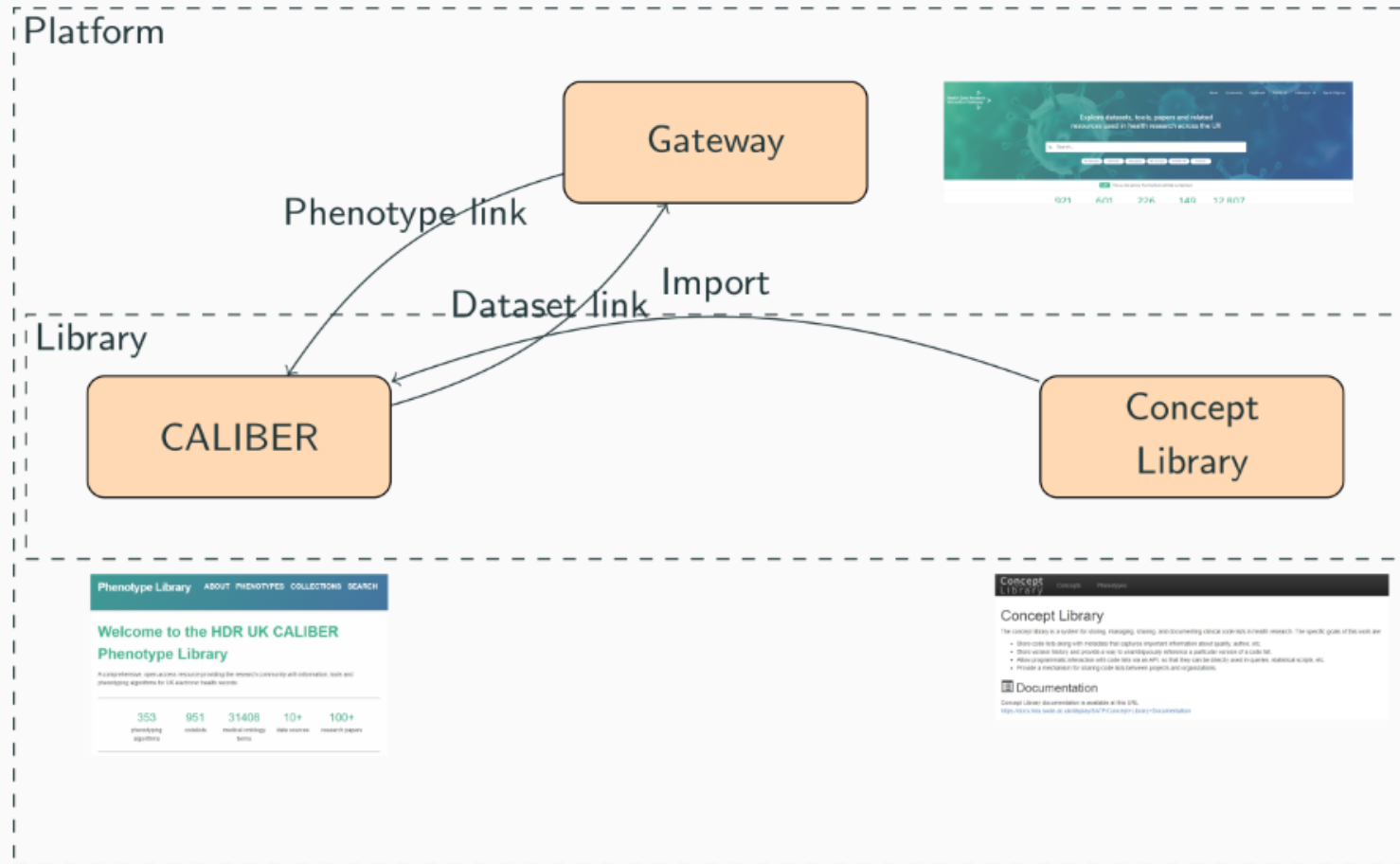


Phenoflow

Martin Chapman and Vasa Curcin

King's College London

HDR Phenotype landscape i



1. CALIBER and the Concept Library – excellent resources for **definitions**.
2. Gateway – excellent resource for **datasets**, which in turn link back to CALIBER (and the Concept Library) for cohort identification.

Q: What if I want to use a definition to identify a cohort in a **local dataset** (i.e. a dataset that isn't available through the Gateway)?

A: Definition **implementation** required.

