Bioinformatics tools and their usage.   |       |           |        |          |

Identify and fill gaps in content



Guidelines document



# How we have applied these

## ▶ H3ABioNet



Pan African bioinformatics network for H3Africa – building bioinformatics capacity

- ▶ Mapping existing courses

- ▶ Used to design new courses

## ▶ NGS Academy

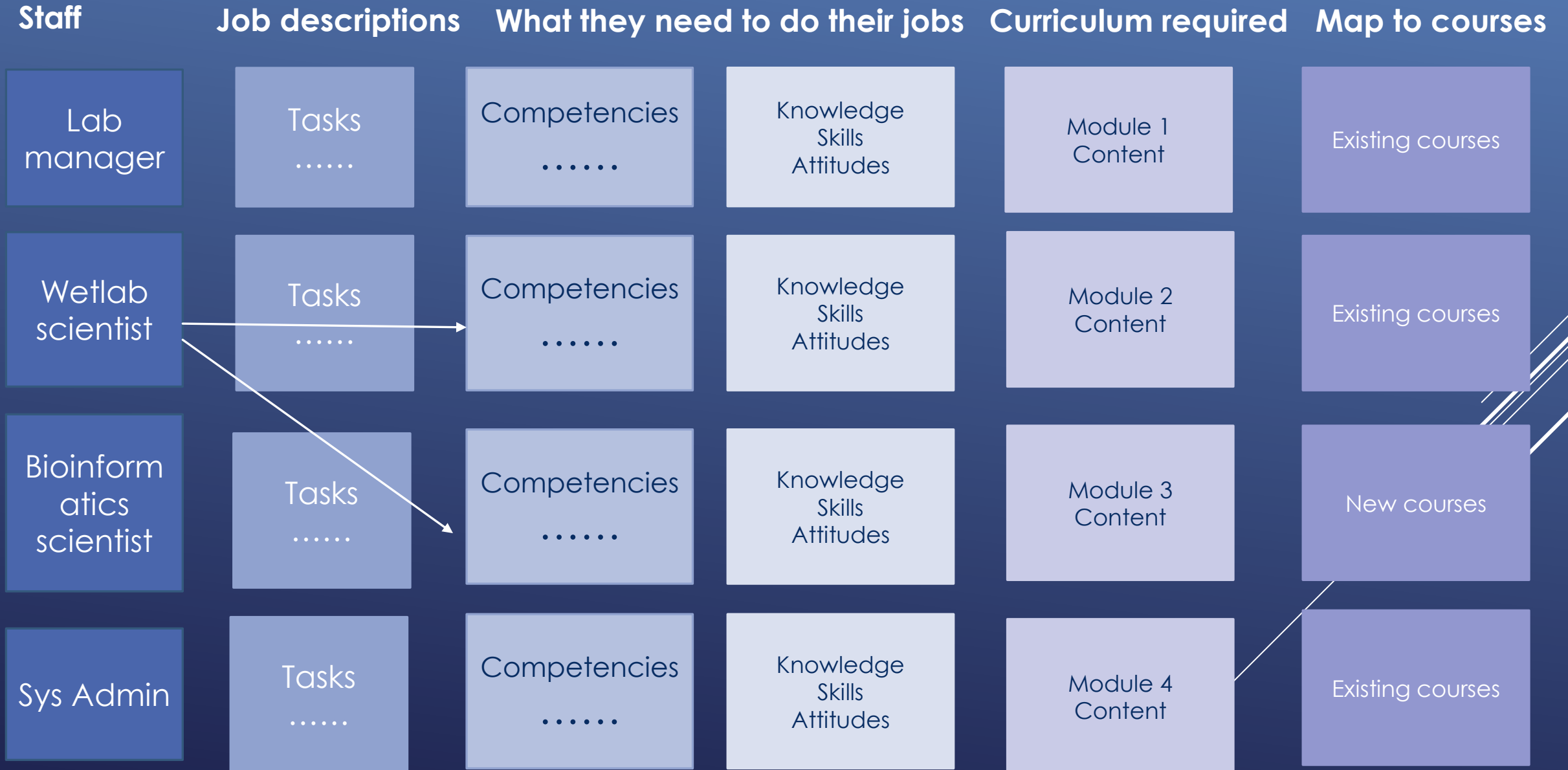
- ▶ Development of a recommended curriculum for pathogen surveillance



NGS for pathogen surveillance Training program  
for the Africa CDC Pathogen Genomics Initiative



# NGS academy curriculum development



# Example of the curriculum

| What is genomic epidemiology, when and how is it used in public health, tracking variants |  |   |                                |                                   |                                  |          |                                     |              |   |
|---|--|---|--------------------------------|-----------------------------------|----------------------------------|----------|-------------------------------------|--------------|---|
|   | A  | B   | C                              | D                                 | E                                | F        | G                                   | H            | I   |
|   | Viral Pathogen Surveillance Topics                                       | Content   | Course Level for Wet-lab Staff | Course Level for Bioinformatician | Course Level for Laboratory Head | Comments | SARS-CoV-2 NGS Training             | Course Level | SARS-CoV-2 Bioinformatics Training          |
| 1   | Introduction to Genomic Epidemiology                                     | What is genomic epidemiology, when and how is it used in public health, tracking variants                                       | Introductory                   | Intermediate                      | Intermediate (adv?)              |          | SARS-CoV-2 NGS Training Module 1    | Beginner     |   |
| 2   | Study Design   | Study design, which samples to sequence, targeted or metagenomics analysis  | Introductory                   | Intermediate                      | Advanced                         |          | SARS-CoV-2 NGS Training Module 2    | Beginner     |   |
| 3   | Introduction to NGS  | Overview of different NGS technologies and workflows  | Intermediate                   | Intermediate                      | Introductory                     |          | SARS-CoV-2 NGS Training Module 1    | Beginner     | SARS-CoV-2 Bioinformatics Training Module 1 |
| 4   | Sample processing and library preparation (Theory and wet lab practical) | Sample collection, handling, storage, DNA/RNA extraction, library preparation, DNA amplification, quantification, normalization | Advanced                       | Introductory                      | Introductory                     |          | SARS-CoV-2 NGS Training Module 2    | Beginner     |   |
| 5   | Lab practical: DNA sequencing  | Introduction to instrument, running sequencing (technology specific)  | Advanced                       |                                   |                                  |          |                                     |              |   |
| 6   | ARTIC amplicon sequencing  | Introduction to ARTIC amplicon sequencing, data processing through workbench  | Advanced                       | Advanced                          | Introductory                     |          |                                     |              |   |
| 7   | NGS data processing  | Sequence data QC, mapping, variant calling, assembly, consensus generation  | Intermediate                   | Advanced                          | Introductory                     |          |                                     |              | SARS-CoV-2 Bioinformatics Training Module 1 |
| 8   | NGS, Illumina, ONT workflows   | Sequencing and data processing workflows specific to technology platform  | Advanced                       | Advanced                          | Introductory                     |          | SARS-CoV-2 NGS Training Module 3, 4 | Intermediate |   |
| 9   | Nextclade and Pangolin   | Introduction to Nextclade and Pangolin, using tools through workbench and online  | Intermediate                   | Advanced                          | Introductory                     |          |                                     |              |   |
| 10  | Galaxy workflow (or other selected workflow(s))                          | Galaxy Concepts: Histories, data and workflows: Viral pathogen workflow(s)  | Intermediate                   | Advanced                          | Introductory                     |          |                                     |              |   |

