

Harnessing Generative AI for PHL



David A. Silva, MS, DLM(ASCP)^{CM}

Quality Manager | Public Health Laboratory

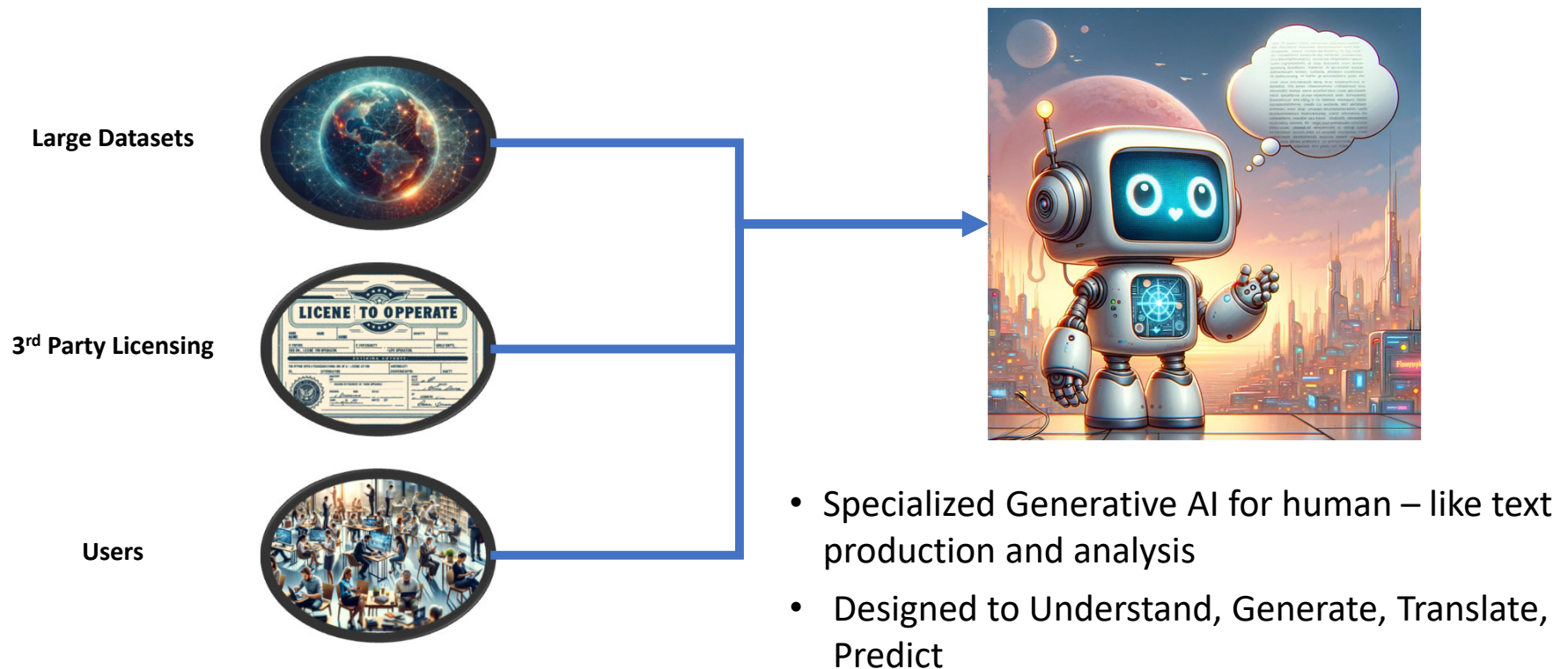
Dallas County Health and Human Services

What is Generative AI?

DALL-E/CHATGPT-4 Prompt: Generate a watercolor image of dapper Latino laboratory manager experiencing writer's block. The writer is completely bald with a full salt and pepper beard, capturing his thoughtful and mildly frustrated expression as he faces the creative challenge.



What is a Large Language Model (LLM)?



Essential skills of a LLM User



Prompt Design



Critical Thinking



Domain Expertise



**Adaptability &
Experimentation**

Art and Science of Prompt Design

The process of creating prompts that elicit the desired response from a LLM



ZERO SHOT

Querying the LLM
knowledge/training set

How does Mexico celebrate the
Christmas Holidays?



FEW SHOT

Provide guiding examples
for the model

David loves hot sandwiches and
nachos. Where should he go
eat?



CHAIN-OF-THOUGHT

Deconstructing goals into
stepwise tasks

Explain the chemical process of
baking a cookie at each step

Analysis of Prompts Examples



Zero Shot

What are the CLIA regulations governing quality control practices?



Few Shot

Outline a policy for quality control management and it should include:
1) Responsibilities 2) CAPA Processes 3) Documentation process



Chain-of-Thought

A patient received an incorrect test result, potentially impacting their treatment plan. Step-by-step, explain how you would investigate this incident, identify the contributing factors, and implement corrective actions to ensure similar errors don't happen again. Consider potential impact on accreditation and patient safety throughout your thought process.