Potential implementation requirements

1. I'm **not** a developer. I want something I can execute straight away.
2. I **am** a developer, but I want a **structure** in which to work (e.g. develop the definition implementations provided by others).
3. I **am** a developer, but I have an **existing** set of **implementation units** that I want to standardise (for others).


The screenshot displays the Phenoflow web application. At the top, a browser tab is labeled "Phenoflow | Library" with a close button (X). The address bar shows the URL "https://kclhi.org/phenoflow/phenotype/all/". On the left, a teal sidebar contains the "phenoflow" logo and two main navigation options: "BROWSE" with a house icon and "DEFINE" with a plus icon. The main content area, titled "Library", lists two phenotypes: "Abdominal-aortic-aneurysm" with ID "NJ2gf6ZTTxjc" and "Abdominal-Hernia" with ID "UmxTP7n8pMNA9vkfrhc".


Phenoflow | Library

← →

https://kclhi.org/phenoflow/phenotype/all/

phenoflow

 **BROWSE**

 **DEFINE**

Library

Abdominal-aortic-aneurysm NJ2gf6ZTTxjc

Abdominal-Hernia UmxTP7n8pMNA9vkfrhc

The library is backed by a workflow-based, **multi-layer** phenotype model:

- **Abstract** Expresses the logic of a phenotype through a set of simple sequential, potentially nested steps, each of which is annotated with multiple descriptions, in order to tackle complexity.
- **Functional** Specifies the metadata of entities passed between the operations within the abstract layer, e.g., the format of an intermediate cohort.
- **Computational** Defines an environment for the execution of one or more implementation units (e.g. a script, data pipeline module, etc.) for each step in the abstract layer, providing a template for development.

Supported by multi-layer model ii

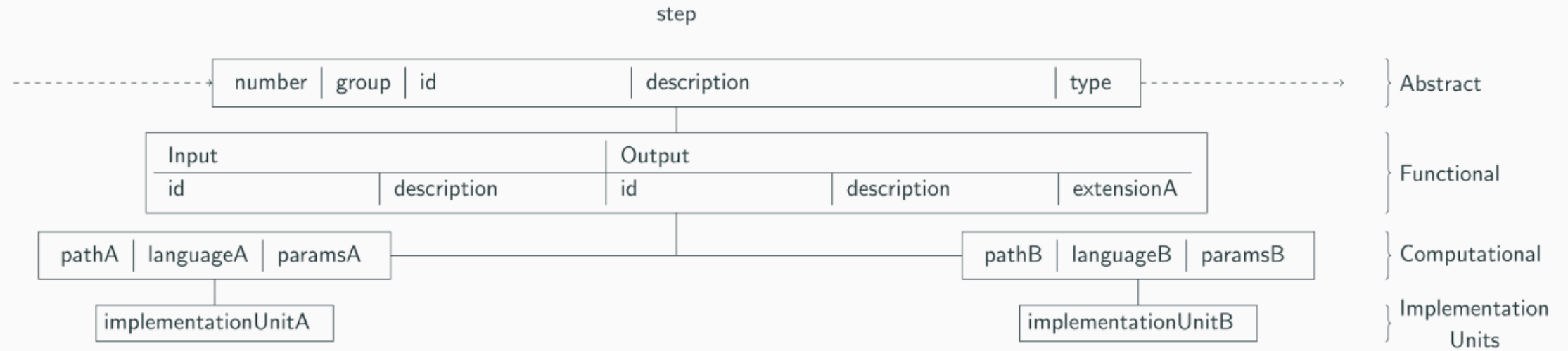


Figure 1: Structured phenotype definition model (step) and implementation units.