Mapping to competencies developed from: job descriptions, existing competency frameworks

# Using the curriculum

## Foundational skills

Lab manager

Wetlab scientist

Bioinformatics scientist

Sys Admin

Health & Safety

Ethics & policy

Intro to Genetic Epi

Unix, scripting

Programming

Bioinformatics

Workflows, HPC

## Applied skills

Introduction to NGS, Study design

### Viral pathogens

Sample preparation

Sequencing

Data QC processing

Analysis workflow

Phylogenetics, Bioinformatics

Data curation & submission

Presenting data for policy/action

SARS-COV-2

### Bacterial pathogens

Sample preparation

Sequencing

Data QC processing

Analysis workflow

Phylogenetics, Bioinformatics

Data curation & submission

Tuberculosis

# Using the curriculum

## Applied skills

## Foundational skills

Lab manager

Wetlab scientist

Bioinformatics scientist

Sys Admin

Health & Safety

Ethics & policy

Intro to Genetic Epi

Unix, scripting

Programming

Bioinformatics

Workflows, HPC

Introduction to NGS, Study design

### Viral pathogens

Sample preparation

Sequencing

Data QC processing

Analysis workflow

Phylogenetics, Bioinformatics

Data curation & submission

Presenting data for policy/action

SARS-COV-2

### Bacterial pathogens

Sample preparation

Sequencing

Data QC processing

Analysis workflow

Phylogenetics, Bioinformatics

Data curation & submission

Tuberculosis

# Education summit projects

**Competencies**

**Course endorsement**

**Trainer resources**

**Train-the-trainer**

**Going virtual**

**Training in LMICs**

Development of  
process for  
endorsement of  
courses &  
degrees by ISCB

# Education summit projects

**Competencies**

**Course endorsement**

**Trainer resources**

**Train-the-trainer**

**Going virtual**

**Training in LMICs**

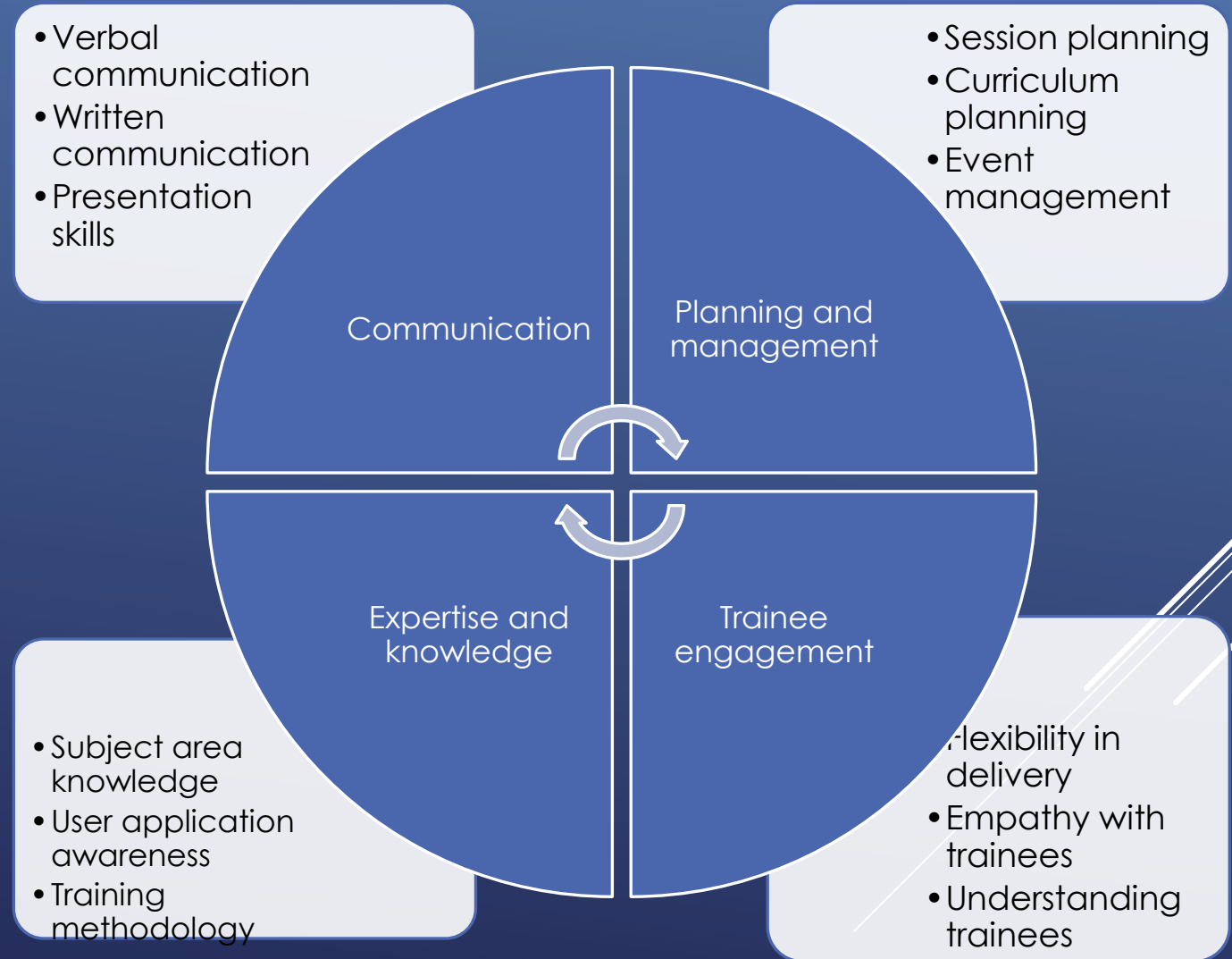
Training trainers  
in developing &  
delivering  
training



