Researcher

Collaborator

Requester

Downloader

Subject
- dbgap_subject_id: int

Subject Consent
- subject_id: int

Family
- family_id: int

Subject Phenotype

Sample

Sample Attributes

Application/Project

Data Access Request (DAR)

Consent Group

Data Use Limitation Document

Data Access Committee

Study

Study version

+creates and maintains

+granted access to

+authorized to access project by

+member 1..* 0..1

+father 0..1 0..1

+mother 0..1 0..1

0..1 1

1 0..* 0..1 0..*

1..* 1

1 1..* 1 1

0..* 1

1..* 1

+is governed by

+governs access to 1

1..* 1

1..*

Classes/Entities in Blue are those discussed in DGDEV-3579.

Note that this is a logical model and not a physical model. Family is a logical entity implemented in the physical model via pedigree - with some denormalization. The purpose of this model is to capture the logical model that was intended.

The user of the data needs to understand the logical model.

Note that the 'complex' tables have been left off for now. There are study-specific models for these. Many (most?) do not have them at all. It makes more sense to add the study specific logical models rather than abstract this. An example for UDN has been added.

This representation was added firstly from hearing Sue Chen articulate her mental model that there were many data access requests in an application and that each request was for one dataset.

So thought where is the model that Sue described written down.

As I drew this up I realized that there is a disconnect between what Sue referred to as a dataset (a phs) and what dbGaP sometimes calls a dataset (a pht).

In fact the more accurate description is to say that an application is made up of multiple Data Access Requests and each request is for a consent group

Building Bridges!

those discussed in DGDEV-3579.

odel and not a physical model. Family is a logical entity
cal model via pedigree - with some denormalization. The
capture the logical model that was intended.

to understand the logical model.

les have been left off for now. There are study-specific models
o not have them at all. It makes more sense to add the study
er than abstract this. An example for UDN has been added.

tion was added firstly from hearing Sue Chen articulate her
hat there were many data access requests in an
that each request was for one dataset.

re is the model that Sue described written down.

p I realized that there is a disconnect between what Sue
dataset (a phs) and what dbGaP sometimes calls a

e accurate description is to say that an application is made
Data Access Requests and each request is for a consent

s!