Anatomy of a Prompt

ESSENTIAL



NICE TO HAVE



Task

Clearly defined goal
Analyze, summarize, generate, etc.



Context

Control the scope of response
Background, Success, Environment



Format

Output Coding language, table, template, bullet points etc.
Constraints (word limitations, etc)



Persona

Adopt a point of view
Delivers nuanced output



Tone

Directly influences diction
Assertive, formal, courteous, etc.

Effective Prompt, Better Output

Weak

How do I address complaints made in my workplace.

"Listen Carefully: Give the complainant your full attention. Listening carefully shows respect and may provide important details about the issue."

Serviceable

You are a laboratorian in a public health laboratory. Generate an outline for a document to manage complaints. Limit the outline to a maximum of 10 headers.

Here's an outline that can serve as a guide:

- 1. Purpose of the document*
- 2. Importance of handling complaints in a public health laboratory*

Better

You are a **clinical quality manager** responsible for ensuring **CLIA** compliance in a public health laboratory. Generate a comprehensive **template for a policy level document** to govern a complaint management system. Ensure it is **compliant with 42 CFR 493.1233** and contains a process for tracking complaints.

1. Introduction

1.1 Purpose

This document outlines the policies and procedures for managing complaints received by [Laboratory Name], ensuring compliance with 42 CFR 493.1233 and promoting continuous quality improvement in laboratory services.

1.2 Scope

This policy applies to all complaints related to the quality, safety, and conduct of laboratory services and operations.

2. Policy Statement

[Laboratory Name] is committed to providing high-quality laboratory services. All complaints are treated seriously, investigated thoroughly, and resolved promptly to maintain the integrity and quality of laboratory operations.

3. Definitions

- **Complaint:** Any written or oral grievance regarding the quality or safety of laboratory testing services.
- **Complainant:** An individual or entity lodging a complaint.

4. Roles and Responsibilities

- **Clinical Quality Manager:** Oversees the complaint management process, ensuring compliance and resolution.
- **Laboratory Staff:** Responsible for reporting complaints to the Clinical Quality

Technical Writer

Summarize information and generate high – quality drafts across the document hierarchy

- Deconstruct/translate published documents into digestible bites (research article, FDA-LDT Proposal, etc.)
- Improve abstracts and grant applications for clarity, thoroughness and meeting constraints
- Identify and improve compliance gaps in policy-level documents
- Identified growth opportunities by comparing test menu to Health Needs Assessment



Communications Specialist

Guide composition, rhetoric and presentation strategies tailored to an intended audience

- Compose comprehensive responses to inquiries or complaints
- Create social media posts or event promotions
- Suggest outlines and scripts for presentations



Educator

Customize learning by converting information into a personalized format

- Explain safety regulations, identify overlaps and key differences in layman terms
- Created study plans for certifications
- Simulate a computer-adaptive testing
- Improved assessments (i.e. quizzes, exercises) based on governing documents



Data Analyst

Fast-track the processing and evaluation of data to aid strategic decision-making and problem-solving

- Produce a modified SQL code for an enhanced LIMS query
- Debug Excel functions, explain itself and suggest alternative approaches
- Consult on determining root cause of Limit of Detection issues



Workforce Consultant

Help express or translate workforce vision into plans, strategies and actionable goals

- Generated question banks for a specific job role that incorporate a manager's vision
- Research retention strategies and draft policy language
- Develop effective SMART goals to tailored to a job role



Pitfalls to Avoid

“The only time I ever went down was when a machine was easing at the wrong time” – Cooper, *Interstellar*