Supported by multi-layer model iii

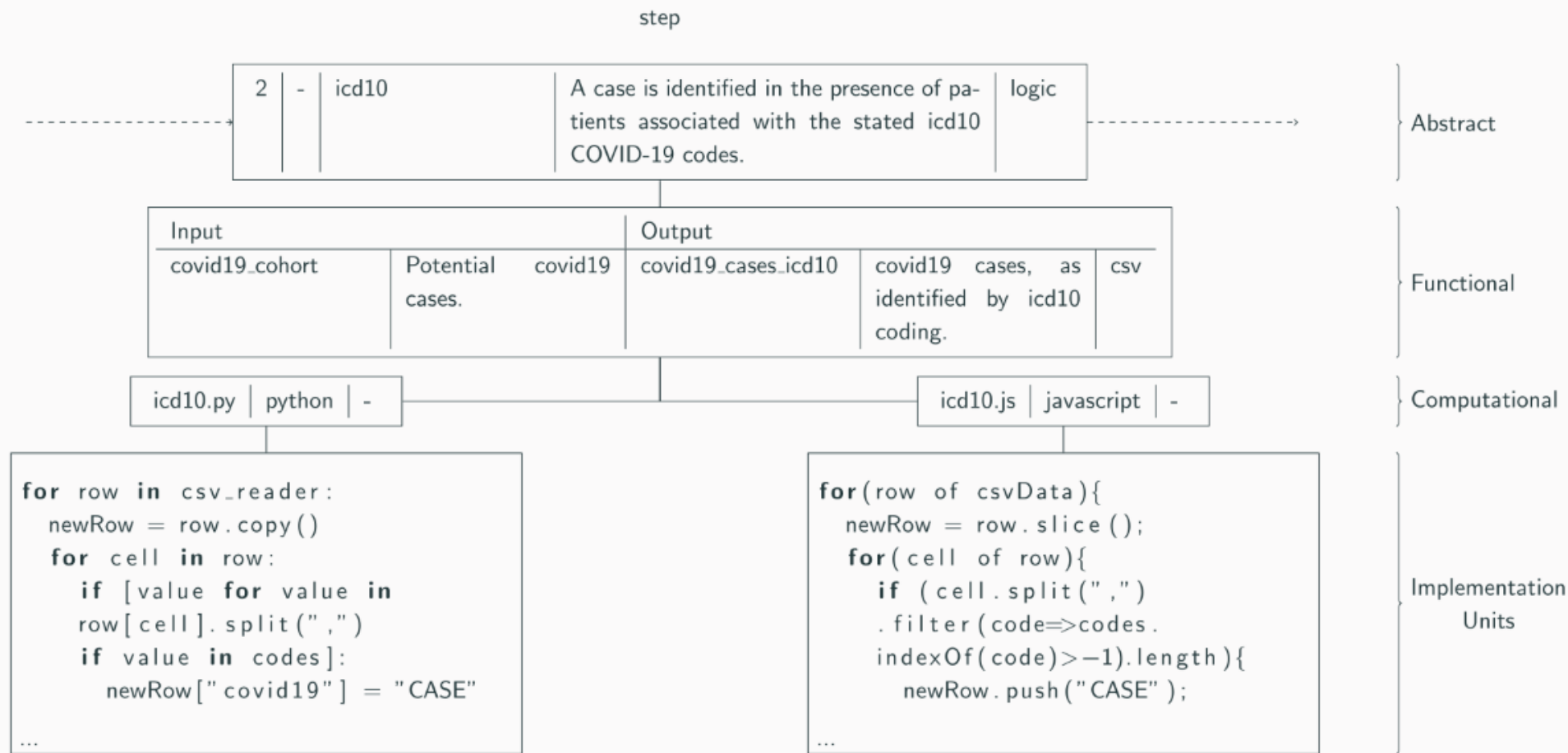
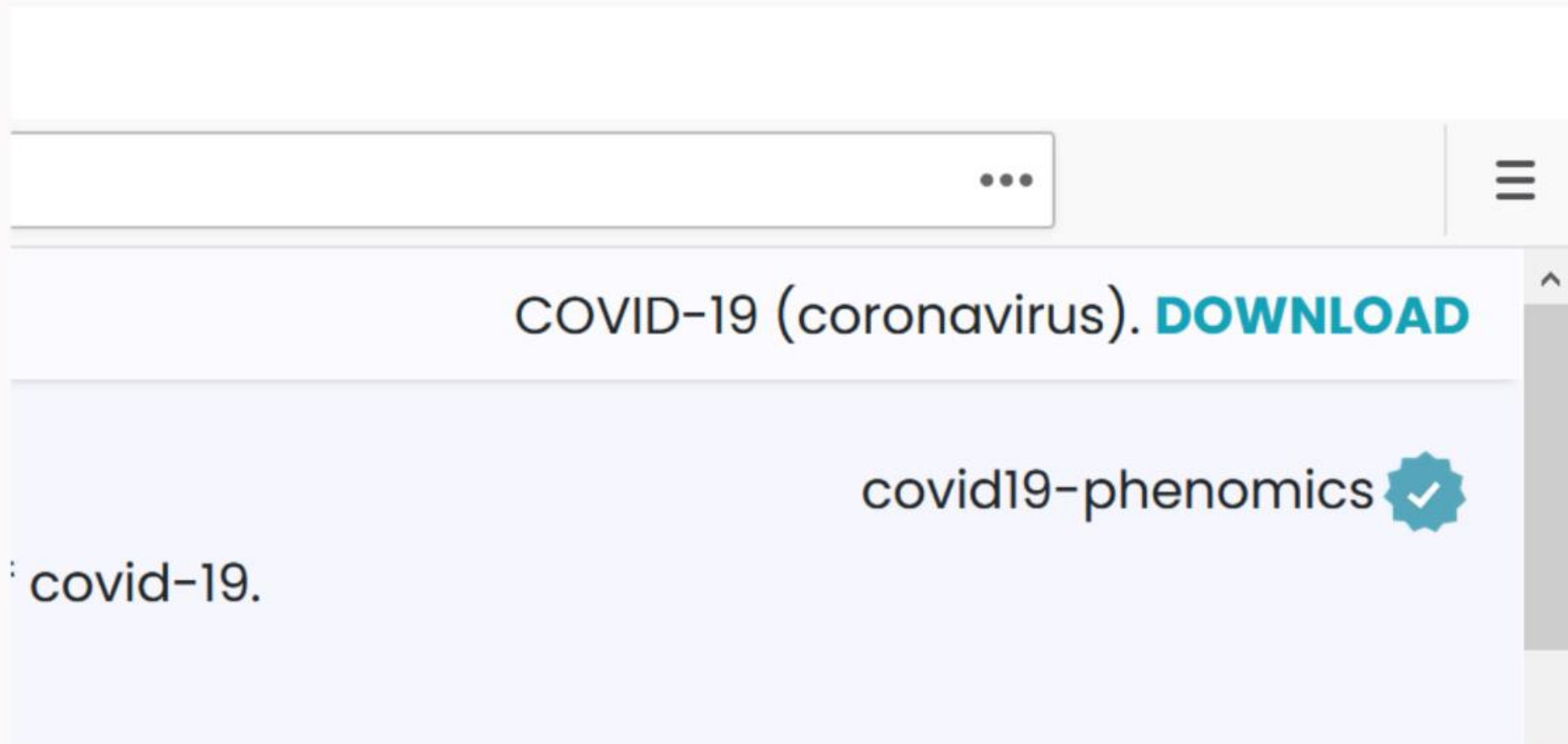