# Online T-t-T course

- Developing along with EBI, Wellcome Connecting Science
- Building TtT curriculum to be run using remote classroom model

**First course will  
run in November  
2022**





# Course outline

- ▶ Week 1: Training theory and practical aspects
  - ▶ Theoretical aspects
  - ▶ Course design elements and factors
- ▶ Week 2: Training design elements
  - ▶ Course design Part 1
  - ▶ Course design Part 2
- ▶ Week 3: Training evaluation
  - ▶ Assessment and evaluation
  - ▶ Review and feedback of course designs



**wellcome connecting science** **H3ABioNet** **EMBL-EBI**

## Train-the-trainer: course design and delivery for bioinformatics trainers

22 November–8 December 2022  
Tuesdays and Thursdays 14:00–17:00 CAT

Learn how to plan, design, and deliver training in bioinformatics to impart skills to students, professionals, and future experts.

- Blended learning format, delivered virtually
- Connect and network with genomic scientists and bioinformaticians.



### Course delivery






- Twice weekly contact sessions with expert bioinformatics and data science trainers, delivered across 3 weeks
- Discussion through highly interactive online forums

### Course content

Join fellow scientists from bioinformatics and data science, and develop your ability to:

- Apply theoretical and pedagogical concepts to

# H3ABioNet Train-the- trainer

|   |  |  |  |  |
|---|--|--|--|--|
|  |  |  |  |  |
| <b>Trainer</b>  | <b>User</b>  | <b>Scientist</b>   | <b>Engineer</b>  | <b>Sys Admin</b>   |
| Introductory +<br>Train-the-<br>Trainer +<br>Specialized<br>Courses               | Introductory<br>Courses +<br>Internships   | Introductory +<br>Intermediate<br>+ Specialized<br>Courses +<br>Internships        | Introductory +<br>Intermediate<br>+ Advanced<br>Courses +<br>Internships           | Sys Admin<br>and Technical<br>+ Data<br>Management<br>Courses                      |



F2F Training

Internship

Train-the-Trainer

Hackathons

Mixed-model



**Community Building**

SOFT SKILLS



**Training Resource Access**

**H3ABioNet**  
1.42K subscribers



# H3ABioNet train-the-trainer activities

- **Focused on developing local trainers throughout by:**
  - Dedicated T-t-T courses and events
  - Building faculty from trainees to TAs to trainers
  - Train-the-trainer web resources
- **Software & Data Carpentries**
  - Software (Command Line, Python, Git), Data (R, Spreadsheets) & Library Carpentries
  - Build up and foster a community of certified Carpentries instructors within H3ABioNet
  - Teach foundational coding and data analysis skills at local and regional institutions



# Carpentries Instructors



## KENYA

*International Centre of Insect Physiology and Ecology (ICIPE)*

## MOROCCO

*University Mohamed First (MFU)  
Institut Pasteur du Maroc (IPM)*

## SOUTH AFRICA

*University of Cape Town (CBIO)*

## TANZANIA

*Muhimbili University of Health and Allied Sciences (MUHAS)*

## TUNISIA

*Institut Pasteur de Tunis (IPT)*

## SUDAN

*University of Khartoum (UofK)*

## RWANDA

*Carnegie Mellon University Africa*

# Education summit projects

**Competencies**

**Course endorsement**

**Trainer resources**

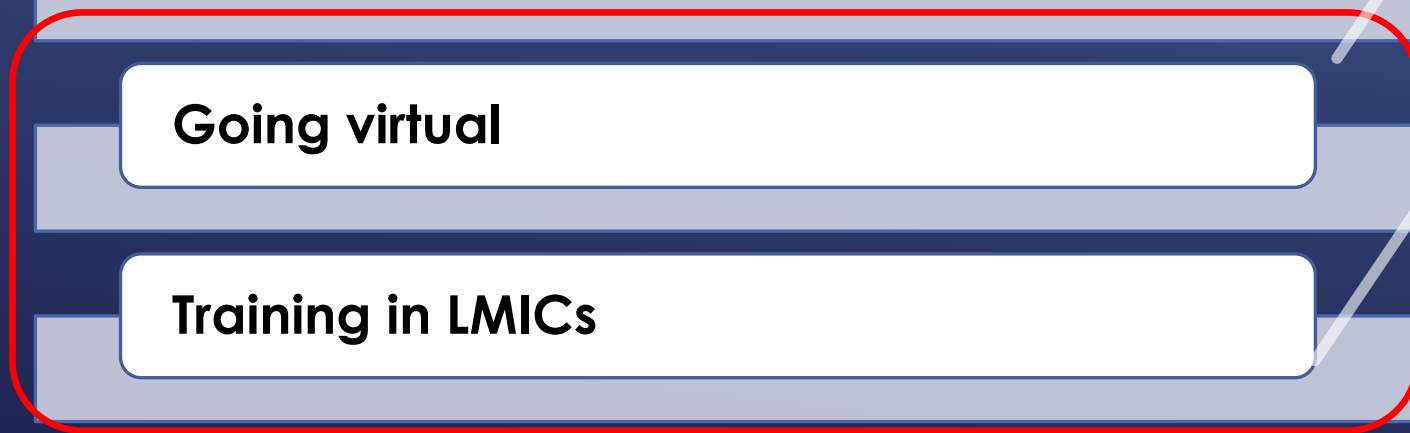
**Train-the-trainer**

**Going virtual**






**Training in LMICs**

Tools for  
delivery;  
challenges,  
assessing  
learning, FAIR


Guidelines;  
building  
infrastructure;  
language  
barriers, EDI




# H3ABioNet Training Environment

|   |   |  |  |  |
|---|---|--|--|--|
|  |  |  |  |  |
| <b>Trainer</b>  | <b>User</b>   | <b>Scientist</b>   | <b>Engineer</b>  | <b>Sys Admin</b>   |
| Introductory +<br>Train-the-Trainer +<br>Specialized<br>Courses                   | Introductory<br>Courses +<br>Internships  | Introductory +<br>Intermediate<br>+ Specialized<br>Courses +<br>Internships        | Introductory +<br>Intermediate<br>+ Advanced<br>Courses +<br>Internships           | Sys Admin<br>and Technical<br>+ Data<br>Management<br>Courses                      |