**TURNING OFF
YOUR BRAIN**



**EXPECTING
PERFECTION**



BIAS CREEP



**WORKING
ALONE**



OVERSHARING

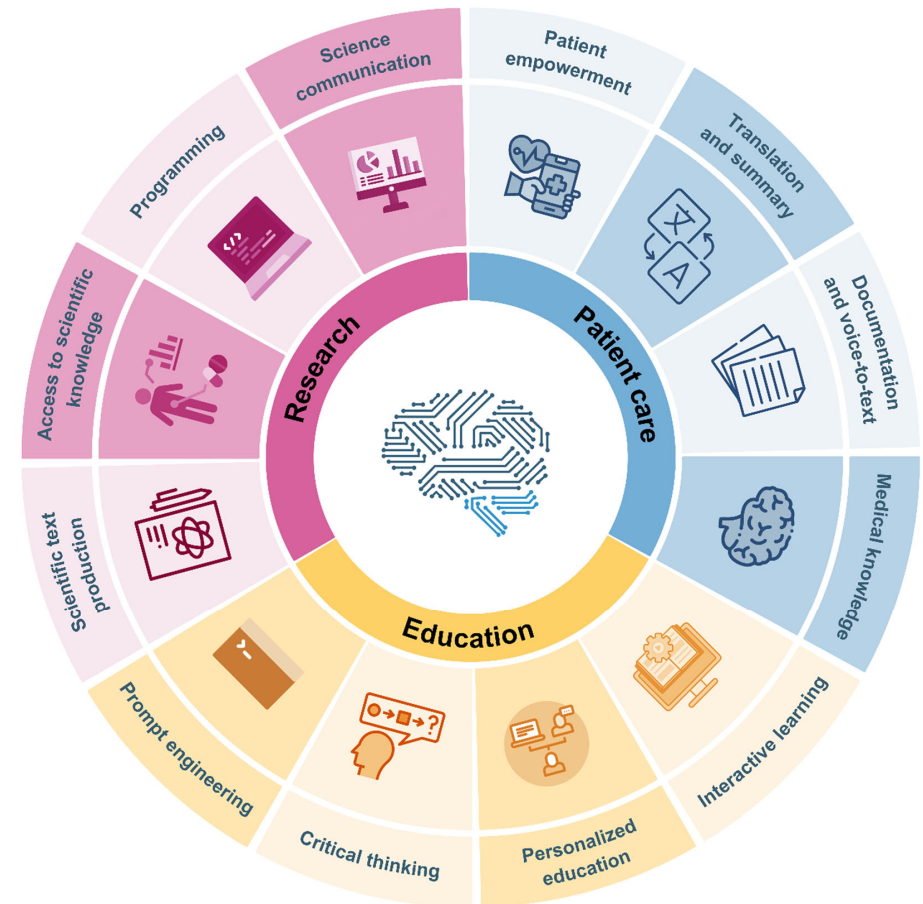


Your conversations are processed by human reviewers to improve the technologies powering Bard. Don't enter anything you wouldn't want reviewed or used.

[How it works](#) [Dismiss](#)

Where are we going?

- Custom LLMs
- Comprehensive Document System Review
- Personalize Training/Competency Assessments
- Strategic Planning (Model scenarios, assess risks, etc.)
- Improving Customer Service



Clusmann, J., Kolbinger, F.R., Muti, H.S. *et al.* The future landscape of large language models in medicine. *Commun Med* 3, 141 (2023). <https://doi.org/10.1038/s43856-023-00370-1>



Artificial Intelligence (AI) in Public Health

FEB-7 2024

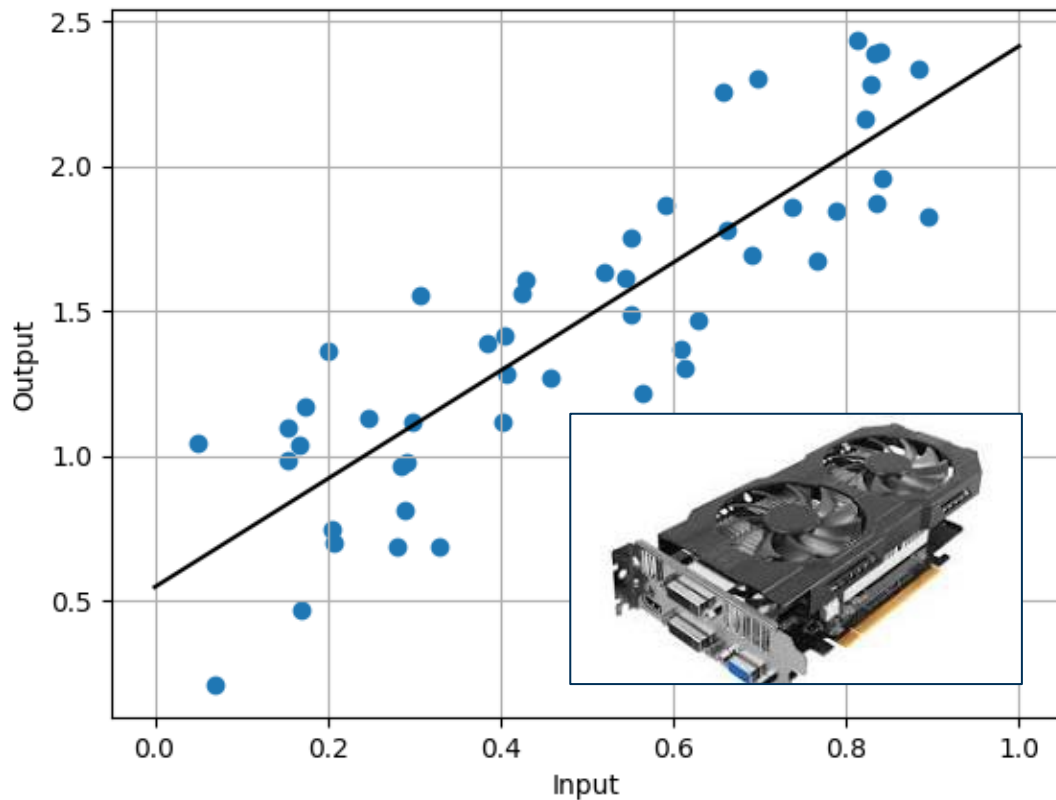
J. True Merrill, Ph.D.

True.Merrill@gtri.gatech.edu | 404-407-7988

Demystifying artificial intelligence

Deep learning is fancy curve fitting ...