Figure 2: Individual step of COVID-19 structured phenotype definition and new implementation units.

(1) Execute straight away

Download

Download each phenotype as a **CWL workflow** and execute with minimal dependencies.



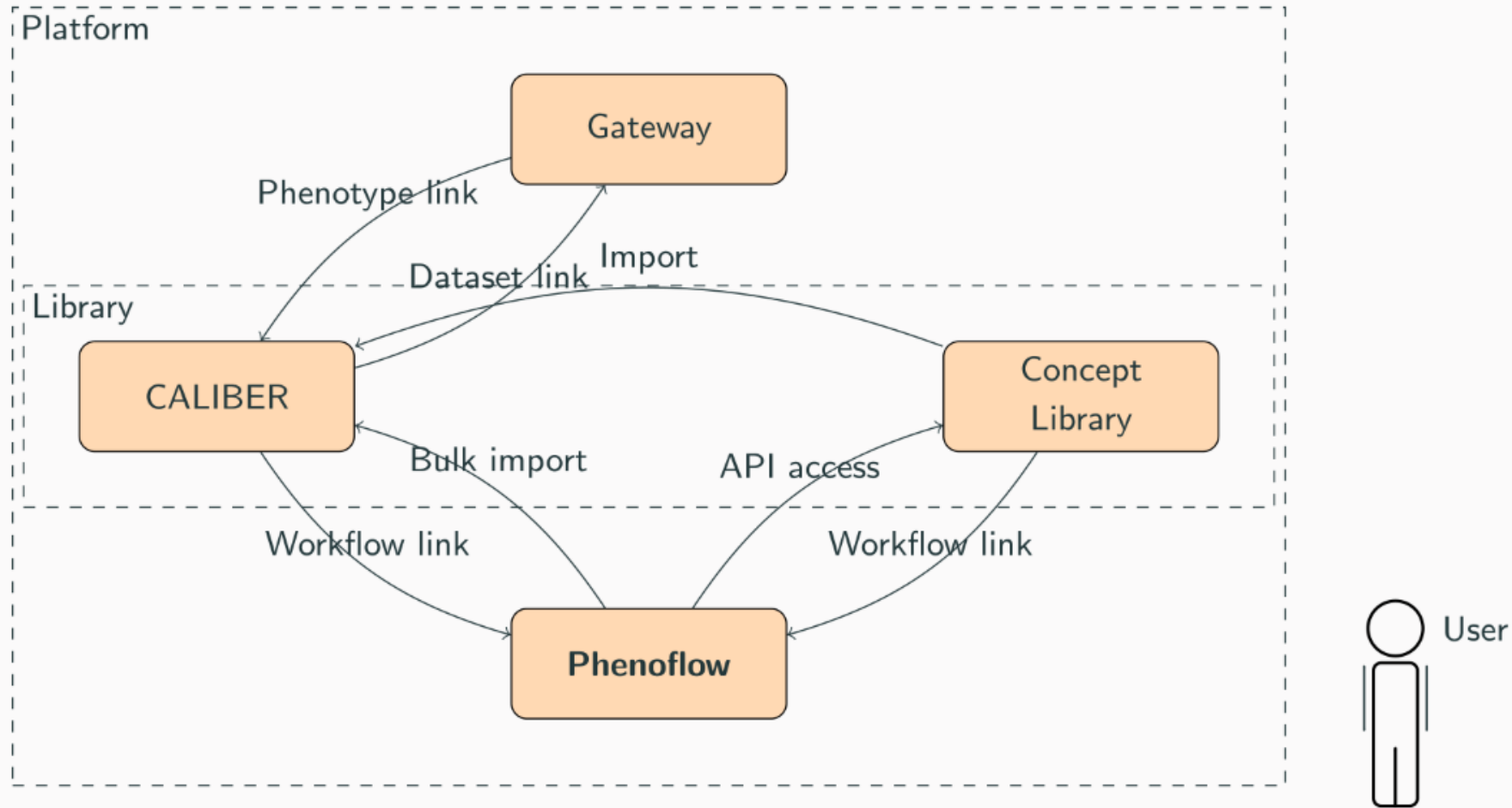
Different data sources?

