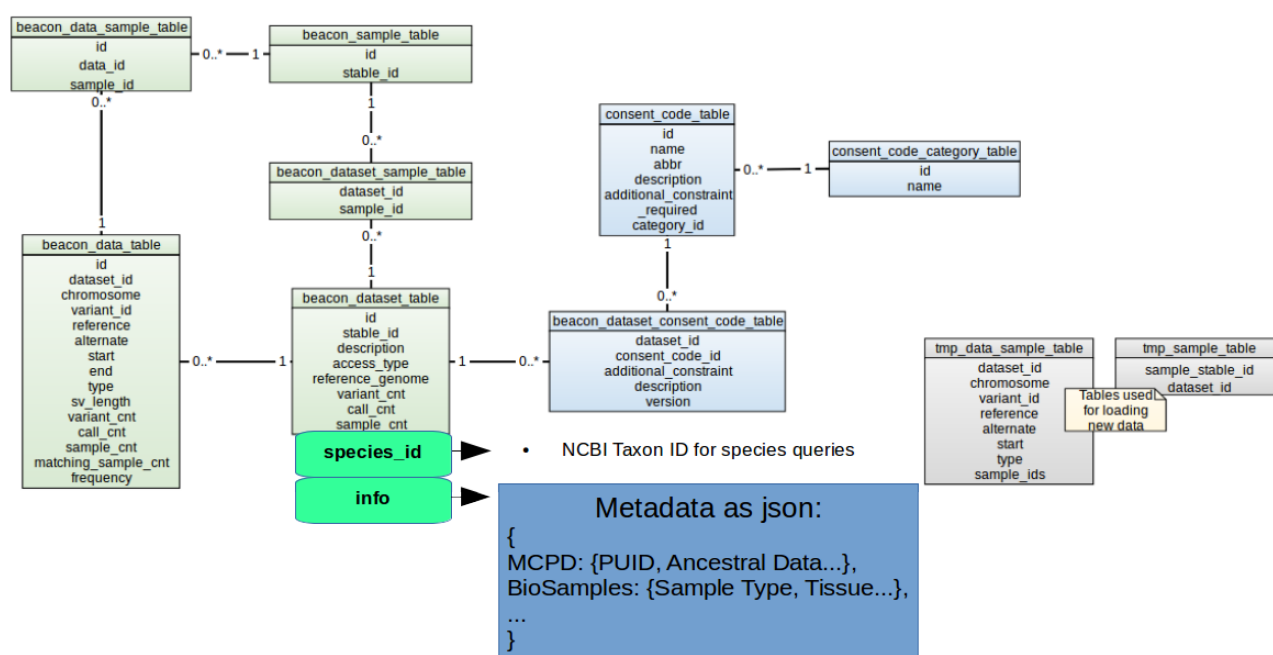# Data schema for the Beacon Plants implementation



To minimize changes to the pre-existing data schema implemented in the ELIXIR Beacon project, we only made two changes, both of them in the *Beacon_dataset_table* SQL table:

- The new **species_id** field is a character field designed to take an NCBI taxon id (for example, "9606" to indicate *homo sapiens Linnaeus*, or "36596" to indicate *Prunus armeniaca L.*). This allows queries to the Beacon database to restrict which datasets they interrogate to only those belonging to a specific taxon id.

- The **info** field was modified to remove any constraints on it's length. This allows it to be used as storage for a json string that can thus contain any metadata associated with the dataset, in the form of indivivdual dictionaries for any particular integrated metadata standard.

For example, for a plant Beacon containing Apricot data and the NCBI BioSamples metadata associated with each dataset:

species_id : "36596"
info : "{BioSamples : { Tissue : 'Leaf', Age : '2 months'} }"