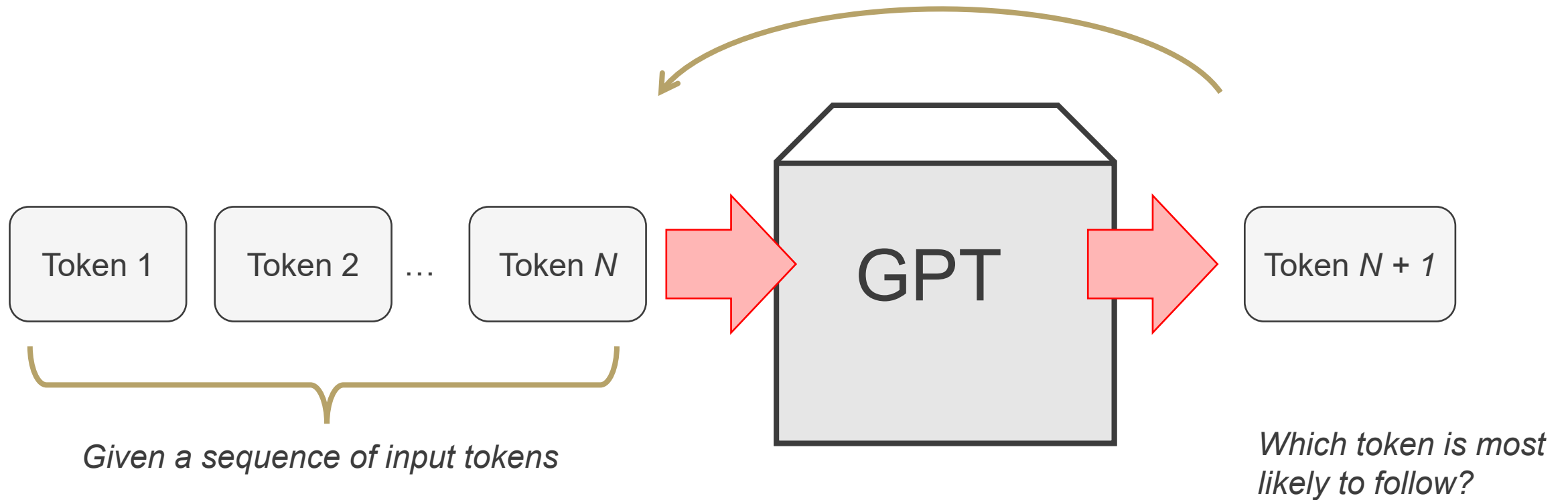
$$y = m x + b$$



- Many AI tasks can be written as a prediction problem. *Given an input x , what is the most likely y ?*
- AI models are functions that map from inputs to outputs
- The function's behavior is controlled by a set parameters, called *weights*.
 - **Linear model**: two weights: (m, b)
 - **Mixtral-8x7B**: 7 billion weights
 - **GPT-3 175B**: 175 billion weights
- Training an AI is a special case of curve fitting.

What are LLM's

Takeaway: Large Language Models (LLM's) are mathematical models that solve a “next token” prediction problem.



Demonstration of Mixtral-8x7B

Locally hosted private general-purpose open source LLM



- **Mixtral-8x7B** LLM hosted on Georgia Tech Research Institute's HPC cluster
- Query 1: I am going to post a protocol for DNA extraction. Later I would like to ask you questions about it. Do you understand?
- Query 2: <the protocol>
- Query 3: Tell me what equipment I need to follow this protocol.
- Query 4: What safety measures do I need to follow while performing this procedure?

Customizing an LLM for public health

For techniques to tune an LLM for a particular technical domain.



- **Prompt engineering:** Construct an optimized input prompt to get domain-specific responses from chatbots.
- **Fine tuning:** Starting from a pre-trained model, retrain it on domain-specific data. Only allow a small fraction of the weights to be adjusted.
- **Retrieval Augmented Generation (RAG):** use a retrieval system to find relevant documents or passages related to a given input, then use those documents to generate a more informed response.
- **Custom agent:** train an LLM from scratch on input data.

GTRI – A core part of the CDC PGCOE

CDC announces \$90M funding to support new Pathogen Genomics Centers of Excellence network

[Print](#)

CDC announced 5-year awards to five public health departments to establish the Pathogen Genomics Centers of Excellence (PGCoE) network on September 20, 2022. The PGCoE network is meant to foster and improve innovation and technical capacity in pathogen genomics, molecular epidemiology, and bioinformatics to better prevent, control, and respond to microbial threats of public health importance. The network also represents an unprecedented opportunity to expand and deepen collaboration between U.S. public health agencies and academic institutions to form a national resource to better prevent, control, and respond to microbial threats of public health importance.