## Going virtual




**F2F Training**




**Internship**




**Train-the-Trainer**




**Hackathons**




**Mixed-model**



**Community Building**



**Training Resource Access**



**H3ABioNet**

1.42K subscribers

# Training mode challenges

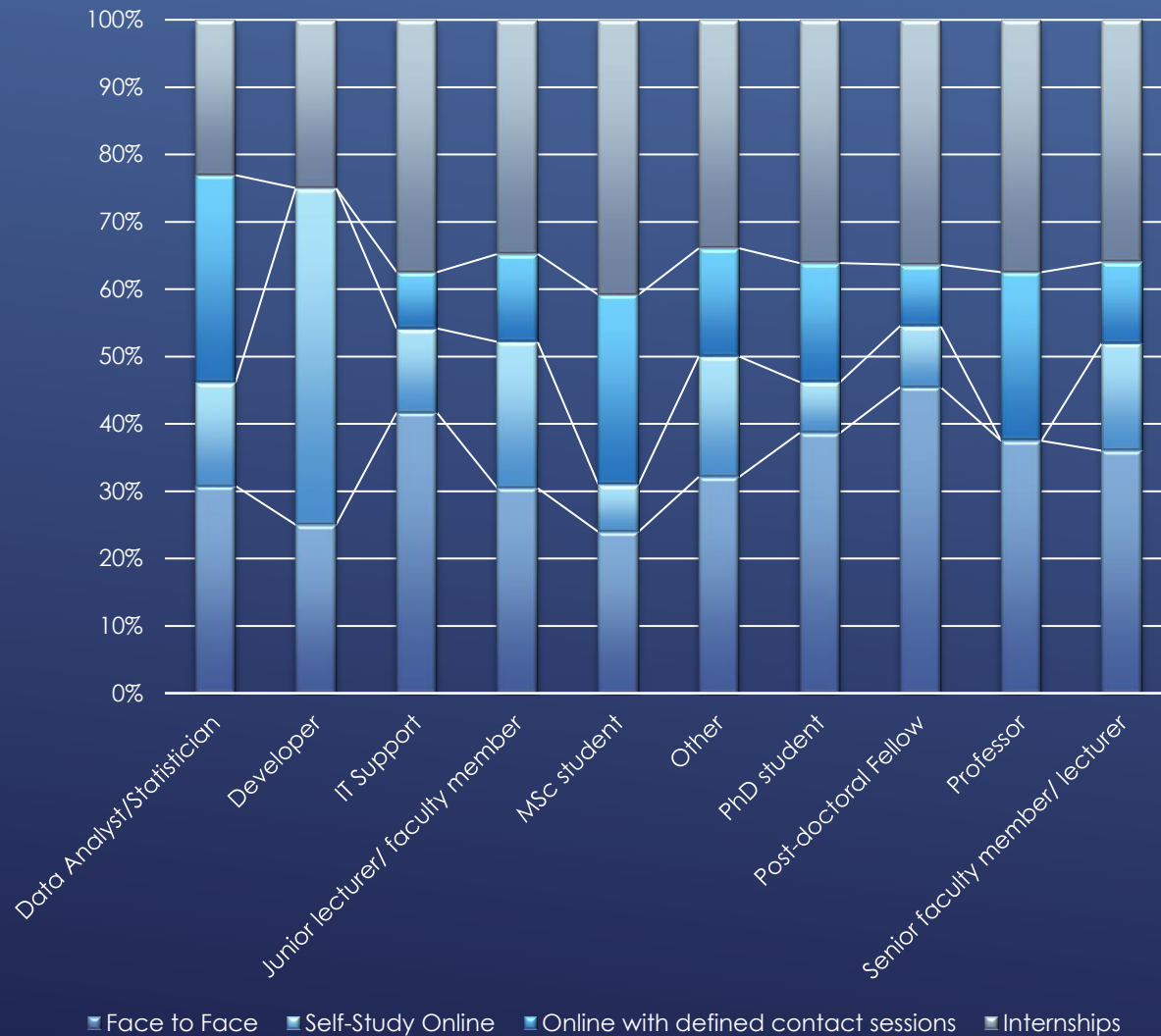


| Online Training            |  |
|----------------------------|--|
| <b>Pros</b>                | <b>Cons</b>                            |
| Cost efficient             | Challenge to foster sense of community |
| Reach a large audience     | Challenge to form collaborations       |
| Easy to share material     |  |
| Live Workshops/Internships |  |
| <b>Pros</b>                | <b>Cons</b>                            |
| Dedicated time             | Costly/unpredictable logistics         |
| Close interaction          | Limited audience reached               |
| Hackathons                 |  |
| <b>Pros</b>                | <b>Cons</b>                            |
| Defined aims and output    | Selection of participants crucial      |
| Develop practical skills   | Base knowledge level required          |
| Cross-disciplinary         | Limited audience                       |



