

Updated ICEES+ YAML file plan, 03.01.2021:

1. Manually create single features.yml file, with all use-case feature variables included and enumerated.
 - a. Conduct systematic manual QC of features.yml file.
2. Manually create a single FHIR_mappings.yml file, with mappings between all ICEES+ API data elements and FHIR elements.
 - a. Conduct systematic manual QC of FHIR_mappings.yml file.
3. Create a single identifiers.yml file, with mappings between all use-case feature variables and relevant identifiers (Translator and otherwise).
 - a. Prioritize identifier systems: ChEBI, ChEMBL, MONDO/HPO, SNOMED?, UMLS?, RxNORM?, LOINC?, ENVO?, MeSH?
 - b. Use SRI normalizer to programmatically look up Translator preferred identifiers for each feature variable and cite that service as the primary source.
 - c. Use Athena to programmatically supplement identifiers, when necessary, and cite that service as the secondary source.
 - d. Conduct a final manual QC step of the identifiers.yml focused on missing variables, missing identifier systems, typos, and a QC check of any identifiers that were manually added.

***See schemas below**

ICEES+ API CONFIG FILES:

<https://github.com/NCATS-Tangerine/icees-api/tree/master/config>

SCHEMA FOR features.yml FILE:

Sex:
 type: string
 enum:
 -Male
 -Female
 -Unknown
 -Other

SCHEMA FOR FHIR_mapping.yml FILE:

feature_variable_name:

 fhir_resource_name:

- system: system
 code: code1
- system: system2
 code: code2

EXAMPLE:

DiabetesDx:

Condition:

- system: <http://hl7.org/fhir/sid/icd-10-cm>
code: O24.42

IDENTIFIER SYSTEMS:

<http://hl7.org/fhir/sid/icd-10-cm>

<http://www.nlm.nih.gov/research/umls/rxnorm>

<http://loinc.org>

<http://terminology.hl7.org/ValueSet/v3-Race>

<http://terminology.hl7.org/ValueSet/v3-Ethnicity>

<http://hl7.org/fhir/2018Sep/valueset-birth-sex.html>

CONCEPT LOOK-UP SERVICE:

<https://athena.ohdsi.org/search-terms/start>

N3C SECURE ENCLAVE FHIR RESOURCES:

<https://atlas.ohdsi.org/#/conceptsets>

IDENTIFIERS:

ChEBI

ChEMBL

MONDO

LOINC

RxNORM

ENVO (will be incomplete)

SNOMED

MeSH (for environmental variables, including socio)

SCHEMA for identifiers.yml FILE:

Table:

Feature:

- identifier
 - biolinkType: PhenotypicFeature (example)

<https://biolink.github.io/biolink-model/>

Example: https://github.com/NCATS-Tangerine/icees-api/blob/cohort_year/config/identifiers.yml