[View Larger](#)

Hierarchy of needs

Data Science Hierarchy of Needs

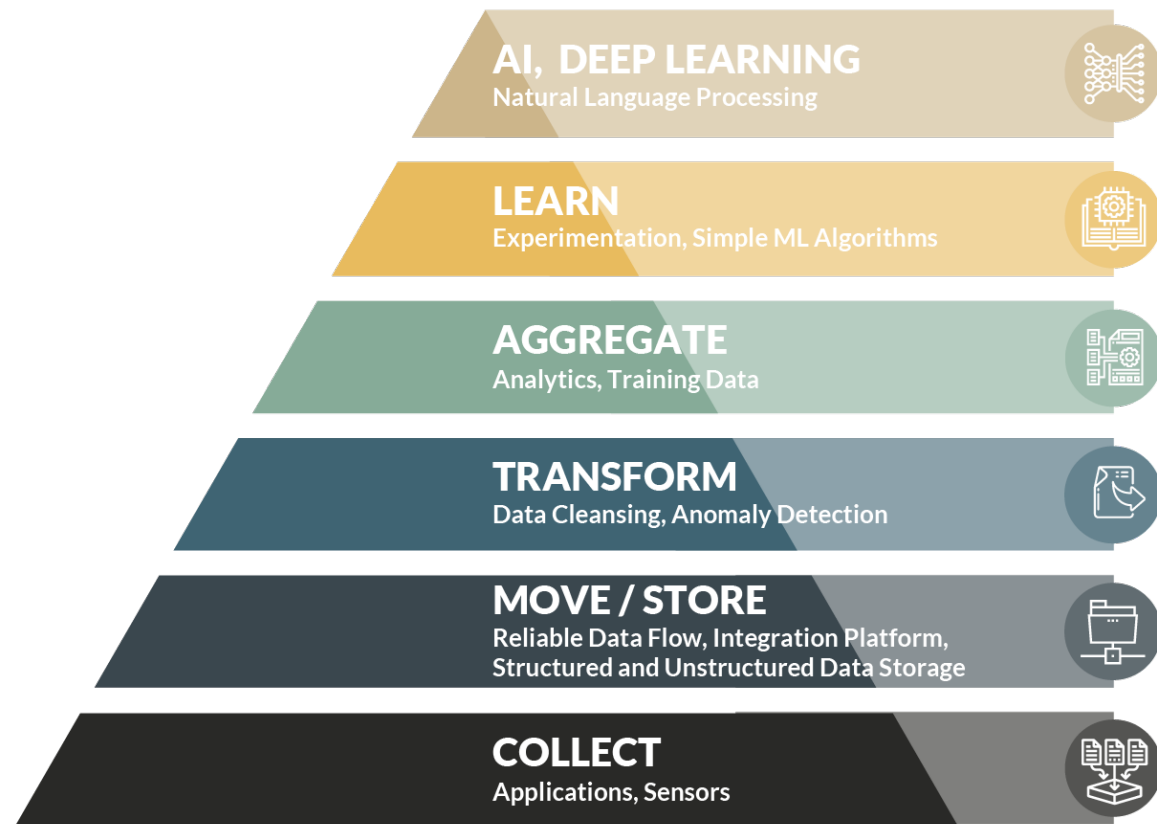
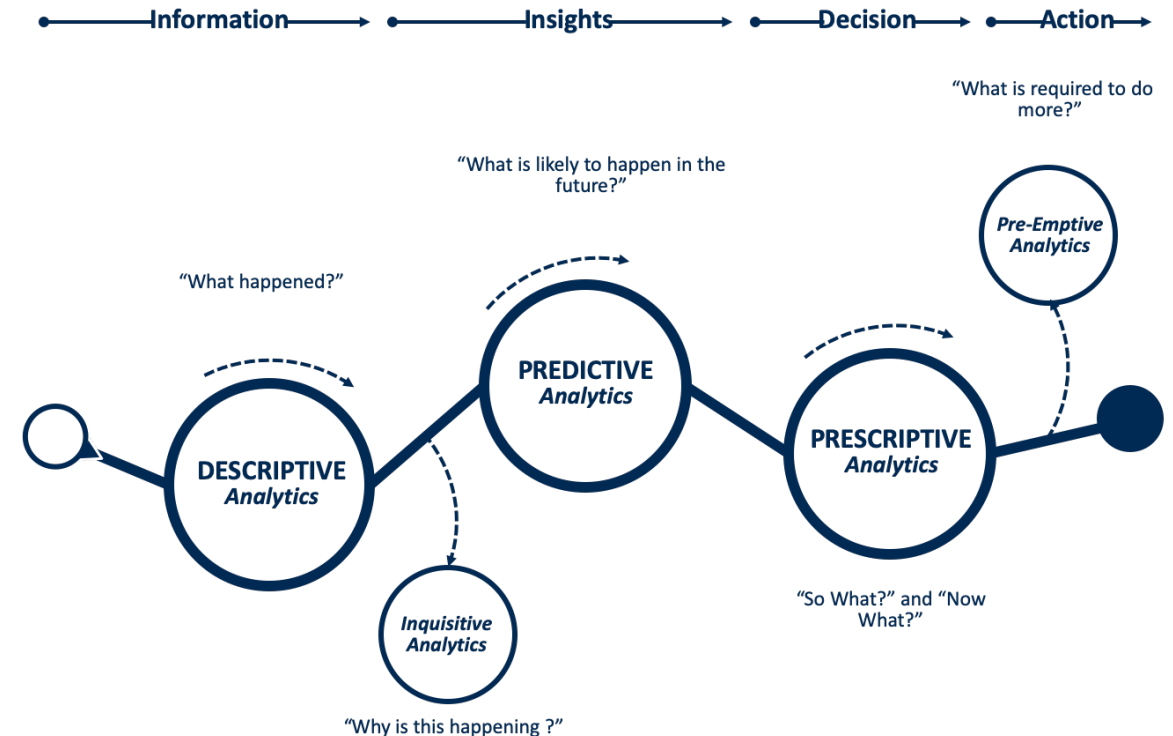