# Training mode requests

## Preferred Mode of Learning by Role



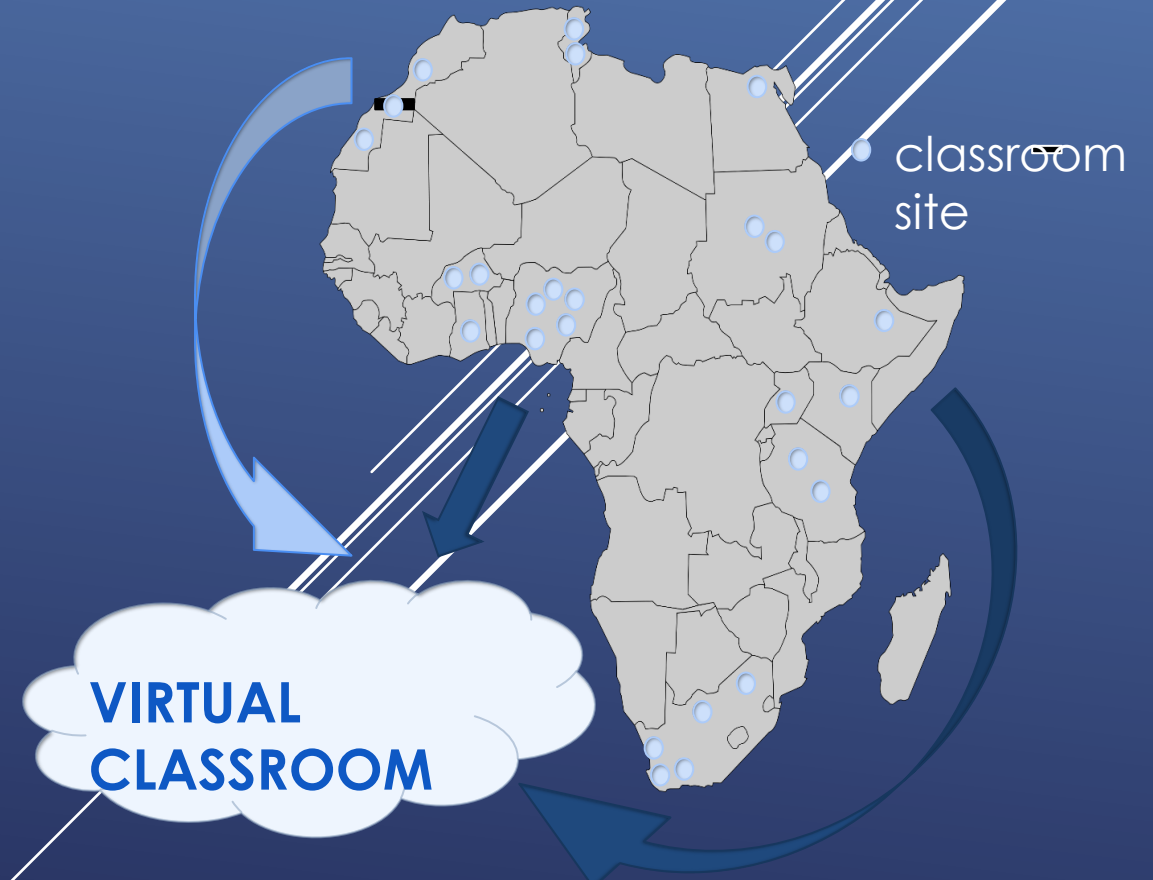
- ▶ No optimal mode suitable for all training
- ▶ Learning styles differ between individuals from different backgrounds
- ▶ Live training events are still the preferred option
- ▶ Live online training has proven successful



# Distributed classroom model

- Run on specific days over 1-2 months,
- set contact time per week (3 hours per contact session)
- Distance learning model – physical classrooms connected to virtual classrooms
- Use video conferencing facility
- Course management system, e.g. Sakai

zoom



# Remote classroom training



Exit Student View

Search

Overview

Announcements

Week 11: Panel Discussion - (A) Microbiomes & (B) Nutrigenomics

Course Outline

Forums

Q&A

Assignments

Calendar

Resources

Chat Room

Help

### Questions and Answers

Ask a question anonymously

View: Most popular

Viewing 1 - 20 of 23 items

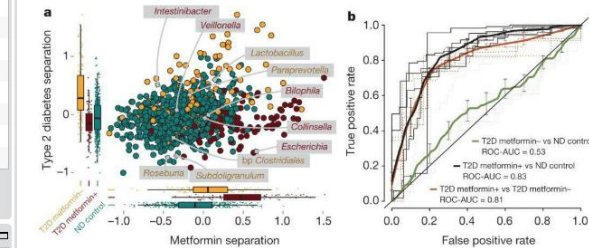
show 20 items per page

| Rank | Questions   | Views |
|------|---|-------|
| 1    | <a href="#">what is autism? what is cause of autism?</a>  | 33    |
| 2    | <a href="#">I'm very confused about what genome, DNA, genetic code are? I need some simplified clarification.</a>                 | 28    |
| 3    | <a href="#">am not understanding the assignment of week 3 due 10th May, as if they are answers to questions, please help</a>      | 13    |
| 4    | <a href="#">Week 2 assignment is not clear, do we draw the pedigree or what? We are not in a position to answer it. Thanks f</a>  | 10    |
| 5    | <a href="#">please help to clarify on the deadline for submission of the class exercise. Is it on 2nd /05/2017 or 16/05/2017?</a> | 9     |
| 6    | <a href="#">Diabetes is assumed to be genetically inherited. Please clarify</a>   | 7     |

|                         |  |
|-------------------------|--|
| bw                      |  |
| CBSB, UofK, Sudan       |  |
| fatma                   |  |
| gmochamah@kemri-w       |  |
| IPM                     |  |
| mamadou.kaba@uct.ac     |  |
| NABDA_Abuja             |  |
| NABDA_Ibadan            |  |
| shimane lekakake        |  |
| Tracy Meiring (Trainer) |  |

## Good phenotypic data are needed in order to assess and control for confounders

- **Example:** The association between **type 2 diabetes and the gut microbiome** is confounded by **metformin treatment**.



Forslund et al. *Nature*. 2015;528(7581):262-6.

Public CBSB, UofK, Sudan Options

Victoria Nembaware 13:38  
Please do ask questions about Nutrigenomics so we can pass these on to Prof Samar? @Kemri we have noted your questions

NABDA\_Ibadan 21:38  
Considering the situation in Africa, can a relatively healthy individual have a knowledge of her genome in order to know what to eat and what not to eat?

Victoria Nembaware 13:39  
Welcome Botswana and Abuja!

Send

Webcams

Run successfully for:

