Improving Openness And Interoperability In Public Health Bioinformatics

A Global Coalition.

Designed by Charlie Barclay, Emma Griffiths and Rhiannon Cameron as part of PHA4GECon 2025 pre-conference workshops

PHA4GECon 2025: Exercise Data Needs Assessment

Identify the data needs and get used to asking questions!

Objectives

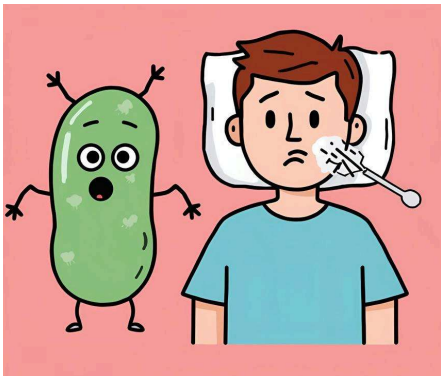
Participants will **simulate a needs-assessment interview** to elicit requirements for a new **contextual data standard** for *E. griffithsyi*.

- "**Interviewers**" will practice asking targeted, open-ended questions to uncover data needs, priorities, and practical constraints.
- "**Stakeholders**" will role-play experts from different domains with specific perspectives and needs.

By the end, participants should:

- Understand how stakeholder priorities shape a data specification.
- Recognise the diversity of contextual data requirements across disciplines.
- Practice recording and categorizing identified needs (essential vs. optional, operational vs. research).

Scenario



A novel pathogen, *E. griffithsyi*, has recently been detected in multiple regions. Early evidence suggests it may spread through both human-to-human contact and environmental reservoirs, but its virulence, incubation period, and transmission routes are still unclear. Public health agencies and research networks agree that rapid data harmonisation is essential to coordinate surveillance, laboratory investigations, and research responses.

Rather than rushing directly into data collection, we have committed to performing a Data Needs Assessment first — to identify what contextual information must be captured to support surveillance, risk assessment, and genomic analyses

Instructions

Split into groups of five:

- Two interviewers (needs assessment team)
- Three stakeholders (roles below)

Conduct a needs assessment interview

- Interviewers should ask about:
  - The data each stakeholder collects or needs.
  - Challenges with current data capture or sharing.
  - What metadata or context would help interpret their results.
  - Any constraints (technical, ethical, policy).
- Stakeholders answer according to their role card. Feel free to add realistic details and challenges!

Summarise findings

- Interviewers note:
  - **Essential** vs. **nice-to-have** data elements.
  - Common **themes** across roles.
  - **Gaps** that would need to be resolved by a new schema.
  - **Data flow** across the lifecycle.