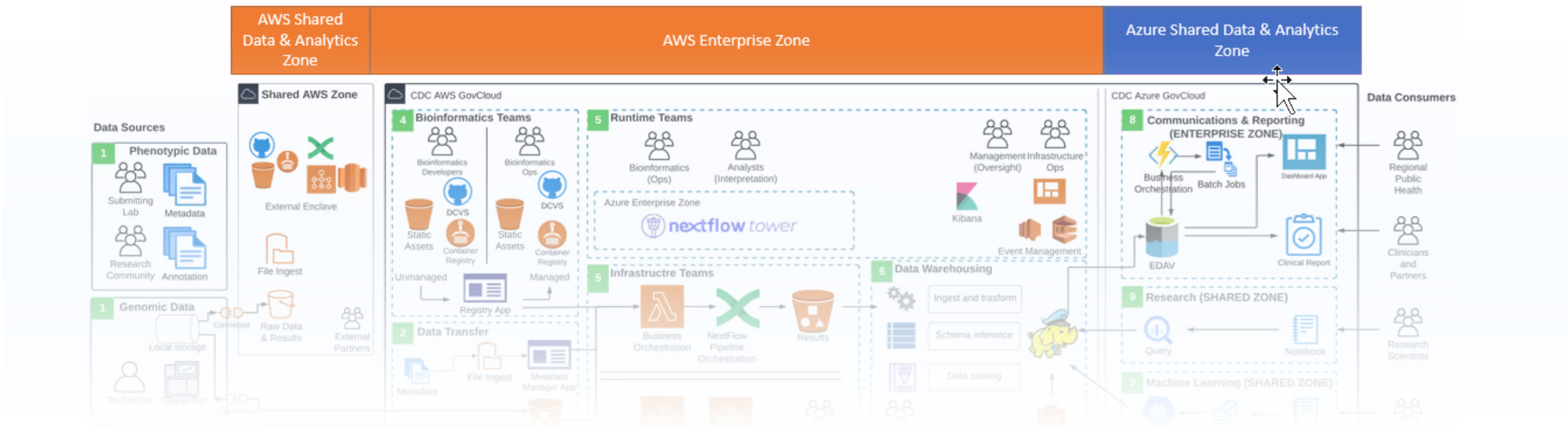
Diagram adapted from [The AI Hierarchy of Needs](#)

Analytics Alignment



Sivarajah, U., Kamal M., et.al., Critical analysis of Big Data challenges and analytical methods, Journal of Business Research, Vol 70, 2017, Pages 263-286 - <http://dx.doi.org/10.1016/j.jbusres.2016.08.001>