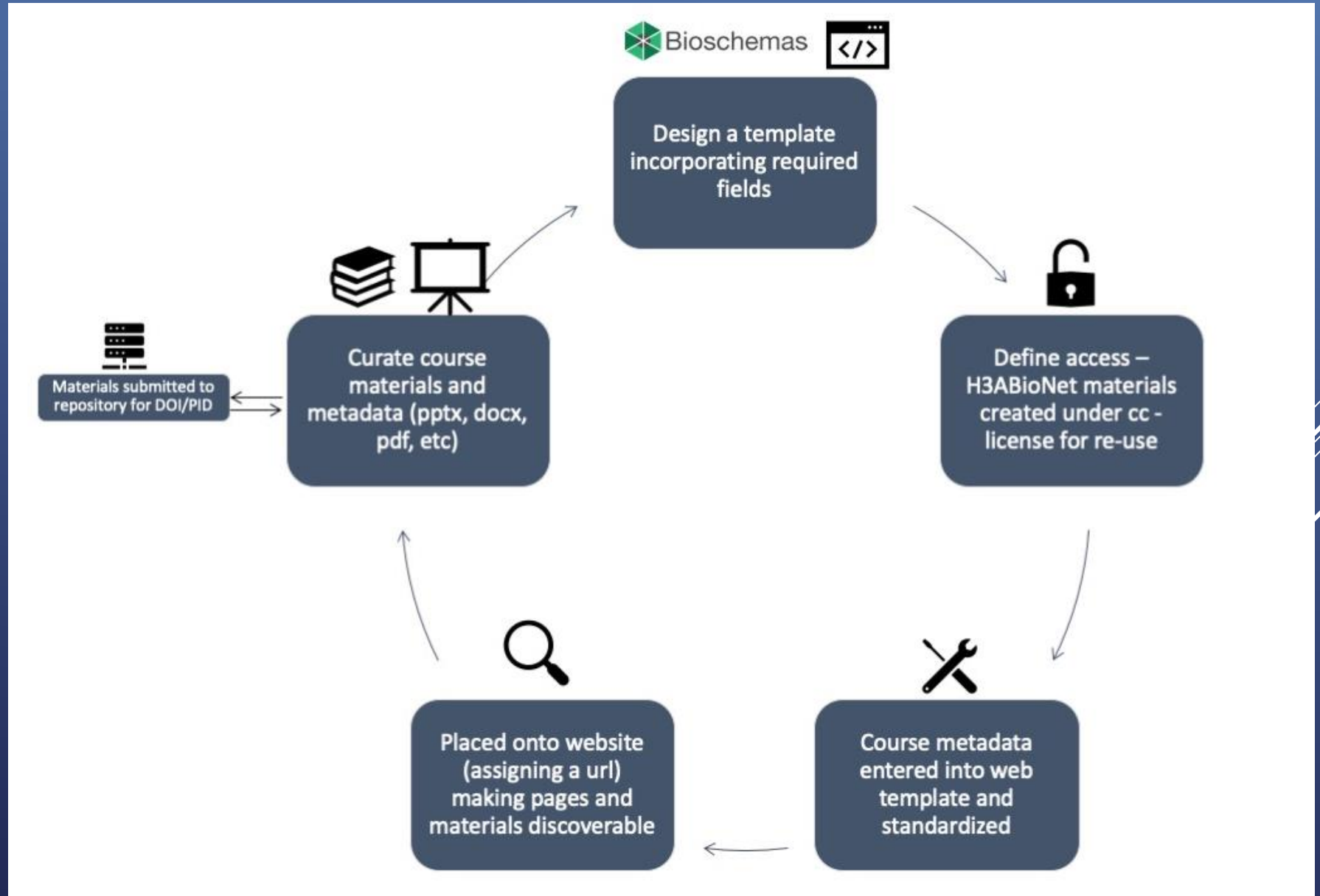
- Introduction to Bioinformatics (>1000 participants)
- Intermediate Bioinformatics (~300-400 participants)
- Genomic Medicine

# Increasing impact -making training materials FAIR






# Curating training materials

Need to curate materials and make them accessible for wider reach



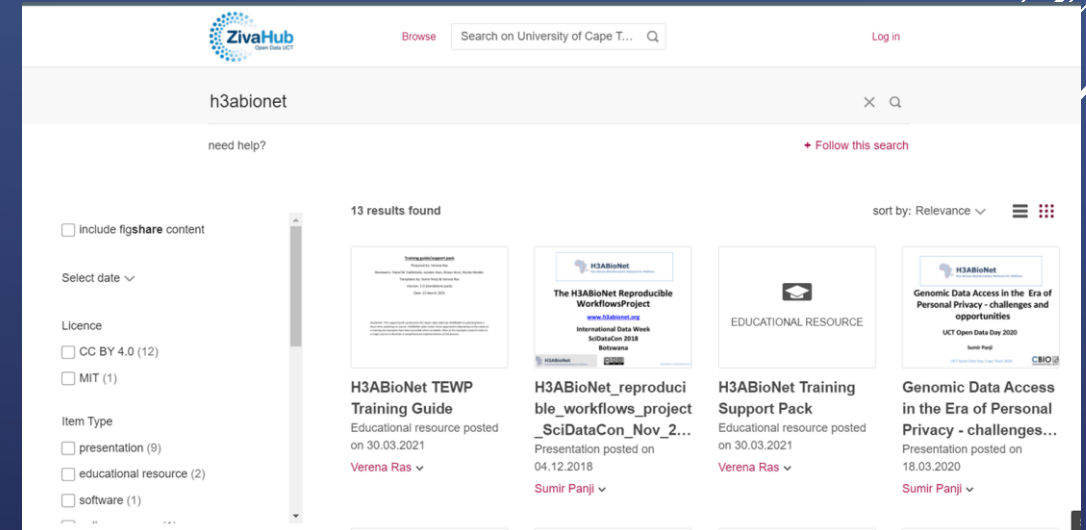
# Bioschemas example

- ▶ Making materials/content searchable online
- ▶ A lightweight way of structuring data online
- ▶ Created by a consortium of search engines to improve experience and search efficacy
- ▶ Training:
  - ▶ Course Instance
  - ▶ Course
  - ▶ Training Material

| Property                         | Expected Type                  | Description   | CD   | Controlled Vocabulary   | Example   |
|----------------------------------|--------------------------------|---|------|---|---|
| <b>Marginality: Minimum.</b>     |                                |   |      |   |   |
| <u>courseMode</u>                | Text<br>URL                    | <b>Schema:</b><br>The medium or means of delivery of the course instance or the mode of study, either as a text label (e.g. "online", "onsite" or "blended"; "synchronous" or "asynchronous"; "full-time" or "part-time") or as a URL reference to a term from a controlled vocabulary (e.g. <a href="https://ceds.ed.gov/element/001311#Asynchronous">https://ceds.ed.gov/element/001311#Asynchronous</a> ). | MANY | <ul style="list-style-type: none"><li>• "online", "onsite" or "blend of online and onsite"</li><li>• "synchronous" or "asynchronous"</li><li>• "full-time" or "part-time"</li></ul> |    |
| <u>location</u>                  | Place<br>PostalAddress<br>Text | <b>Schema:</b><br>The location of for example where the event is happening, an organization is located, or where an action takes place.<br><br><b>Bioschemas:</b><br>Location of the Course Instance. If the Course Instance is online, add the connection details as text  | ONE  |   |    |
| <b>Marginality: Recommended.</b> |                                |   |      |   |   |
| <u>endDate</u>                   | Date<br>DateTime               | <b>Schema:</b><br>The end date and time of the item (in ISO 8601 date format).  | ONE  | ISO 8601  |  |

# Training Material Curation

- H3ABioNet delivers a range of training courses
- Courses result in a range of training materials that could be beneficial to the community
- Materials currently only located on website, not easy to find or know when to re-use
- Aim to curate training materials and making them more accessible



# NGS academy course pages

## SARS-CoV-2 Bioinformatics Training

- Home
- About
- Courses ▾
- Resources ▾
- Contact

Click [here](#) for the training survey





The NGS Academy forms part of the Africa CDC Pathogen Genomics Initiative (PGI). This initiative is funded by the Bill & Melinda Gates Foundation. Our training initiatives are carried out in partnership with the African Society for Laboratory Medicine (ASLM).

