Contribution ID: **116** Type: **Talk**

Implementation of data standards for African genomics data by H3ABioNet

Tuesday, 1 December 2020 14:00 (30 minutes)

The reducing cost of genomic data generation has outpaced Moore's Law resulting in many countries starting national genomic initiatives to better represent their populations and create the foundations for precision medicine programs. Programs such as H3Africa seek to generate and leverage genomic data for specific health related research within diverse African populations. Genomic data, unlike biological samples, is captured and stored digitally, and copies of these data can indefinitely be provided to multiple researchers for use in multiple studies.

In order for genomics data to be integrated and be used meaningfully by the scientific community, standardized attributes that define its collection provenance, rich meta-data and conditions of use need to be explicitly provided. H3ABioNet, a Pan-African Bioinformatics Network to support the H3Africa program, has been working on developing various standards such as Case Reporting Forms for phenotype collection provided as REDCap instruments and mapped to ontologies. Working with international initiatives such as the Global Alliance for Genome Health (GA4GH), H3ABioNet is contributing to the adoption and refinement of standards such as the Data Use Ontology for H3Africa genomics data. To facilitate the ease of finding African specific genomic and genetic variation datasets, an African Microbiome portal and an African Precision Medicine portal are being created by H3ABioNet that curate African specific data from various sources for inclusion in these portals. An African Genome Variation Database using Open CGA standards to house H3Africa specific data is being created for research groups with specific focus areas such as rare diseases. Genomic data processing and curation for African COVD-19 is being undertaken by network members and data submitted to the international GISAID platform.

Various outputs from different H3ABioNet projects are being assessed for Findable, Accessible, Interoperable and Reusable (FAIR) standards and what is required for these outputs to be FAIR such as the adoption of Bioschemas for training materials and versioning of code. Creation of a robust data ecosystem that utilizes established standards requires human capacity to be developed. H3ABioNet has been providing a series of data management planning training workshops to H3Africa personnel and students to enable better planning and preservation of research outputs.

Student?

No

Supervisor name

Supervisor email

Primary author: PANJI, Sumir (H3ABioNet / CBIO / UCT)

Presenter: PANJI, Sumir (H3ABioNet / CBIO / UCT)

Session Classification: DIRISA

Track Classification: DIRISA