Infrastructure needs



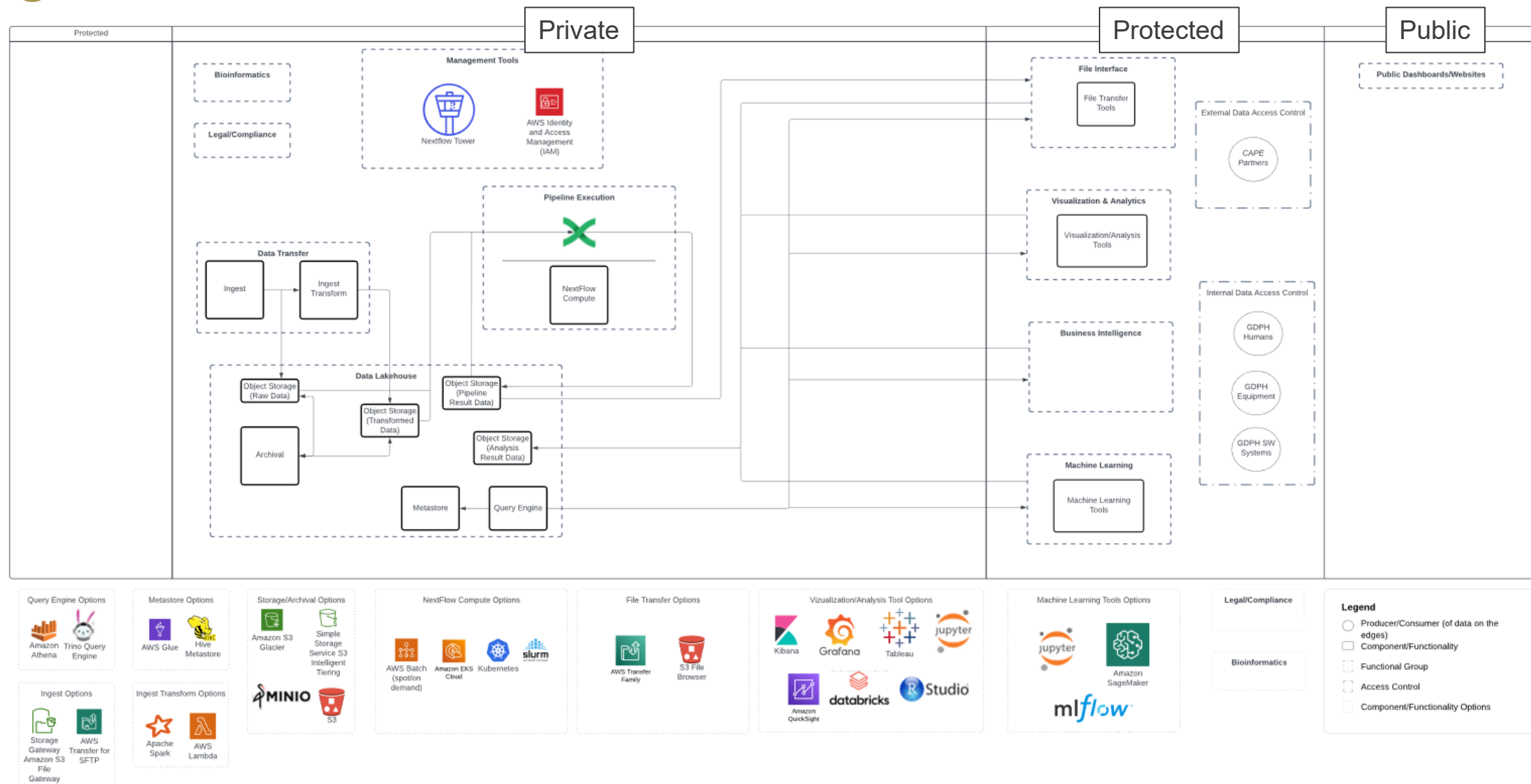
You just want to

- Reproducibly run bioinformatics pipelines
- Provision identical compute architectures across multiple laboratories
- Upload, process, share, and archive data
- Combine data across multiple heterogeneous sources
- Discover unobserved patterns
- Report critical results in near real-time
- Share pipelines and best practices within the community

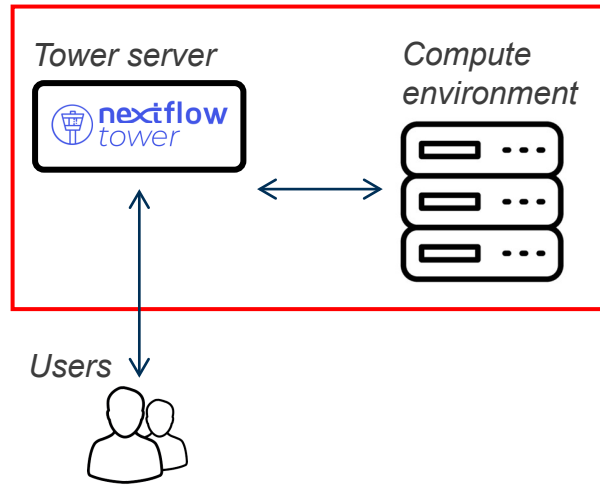
You don't want to

- Fail to deliver results on *time* or at *scale* relevant to decision makers and the US public
- Become a computer scientist
- Manually configure compute infrastructures
- Rely on manual data sharing mechanisms (FTP, email, secure download, etc.)
- Violate HIPAA or other Federal Law
- Develop incompatible “Silos of Excellence”

High-level Infrastructure Functional Architecture

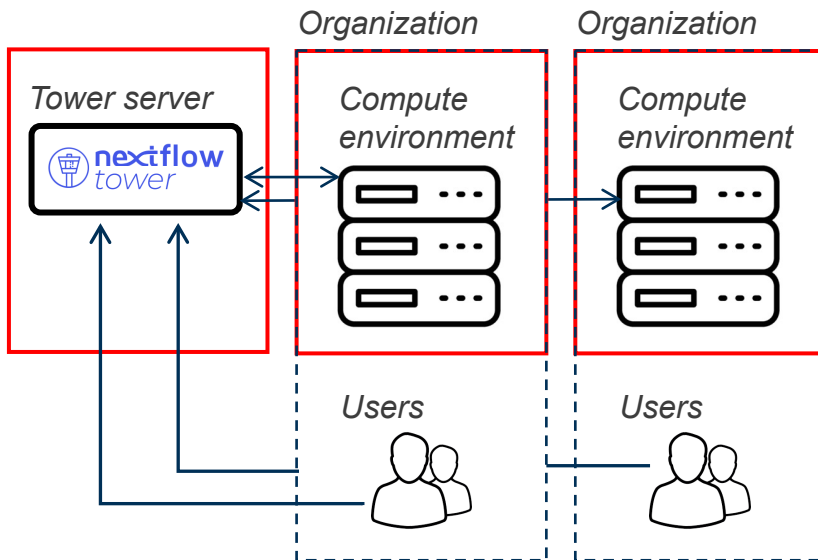


Cost models



Monolithic cloud

- One cloud hosts all infrastructure
- Single party (the owner) responsible for paying cloud provider and tower licenses.
- Single party (the owner) bills participating organization for usage