### Course overview:

The SARS-CoV-2 Bioinformatics Training was a follow-up to our SARS-CoV-2 NGS training. The modules were more advanced bioinformatics analysis, data curation and dissemination from SARS-CoV-2 NGS data.

### Intended audience:

Personnel of national public health institutions on the African continent, carrying out analyses, and individuals involved in SARS-CoV-2 research.

|   |   |
|---|---|
| Pathogen Surveillance<br>            | Data Curation and Sharing<br>        |
| Advanced Bioinformatics Analysis<br> | Presenting Data to Inform Policy<br> |

A pre-course workshop was held in October 2021. A number of topics for the course were introduced, including:

1. Unix, command line
2. Git, version control
3. High performance computing

### Course curriculum:

Module 1 Workflows for SARS-CoV-2 analysis

Module 2 Data curation and sharing

Module 3 Advanced bioinformatics analysis

Module 4 Presenting data to inform policy

Recommended materials: Model bioinformatics analysis

### Course schedule:

[Pre-course workshop](#)   [Module 1](#)   [Module 2](#)   [Preparatory sessions](#)   [Module 3](#)

[Module 4](#)

| Date         | Course materials  | Training partner          | Interactive Sessions with Instructor/s   |
|--------------|---|---------------------------|--|
| 5-7 Oct 2021 | Unix, command line, git and version control. Introduction to HPC, software containerization, workflow tools | H3ABioNet/The Carpentries | Trainers: Gerrit Botha, Kauthar Omar, Lyndon Zass, Ruth Nanjala, Verena Ras, Ziyaad Parker<br>Helpers: Chaimae Samtal, Nihad Alsayed |

# Education summit projects

**Competencies**

**Course endorsement**

**Trainer resources**

**Train-the-trainer**

**Going virtual**

**Training in LMICs**

Development of  
resources to  
support trainers





# Trainer portal



# Trainer portal

## Trainer resources

### What skills are needed to be a good trainer?

Just because you are very familiar with a topic, doesn't mean you are necessarily able to teach it well. GOBLET has put together some useful resources for improving skills to become a good trainer. This includes a guidance document for new trainers, links to materials and papers from train-the-trainer (TTT) initiatives and to an online course developed collaboratively by GOBLET, H3ABioNet, ISCB, ELIXIR and EBI Training. Many of the materials are applicable for trainers across a range of disciplines but there is a focus on bioinformatics.

1. A [Trainer Guidelines Document](#) was developed at a Bioinformatics Education Summit in 2019 which has a number of useful tips for trainers
2. The key skills and knowledge bioinformatics trainers should have are included in the [GOBLET skills matrix](#)
3. Train-the-Trainer courses materials are available from [ELIXIR](#), [EMBL-EBI](#), [Carpentries](#) and other organizations who offer Train-the-Trainer courses. Courses by many of these providers are advertised on [ELIXIR TeSS](#).
4. An Online Train-the-Trainer course is currently being developed collaboratively by GOBLET, H3ABioNet, ISCB, ELIXIR and EBI Training. [COMING SOON]
5. Several papers or blogs have been published about Train-the-Trainer activities, both courses and resources, these include:
  1. [ELIXIR papers \(Via et al, 2019, Morgan et al, 2017\)](#).

This site uses functional cookies and external scripts to improve your experience. [My settings](#) **Accept**

4. Others:

1. [Madlung 2018](#)
2. [Attwood et al, 2019](#)

6. There are several resources for improving your presentation skills such as:

EMBL-EBI User Training Working Group (UTWG) [training tips](#)

How do I design and develop course/training materials? ⋮

How can I make my training materials Findable, Accessible, Interoperable, and Reusable (FAIR)? ⋮

How should I organise and facilitate training? ⋮

How should I deliver training? ⋮

How should I assess trainees? ⋮

How do I evaluate a course? ⋮

How do I endorse and accredit a course? ⋮

What should I consider when teaching bioinformatics to high school students? ⋮

Where can I find other sources of support? ⋮

Missing a Resource? ⋮

Acknowledgements ⋮

# Additional resources

- ▶ Need to identify suitable trainers
- ▶ Requests for longer term support
- ▶ Requests for mentorship



- ▶ Trainer Database

African Union AFRICA CDC NGS Academy Helpdesk Ticketing System

HOME ABOUT RAISE TICKET TRACK TICKET REGISTER TRAINER LOGIN

HOME > REGISTER TRAINER

## Register Trainer

Username\*

Required: 150 charac

Email\*

Postgraduate teaching level

Postgraduate

Professional Development

Undergraduate

First name\*

Last name\*

Category of specialisation  Years of experience  Remove

Category of specialisation  Years of experience  Remove

Category of specialisation  Years of experience  Remove


Add Area of Specialisation

Password\*

Note: By submitting your details, you consent to have your information shared with the helpdesk team and helpdesk users for the purpose of addressing their queries and related communication. Your de-, research and reporting purposes, with all personally identifiable information removed. Your personal information will only be shared with 3rd party contractors for the database please contact NGS\_Academy\_info@h3abionet.org. The Human Research Ethics Committee that has approved the database

### My Profile

Personal Details



**Perceval Maturure**  
Software Developer  
Qualification: None  
Nationality: Zimbabwe  
Timezone: None

[Website](https://uct-cbio.github.io/ngs-academy/)  
[github](#)  
[twitter](#)

Name: Perceval Maturure  
Email: mtrper001@myuct.ac.za  
Gender: Male  
Bio: Perceval joined CBIO as a Software Developer in 2021. He has served several roles which include Website and database-driven application development, IT Management, Tutoring Programming, and Mathematics. He is passionate about the use of Software, Database Technologies for scientific research and enthusiastic about data analysis. He is currently studying a PGDip in Cybersecurity (part-time) with the University of Cape Town in the School of Information Technology.  
Organisation: University of Cape Town  
Resides in: South Africa  
Receive emails: yes

| Username | Area of Specialisation                      | Years of Experience |        |
|----------|---|---------------------|--------|
| perceval | Bioinformatics and Computing Infrastructure | 2                   | delete |
| perceval | General Bioinformatics                      | 1                   | delete |
| perceval | Training & Mentorship                       | 2                   | delete |

