**User-entry fields for each form and relevant info**

(\*here, a form is conceptualized as all the fields for one “submit” - probably corresponds to a node)

Structure of the info in this doc:

**Name of field for user interface:**

PBot: the PBot backend field name+code from the list Claire generated

dwc: the DarwinCore field term name and link to term info. <https://dwc.tdwg.org/terms/>

Info: the definition of the field. In addition to being for our internal use, these will serve as the draft for later definitions used in the guidelines/tutorials for users.

Entry type: describe how users will enter data into the field

Considerations: this is an optional section, if there is info to convey to the PBot team (Botany, Dev, or both) about the field. Notes, wishes, questions, etc.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Operational Taxonomic Unit Entry Form**

**OTU name:**

PBot: name: String (change in backend to "otuName" or OTUName"? Not sure how camel case works for upper-case acronyms!)

dwc: taxonID

Info: The name of the Operational Taxonomic Unit. The name can be Linnean (genus + specific epithet) or any operational name for the taxon that is being used in lieu of Linnean name. The name should be unique.

Entry type: free text. **Required.**

Consideration: how to ensure we don’t have duplicate names? We need to provide guidance on when to make a new OTU vs add specimens (ie occurrences) to an existing OTU.

**OTU authority:**

PBot:

dwc: nameAccordingTo

Info: The authority for the taxon, following ICN rules if a named taxon, or if an informal taxon the authority is the person, or groups of persons, describing the taxon.

Entry type: free text. **Required.**

Consideration: Should it be the short-form or long-form authority names - does it matter? Same for latinized things like and vs et

**References:**

PBot: references: [OTUCitedBy]. Defines a relationship to another node.

dwc:

Info: The reference for the establishment of the OTU. For a taxon that has been described in existing literature, the reference would be the primary citation as well as any citations to published revisions. For un-published taxa, the reference should be to the person or group that has generated the OTU (e.g., an internal reference). **Required.**

Entry type: select from existing references, prompt to create a new reference if one does not already exist.

Consideration: We have ‘reference’ field on ‘description’ node too - this will be duplicate info to some level. But having “reference” on the description nodes allows for references to people specifically adding to the OTU description, for example, by adding a cuticle description.

**Exemplar specimen:**

PBot: holotype: HolotypeOf!. Defines a relationship to another node..

dwc:

Info: A single specimen serving as the representative of the taxon. For a species described in peer-reviewed literature following ICN standards, this is a holotype specimen. For taxa that are not yet formally described (such as morphotypes), then this is an exemplar specimen similar to a holotype.

Entry type: Select a specimen, or prompt to enter a new specimen. **Required.**

Consideration:

**Additional specimens:**

PBot: exampleSpecimens: [ExampleOf!]!. Defines a relationship to another node.

dwc:

Info: All other specimens identified as belonging to the OTU.

Entry type: select from existing specimens, or prompt to enter new specimens.

Consideration: Just as a note - the relationship between specimen and OTU can be created from either the specimen or OTU entry forms. If you are entering a new OTU, that is when you create the exampleOf relationship with specimen(s) that have already been created (note the order of operations here - these specimen(s) would have been created without an exampleOf relationship since the OTU doesn't exist yet!). Conversely, IF the appropriate OTU already exists in the system, you can create the exampleOf relationship from the specimen entry side.

**Parts preserved:**

PBot: organ: Organ @relation(name: "IS\_TYPE", direction: OUT). Defines a relationship to another node.

dwc:

Info: The organs or parts of the plant that are described by the OTU description.

Entry type: multi-select from predefined options. **Required.**

Consideration: formerly called Organ

**Notable features preserved:**

PBot: Defines a relationship to another node..

dwc:

Info: The OTU includes description of these preserved features.

Entry type: multi-select from predefined “notable features preserved” node options:

Consideration: Note: this field is NOT required - the presence of these features can be definitive (if you say they are there then you definitely have observed them), but it is hard to be certain of the absence of the features (i.e., maybe there is cuticle but you are not sure because it would require extra prep work on the specimens, regardless you have not yet observed it so can’t say yes, but can’t say absent either).

**Major taxon group:**

PBot:

dwc:

Info: Choose the major taxonomic group to which the taxon belongs. The selection will allow simplified searches of the database. **Required.**

Entry type: Select from the pre-defined list of taxonomic groups (we are working on finalizing the list of options and will provide them soon)

Consideration:

**PBDB parent taxon:**

PBot:

dwc:

Info: Select the most highly resolved clade that the OTU belongs to from the selection of taxa in the Paleobiology Database (PBDB). This entry allows interoperability with PBDB. Please refer to PBDB if you would like to add taxa to the hierarchy. **Required.**

Entry type: selection from PBDB

Consideration: Will require some information exchange with PBDB. Note that Ellen’s grad student Matt is working on cleaning up and beefing up the taxonomic hierarchy in PBDB from family level up.

**Family:**

PBot: family: String

dwc:

Info: The family to which the OTU belongs.

Entry type: Free text. Not required.

Consideration: This may duplicate info in the Parent Taxon field - but allows a way to enter family if this info isn’t already in PBDB, so that we can capture this info and make changes in PBDB periodically. Alternatively, by having a free-response “clade” field, people could enter the family and everything else there in a comma-separated string. Having the dedicated field means that you would be able to search PBot by “family”.

**Genus:**

PBot: genus: String

dwc:

Info: The genus to which the OTU belongs.

Entry type: Free text. Not required.

Consideration: For named species, this duplicates the “OTU name” field, but enables a dedicated way to search by genus.

**Specific epithet:**

PBot: species: String

dwc:

Info: If applicable, the specific epithet of the OTU.

Entry type: Free text. Not required.

