

Global Health Data Science Hub working group meeting notes

4 April 2023

1. Progress on priority areas and content and curation log review

- No new items were added by the group
- Agreed that a webinar on machine learning and AI is a priority
- Agreed that it would be a priority to focus on developing resources on data sharing, potentially an article on why it is important to have a data sharing policy and how to create one
- icddr,b would be a good case study, in addition to Fiocruz and its Open Science policy
- Development of a data sharing SOP would be very useful

Structural changes to the Hub are:

- A new menu item has been created for the project lifecycle and individual pages for each stage are sub menu items under this heading
- The data re-use project lifecycle diagram is no longer on the Home page and now appears on the project lifecycle main page
- A resources menu item has been created and all resources are now listed in the table. The longer term plan is to make this table searchable.
- A pathfinder studies menu item has been created and a number of resources relating to those have been added to the Hub
- Another layer has been added to the Connect and Collaborate table to capture other data science initiatives in Africa, Asia and Latin America and the Caribbean. The longer-term plan is to add these as different coloured pins on the map.

2. Keeping the Hub current

The group discussed ways in which we can keep the Hub current. Key points were:

- We need to understand what impact the Hub is having
- The group suggested that a 'Pathways to impact' menu heading would be a useful addition to the Hub
- An 'Impact' page is currently being developed for all the TGHN digital Hubs and that we can tailor this to meet our needs
- We have included and sign-posted to other data science programmes and initiatives in a list on the Hub. The next priority will be to consider how we can engage with other initiatives and encourage them to use the Hub to showcase their work
- We could try to invite some data science programme leads to be members of the working group
- Could we explore engaging with the HDSS (Health and Demographic Surveillance System) across regions given the importance of the data that they collect to so many researchers in health

3. Hub metrics

Hub metrics for March were not available at the time of the meeting and will be reviewed at the May meeting.

4. Wider initiatives

The team are undertaking a process to identify essential data science skills by crowd consensus, and develop an essential data science skills curriculum.

It is a 3-stage process:

1. Landscaping and knowledge gap analysis
2. 2 crowd consensus surveys
3. Validation of results with the community of practice

The current status of this work is as follows:

- A protocol has been developed
- A landscaping and knowledge gap analysis survey is in development. Landscaping will be supported through interviews with key stakeholders and a literature review. The findings from this survey will help us to build out content on the Hub and signpost members to existing resources
- Regional teams will shape the protocol and the surveys
- It will be important to factor in the fast-paced developments in the use of AI in data science to ensure there is an evolutionary element.
- We will consider ways in which the working group can contribute to this process