# Additional resources

- ▶ Need to identify suitable trainers
- ▶ Requests for longer term support
- ▶ Requests for mentorship



- ▶ Trainer Database



- ▶ Trainee Helpdesk



# Helpdesk/Trainer database

## Trainer responses & the knowledgebase

### Submitting a helpdesk ticket

The screenshot shows the 'Raise Ticket' page. At the top, there are logos for African Union, AFRICA CDC, and NGS Academy. Below the logos is a navigation bar with links: HOME, ABOUT, RAISE TICKET, TRACK TICKET, REGISTER TRAINER, LOGIN. The main heading is 'Raise Ticket'. On the left, there is a sidebar menu with various categories, including 'Software Containerisation & Workflow Tools' which is currently selected. The page content is mostly blank, suggesting a form for submitting a ticket.

The screenshot shows a ticket details page. At the top, there is a table with ticket information:

|                    |  |                          |                      |
|--------------------|--|--------------------------|----------------------|
| Ticket ID          | NGS001                                     | Date & Time              | 2022-08-15 11:26     |
| Ticket Subject     | Docker                                     | Email                    | drperceval@gmail.com |
| Ticket Description | Containerization of workflows              | Organisation             | uct                  |
| Ticket Category    | Software Containerisation & Workflow Tools | Category of Organisation | Academia             |
| Ticket tags        | container                                  | Ticket Completed         | True                 |

Below the table, there is a section for the trainer's response:

**Subject: Containers**  
Reply: Each time u run a docker compose down, you lose data in the db container

Trainer: Perceval Maturure  
Date: Sept. 1, 2022, 11:34 a.m.

At the bottom right, there are two buttons: 'Tag and Close Ticket' and '<< Back to Full Ticket List'.

The screenshot shows the 'Track Software Containerisation & Workflow Tools Tickets' page. At the top, there are logos for African Union, AFRICA CDC, and NGS Academy. Below the logos is a navigation bar with links: HOME, ABOUT, RAISE TICKET, TRACK TICKET, REGISTER TRAINER, LOGIN. The main heading is 'Track Software Containerisation & Workflow Tools Tickets'. Below the heading, there is a search bar and a list of tickets. The first ticket is highlighted:

|                     |  |                         |
|---------------------|--|-------------------------|
| Back to all tickets | Software Containerisation & Workflow Tools | FAQ's/Tags<br>container |
|---------------------|--|-------------------------|

Below the search bar, there is a 'Show 10 entries' dropdown and a 'Search:' input field. The main table has the following columns: Ticket ID, Ticket Subject, Ticket Category, Date & Time, and Action. The first row of data is:

| Ticket ID | Ticket Subject | Ticket Category   | Date & Time      | Action                            |
|-----------|----------------|---|------------------|-----------------------------------|
| NGS001    | Docker         | Software Containerisation & Workflow Tools<br>tags: container | 2022-08-15 11:26 | View Details & Trainer's Response |

At the bottom, there is a 'Showing 1 to 1 of 1 entries' message and a pagination control with 'Previous', '1', and 'Next' buttons.

# Other useful resources

- Training guide & support pack

## H3ABioNet TEWP Training Guide

[Cite](#) [Download \(1.58 MB\)](#) [Share](#) [Embed](#) [+ Collect](#)

**Version 1 (old)** [▼](#) [Educational Resource](#) posted on 30.03.2021, 09:20 by [Verena Ras](#), [Sumir Panji](#), [Faisal M. Fadlemolma](#), [Shaun Aron](#), [Lyndon Zass](#), [Nicola Mulder](#)

## H3ABioNet Training Support Pack

[Cite](#) [Download \(1.14 MB\)](#) [Share](#) [Embed](#) [+ Collect](#)

**Version 2** [▼](#) [Educational Resource](#) posted on 30.08.2021, 11:32 by [Verena Ras](#), [Sumir Panji](#), [Nicola Mulder](#), [Faisal M. Fadlemolma](#), [Lyndon Zass](#), [Shaun Aron](#)

- Starting Bioinformatics from scratch
- Building infrastructure and training in LMICs

## Challenges and Considerations for Delivering Bioinformatics Training in LMICs: Perspectives From Pan-African and Latin American Bioinformatics Networks

[Verena Ras](#)<sup>1</sup>, [Patricia Carvajal-López](#)<sup>2</sup>, [Piraveen Gopalasingam](#)<sup>2</sup>, [Alice Matimba](#)<sup>2</sup>, [Paballo Abel Chauke](#)<sup>1</sup>, [Nicola Mulder](#)<sup>1</sup>, [Fatma Guerfali](#)<sup>3</sup>, [Victoria Dominguez Del Angel](#)<sup>4</sup>, [Alejandro Reyes](#)<sup>5</sup>, [Guilherme Oliveira](#)<sup>6</sup>, [Javier De Las Rivas](#)<sup>7</sup> and [Marco Cristancho](#)<sup>8\*</sup>

OPEN ACCESS

EDITORIAL

## Ten simple rules for organizing a bioinformatics training course in low- and middle-income countries

[Benjamin Moore](#) , [Patricia Carvajal-López](#), [Paballo Abel Chauke](#), [Marco Cristancho](#), [Victoria Dominguez Del Angel](#), [Selene L. Fernandez-Valverde](#), [Amel Ghouila](#), [Piraveen Gopalasingam](#), [Fatma Zahra Guerfali](#), [Alice Matimba](#), [Sarah L. Morgan](#), [Guilherme Oliveira](#), [Verena Ras](#), [ ... ], [Nicola Mulder](#) [ [view all](#) ]

Published: August 19, 2021 • <https://doi.org/10.1371/journal.pcbi.1009218>

# Acknowledgements

Education summit community and project leads

Summit organisers, session chairs, facilitators and participants

H3ABioNet training team & funder (NIH)

NGS Academy training team & funder (BMGF)

Project leads: Cath Brooksbank, Bruno Gaeta, Russell Swartz, Sarah Morgan, Alice Matimba, Dusanka Nikolic, Michelle Brazas, Celia van Gelder, Verena Ras, Patricia Carvajal Lopez, Venkata Satagopam