Consideration: For named species, this duplicates the “OTU name” field, seems right to have it though. Question about authority: Do people need to type the authority after the name, or is authority a separate field, and if so do we need one for genus and family too? (I don’t really want extra fields, and hate looking up authorities, but I understand their utility and that some people may want that info - it overlaps with the references for a taxon, if people enter the full reference history for a taxonomic description, but not fully)

**Additional clades:**

PBot:

dwc:

Info: Additional clades to which the taxon belongs, other than already listed in other fields (e.g., tracheophytes, eudicots). Enter multiple clades separated by commas.

Entry type: text entry, multiple entries allowed. Not required.

Consideration: This allows users to enter useful Linnean or non-linnean clade names that are useful, but not captured in the other fields. We are creating this field in lieu of having multiple specific fields for other hierarchy data (for example a sub-genus, or “section” of a family or genus, or broad informal/non-monophyletic group that is still a clade such as ‘conifer’).

**Proposed synonym:**

PBot: synonyms: [Synonym] @relation(name: "SAME\_AS", direction: OUT). Defines a relationship to another node.

dwc:

Info: Proposes a synonymy with another OTU in the database.

Entry type: Select from existing OTUs.

Consideration: This option won’t be on the initial entry of an OTU, right? But after an OTU is created then a synonymy can be proposed, as a button/function from the OTU landing page? We need explanatory materials about when to propose a synonymy vs add notes to an OTU - e.g., how confident should you be??

**OTU quality index:**

PBot:

dwc:

Info: Indicates the author’s confidence level that this OTU represents a valid distinct taxon. For example, how confident are you that this is a real OTU? Are there enough diagnostic characters? Uncertainty in the taxonomic assignment (e.g., what family/genus/etc) should be written into the notes with explanation. [we need to work on this definition to include examples]

Entry type: Select from the following list of options: high, medium, low. **Required.**

Consideration: certainty level of a taxonomic assignment (e.g., what family/genus/etc) should be written into the notes with explanation.

**Diagnosis:**

PBot:

dwc:

Info: The diagnosis for the taxon. Should be a concise statement of the diagnostic features of the taxon that differentiates it from other similar taxa.

Entry type: Free text. **Required.**

Consideration: The diagnosis can include features that are not captured/described by the schema character entries. For example: (1) a published taxon that was described based on leaves and flowers, but we only have a schema currently to accommodate leaf description. The floral parts are still part of the diagnosis of the taxon. Hopefully a flower schema will eventually be developed that could accommodate the remainder of the taxon description, but we can’t hold up the important inclusion of leaf data because of that (note the parts preserved field for this OTU would include leaves and flowers). (2) a leaf described by Brown pre-MLA that has some descriptive features in its diagnosis that are not accommodated by the MLA, but for which the user is not equipped to create a new character for (maybe it’s a weird unique thing and there is not comparative study yet to know what the other states of that character would/whould be). That unique diagnostic feature can be included in the text diagnosis, but would not be in the MLA description.

**[Merged] description:**

PBot: mergedDescription: [PseudoCharacterInstance!]! Defines a relationship to another node(s)?

dwc:

Info:

Entry type: takes user to the form/interface for entering descriptive data.

Consideration: I am not sure how this works vs the "description" field that is on the specimen node. This is probably not a part of the initial creation of an OTU node, but once you enter an OTU then you can describe it. Also, because of that order of operations, the description cannot be required. In some cases this may make sense, even though it is not preferable. For example, I am working on a new flora and have to enter an OTU that is an uncommon leaf form (for example, a czechanowskia, or a bryophtyte) that does not yet have a schema appropriate for describing it - a user can enter the taxon in their private workbench using just a diagnosis. The hope would be that a schema is at some point, but this prevents a total roadblock on the user for working on their flora. We may need to make a decision on whether this is only ok in a private workbench, but to go public the OTU needs a character-based description, at some level (e.g., fill out the few possible characters of the MLA or other schema that can be applied, like leaf shape).

**Notes:**

PBot: Not yet defined

dwc: similar to taxonRemarks or

Info: Users can enter “Notes” on the specimen (formerly we called these “comments”).

Entry type: free text.

Considerations: need to figure out comment entry. This is the forum that users can enter notes (ie comments) on the taxon. Can be entered by the creator of the OTU during the initial OTU form submission, and added onto later by any user to make notes about the taxon (preferably via a button on the OTU landing page, as in the mock-up).

**Data access groups:**

PBot: elementOf: [Group!]! @relation(name: "ELEMENT\_OF", direction: OUT). Defines a relationship to another node.

dwc: “groups” are sorta similar to datasetID <http://rs.tdwg.org/dwc/terms/datasetID>, or datasetName <http://rs.tdwg.org/dwc/terms/datasetName>

Info: Defines what project groups have access to view the record. Published OTUs should always be part of the ‘public’ access group. OTUs can belong to multiple groups. If an OTU is a part of a designated project group, that OTU will display in the taxon list for the project on the users workbench. **Required.**

Entry type: multi-select from the users project groups and the option Public

Considerations: can be public AND in another defined group

**Entered by:**

PBot: enteredBy: [OTUEnteredBy]

dwc:

Info:

Entry type: auto-generated.

Consideration: This is display on the taxon history on the OTU landing page, but is not something users interact with.

Currently in Dev database:

name: String

family: String

genus: String

species: String

exampleSpecimens: [ExampleOf!]!

holotype: HolotypeOf!

synonyms: [Synonym] @relation(name: "SAME\_AS", direction: OUT)

references: [OTUCitedBy]

mergedDescription: [PseudoCharacterInstance!]!

elementOf: [Group!]! @relation(name: "ELEMENT\_OF", direction: OUT)

enteredBy: [OTUEnteredBy]