



Data Quality WG

Q1 2022 Updates



Clair Blacketer



Let's Go Back...To The Future





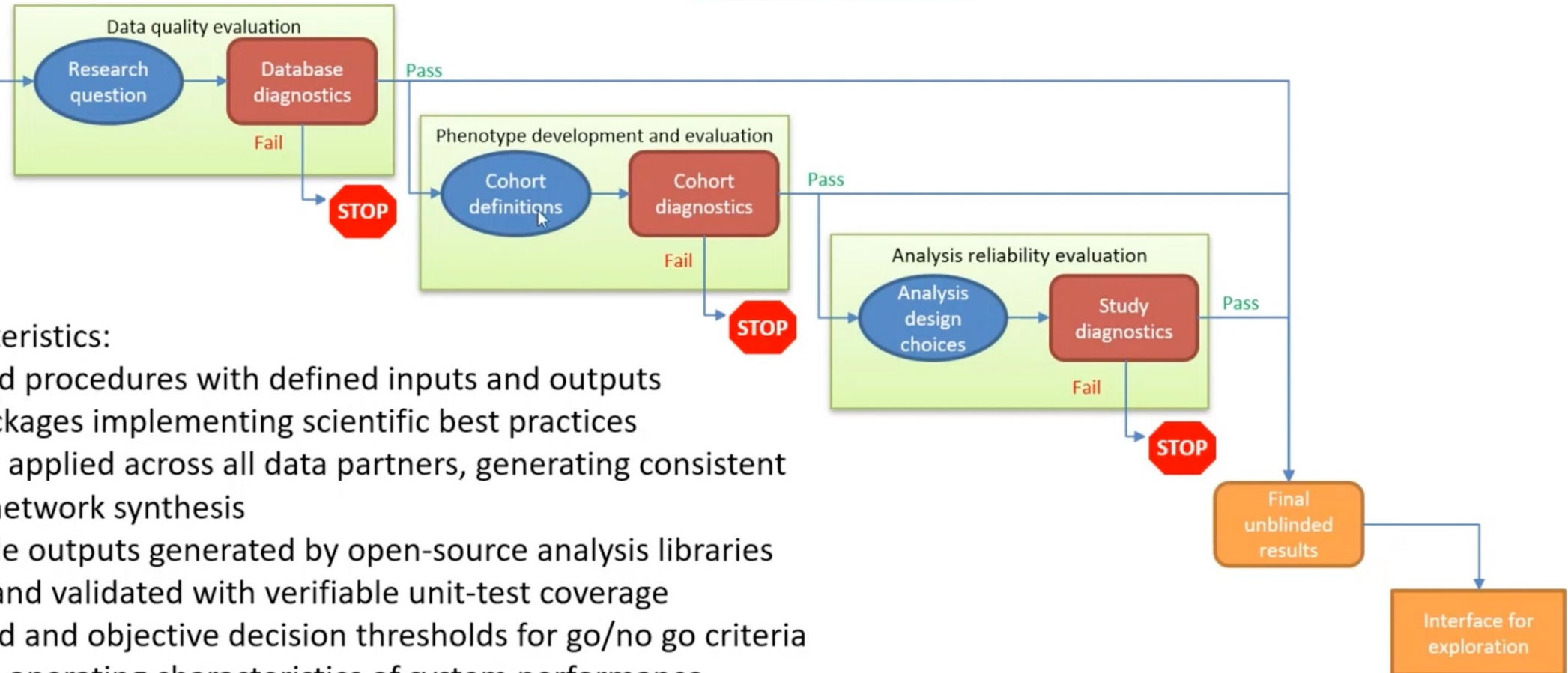
Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:

- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

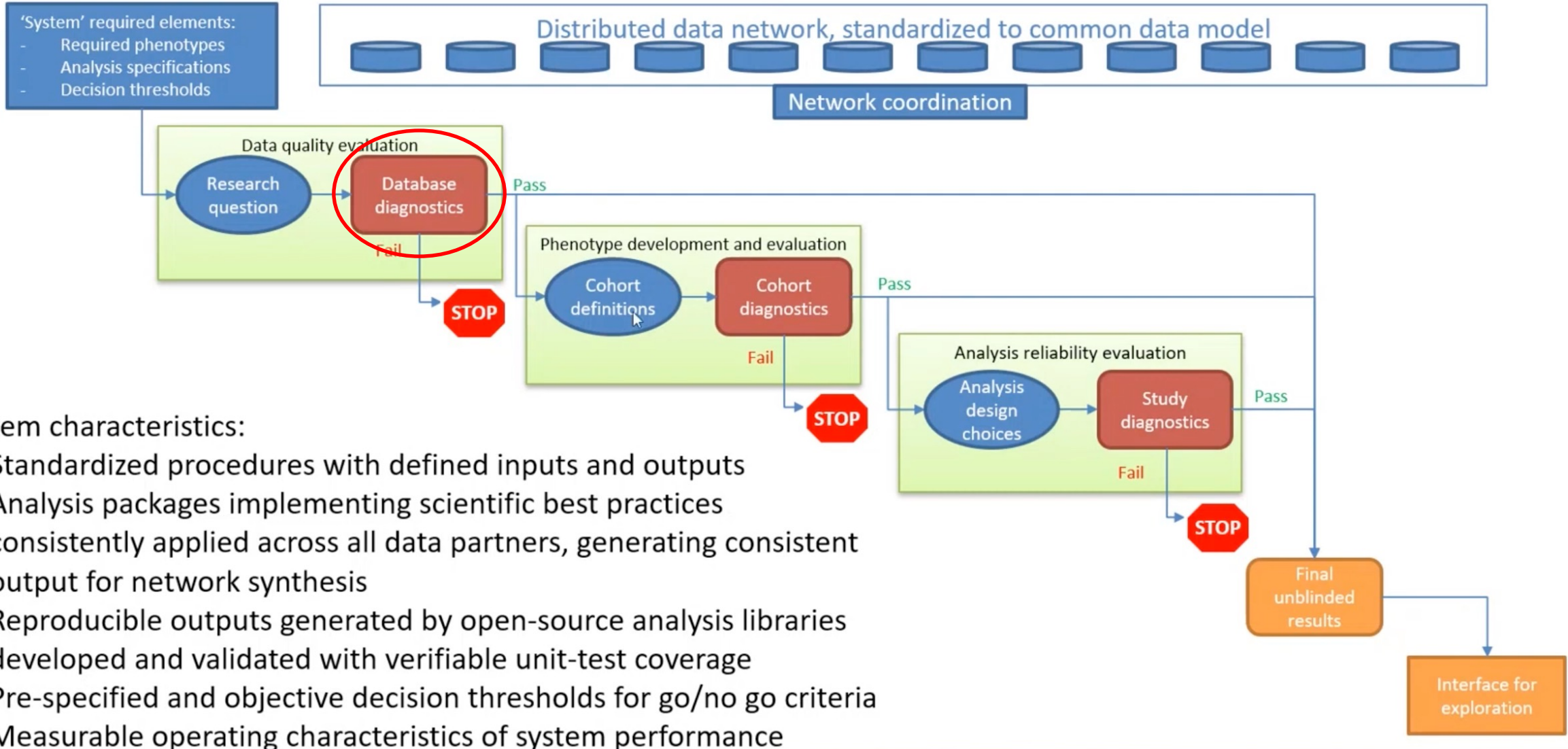


System characteristics:

- Standardized procedures with defined inputs and outputs
- Analysis packages implementing scientific best practices consistently applied across all data partners, generating consistent output for network synthesis
- Reproducible outputs generated by open-source analysis libraries developed and validated with verifiable unit-test coverage
- Pre-specified and objective decision thresholds for go/no go criteria
- Measurable operating characteristics of system performance



Engineering open science systems that build trust into the real-world evidence generation and dissemination process





Database Diagnostics

Does a database
have the necessary
elements required
to perform a
study?



Database Diagnostics

Does a database
have the necessary
elements required
to perform a
study?

Are those
elements of a high
enough quality
such that we can
trust the evidence
generated?



Database Diagnostics

Does a database
have the necessary
elements required
to perform a
study?

Are those
elements of a high
enough quality
such that we can
trust the evidence
generated?



2022 OKR: Data Quality

1. Design and Implement the Data Quality Evaluation Step of the OHDSI Evidence System

1Q2022 Key Results:

1. Given a database has necessary elements to run a study, design a study-specific filter for the DQD.
2. Close at least 50% of issues and pull requests on the DQD github, prioritizing high-need bugs and new features added by Odysseus.
3. Update documentation to reflect the new features in the DQD application.