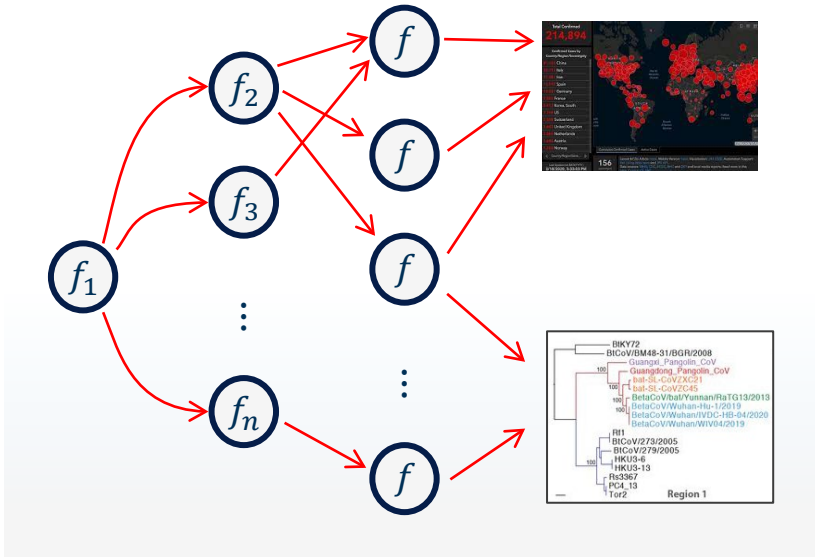


“Bring your own” compute

- Single party (the owner) responsible for paying tower licenses.
- Single nextflow tower instance across CoE allows efficient sharing of pipelines and monitoring best practices
- Each participating organization configures its own compute environment and authorized users
- Compute costs are local to the organization

Analysis pipelines and processes

Goal: reusable units of calculation composed in a computational graph



Process: a node in a pipeline

Pipeline: the computational graph

Desired properties

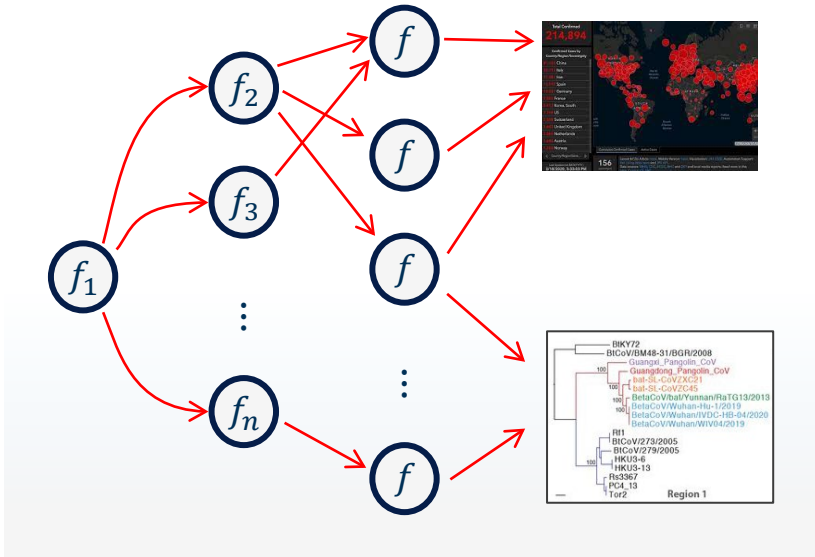
- Processes act as pure functions (no side effects).
- Standardized data IO between processes.
- Documentation standards and independent review.
- Classification metadata for both processes and process data.
- Efficient execution:
 - *Re-entrancy*: process can be paused and restarted
 - *Caching*: reuse a prior result if the inputs have not changed.
 - *Lazy evaluation*: only compute when the result is needed.
 - *Complexity*: efficient scaling (in space and time)
 - *Streaming*: recurrent functions.

$$f(x_n, y_{n-1}) = y_n$$

New data $\nearrow$ $\uparrow$ Prior result $\nwarrow$ Posterior result

Analysis pipelines and processes

Goal: reusable units of calculation composed in a computational graph



Process: a node in a pipeline

Pipeline: the computational graph

Our approach

- Docker containers.
- Functional and unit testing.
- Code formatting standards.
- Autodocs.
- Continuous integration.
- CoE code review.
- Nextflow and nf-core.