Centre for Infectious Disease Genomics and One Health
Blusson Hall,
Simon Fraser University,
Burnaby

Meeting Minutes – Needs Assessment Interview

Project Name: [] Data Standards Development
Interviewee Name: Name
Affiliation / Role:
Date: 2025

Overview

Summary of the Project/Pilot aims and context

The Centre for Infectious Disease Genomics and One Health (CIDGOH) has been engaged to develop a data specification/conduct a pilot assessment for the [project name].
The goal is to assess the current (contextual) data ecosystem supporting the [project name], in order to develop an interoperable data specification/recommendations and a conceptual model for contextual data.

Purpose of the interview

To identify current data practises, data size, type and variety, as well as the processes involved to inform the needs assessment component of developing a data standard. We aim to enhance collaboration, improve data quality, and support informed decision-making in [project name]

Consent (if applicable)

Notes to be taken, no recording

Stakeholder context

	Details
Job Title	Researcher
Organisation type	(e.g. Research, Government, Policy, Agriculture etc)

Role in data lifecycle	(e.g. sample/data collection, curation, analysis, use, dissemination etc)

Needs Assessment

The following questions are design to assess the complexity and variety of processes that affect and can offer insight to useful contextual data.

Scope

1. What are the questions your project/initiative is trying to answer?
   Research, surveillance? Pathogen/target?
   Assay - amplicon, metagenomics, marker-based, culture & WGS?

2. Who are your partners?
   Informs data flow, sharing, harmonization needs

Sampling Strategy

3. What kinds of samples are you collecting?
   Hosts (anatomical materials/parts/body products)
   Environments (materials/sites)
   Collection devices/methods

4. Are you using any previously collected samples?
   Purpose of sampling/sample bias, data reuse

5. What contextual data associated with the samples is being collected/considered?
   a. Host
   b. Physico-chemical attributes
   c. Land use, water table and/or tidal data
   d. Presampling activities
   e. Experimental interventions
   f. Geo-spatial data
   g. Sampling frequency and duration

6. Do you have a sample plan? Is the sample plan, experimental design, or any other material available online?

Specimen processing and Storage

7. Are the samples subjected to any special processing before sequencing libraries are prepared?
   Filtering, enrichment, culturing, etc

8. If samples are cultured, what kinds of media and conditions are being used?

9. How are samples stored, and for how long
   Cold chain, comparability, audit

Library Preparation

10. What are your library preparation processes?
    Extraction, enrichment

11. How are libraries being generated?

12. What are your controls?
    endogenous controls, replicates, synthetic constructs/data

Sequencing and Bioinformatics

13. What sequencing instruments are being used?

14. What bioinformatics processing is being performed
    a. Are workflows custom scripts or established/community-developed?

    b. Specifically what tools are being used for:
        i. Raw sequence processing/dehosting

        ii. Consensus

        iii. Callers

        iv. Databases (reference or otherwise)

        v. QC metrics

Associated Data Types

15. Are there other data types or characterizations available for the samples/sequences?
    a. Physico-chemical properties:
       Temp, pH, oxygen, salinity, nitrogen
    b. Biological properties:
       Enzymatic, staining, marker identification, colony counts, AMR, typing (serotype), matter composition, mollusc shell length, therapeutic history, etc

Data Management

16. How is the contextual data being captured?
    Spreadsheets, LIMS, RedCap, other?

17. Are any data standards being used to structure the data?

18. Where is the data being stored?

19. What is the data flow/life cycle

20. During the data life cycle, will it need to be transformed?

Data Sharing & Governance

21. Among your network(s), are there any data use limitations/restrictions? Due to:
    a. Limitations/biases of the data

    b. Data governance

22. During data flow, will the data need to be transformed?
    Curation and automation needs

23. Is there a plan to share the data publicly? If no, why? What are the perceived challenges?
    Governance/privacy/equity/policies/technical issues

Future Work (anticipating future needs)

24. What data types would you want to collect/foresee integrating in the future?

25. What are your main contextual data challenges?

26. What would be on your contextual data wishlist?

Summary

Summary to be included here

Data ecosystem	**Details**
Types of data currently collected	
Current data workflow	
Systems, standards and/or tools currently used	
Key outputs/reports that data supports	

Challenges	**Details**
Pain points	
Missing data fields	
Integration challenges	
Data governance and permissions	