Select appropriate **data connector** (e.g. i2b2, OMOP, CSV) prior to download.



Available phenotypes

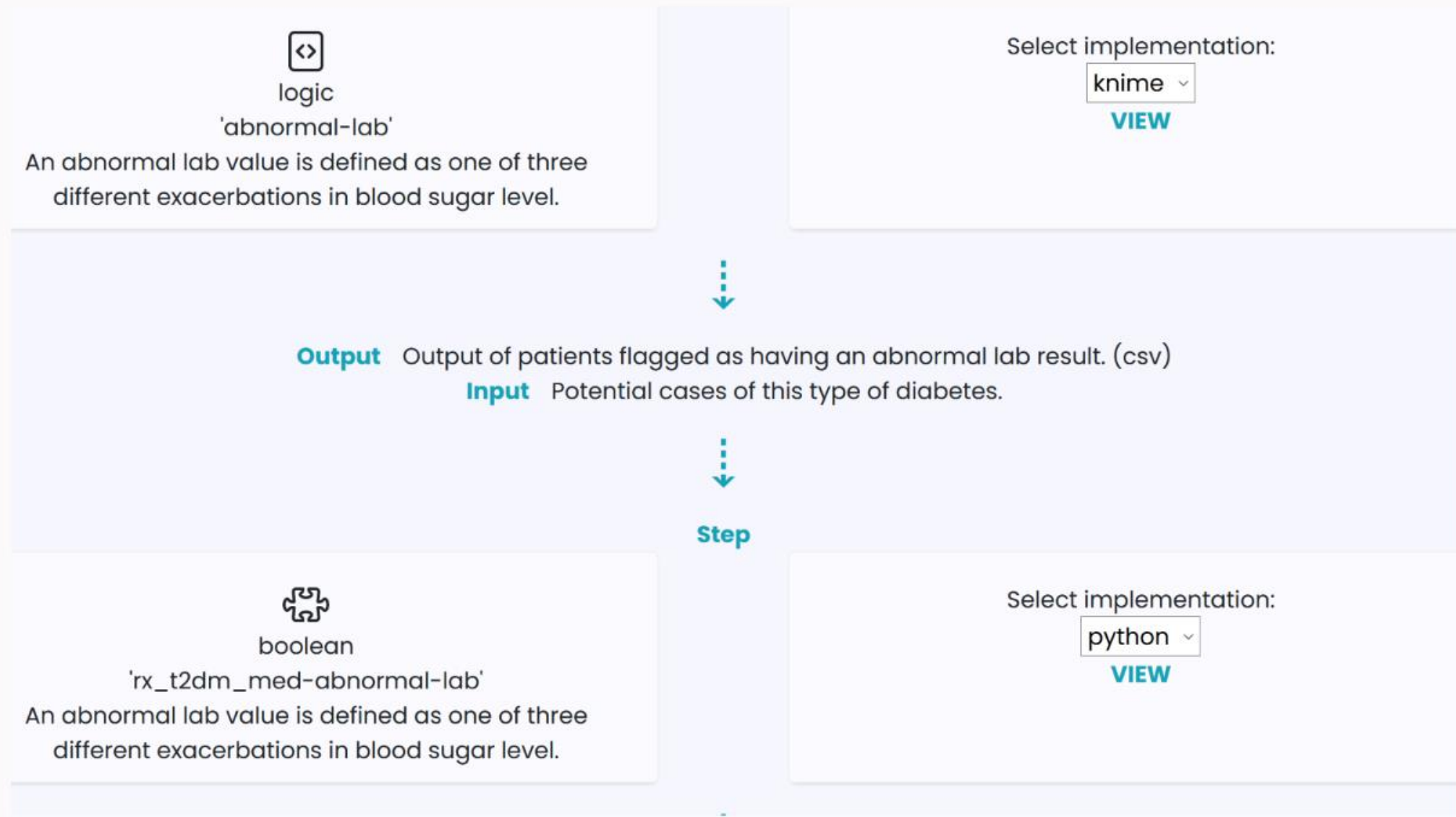
Have **imported**, **standardised** and provided **implementations** for 333 existing definitions:



(2) Structure in which to work

Configure download


Configure a set of **existing implementations** you're comfortable editing prior to download:





(3) Use of local implementations


Multiple implementations


Upload **alternate implementations** for a given abstract step, based on locally developed units:


load


external


logic


boolean


output

Name

rx_t2dm_med-abnormal-lab

Describe

An abnormal lab value is defined as on

rx_t2dm_med-abnormal-lab.knwf

Browse...

No file selected.

knime

rx_t2dm_med-abnormal-lab.py

Browse...

No file selected.

python

ADD

Author **new** definitions:

The screenshot shows the Phenoflow web application interface for defining a new phenotype. The browser address bar shows the URL <https://kclhi.org/phenoflow/phenotype/define>. The interface is divided into several sections:

- Header:** "Phenoflow | Phenotype" and a "X" icon.
- Left Sidebar:** Contains navigation links: "BROWSE", "DEFINE" (highlighted), "EDIT", and "LOGOUT". At the bottom, it features the "HDR UK" logo and "© kclhi 2020".
- Form Fields:**
 - Phenotype:** Two input fields labeled "name" and "about", each containing the text "name" and "about" respectively. A "SAVE" button is to the right.
 - Input:** A single input field labeled "Describe" containing the text "Describe".
 - Step:** A section with five icons: "load", "external", "logic", "boolean", and "output". Below these are two input fields labeled "Name" and "Describe", both containing the text "Name" and "Describe" respectively.
 - Output:** Two input fields labeled "Describe" and "Extension", both containing the text "Describe" and "Extension" respectively. An "ADD" button is at the bottom right.
- Flow:** Vertical dashed arrows indicate a flow from the "Input" section to the "Step" section, and from the "Step" section to the "Output" section.

- Further leveraging the multi-layer model to express **relationships between phenotypes** (e.g. sub-phenotypes) at each layer of the model.
- Publish more implementations for complex disease-specific phenotypes - Covid-19, Stroke.
- Explore the use of Phenoflow as a validation tool.
 - Existing work: AMIA Summits 21 paper with Northwestern [1].

- Increase the library of **workflow modules** (e.g. types of dataset connectors) ready for download and use.
- Automatic **data conversion** to enable use of different implementation techniques on same dataset, e.g. conversion from CSV to DB to allow use of SQL scripts.
- Trace the evolution of phenotype implementations through use of data provenance.
- Further explore the impact of phenotypes on **clinical trials**.
 - Existing work: MIE 21 paper [2].

-  Martin Chapman, Luke Rasmussen, Jennifer Pacheco, and Vasa Curcin.
Phenoflow: A Microservice Architecture for Portable Workflow-based Phenotype Definitions.
AMIA Summits on Translational Science (forthcoming), 2021.
-  Martin Chapman, Jesús Domínguez, Elliot Fairweather, Brendan C. Delaney, and Vasa Curcin.
Using Computable Phenotypes in Point-of-Care Clinical Trial Recruitment.
Medical Informatics Europe (MIE) (under review), 2021.