# Specify 6 Model for Institution Organization

Specify Software Project Staff
25 February 2009

During our requirements analysis for Specify 6, we have observed many different organizational models for collections institutions. As a consequence, we designed Specify 6 to handle various models for the administrative organization of biological repositories. This is a description of the major components of Specify 6's model for institution organization.

1. Each Specify database installation (including the underlying database management system), whether it contains records from a single collection of specimens representing one taxonomic discipline or several collections representing different taxonomic groups, must have one **Institution**. For university collections, the **Institution** in a Specify sense would typically be the museum, herbarium or research center which operates the collection itself, not the College or University or higher level group.

2. Each Specify **Institution** has one or more of each of these three units: **Division, Discipline** and **Collection.** These administrative units are represented as distinct database tables in Specify, with their own descriptors, and they are related to each other in that order, with one-to-many relationships. Namely, an **Institution** must have one or more **Divisions** which must have one or more **Disciplines**, which in turn has one or more **Collections**.

3. Each **Institution** in Specify must have one or more **Divisions**. For example, at the Sam Noble Museum of Oklahoma Natural History, there are administrative Divisions of Biological, Geological and Social Sciences. Or museums may divide collections into groups of Vertebrate Zoology and Invertebrate Zoology. If an institution does not formally recognize Division (or Department, Section, etc.) as an administrative unit, Specify requires that one nominally be named. If it is not a real, existing unit, the name will have no effect on using the software. But if Divisions exist, it is very important to understand the how the identification of them in Specify will effect access to certain data types.

4. Each **Division** (or whatever the administrative unit below Institution may be called) must a have one or more **Disciplines.** Commonly there will be one discipline for each division, but some institutions have multiple disciplines per division. Disciplines in Specify correspond to high order taxonomic groups: ichthyology, herpetology, non-vascular plants, entomology, paleobotany, etc. Records for some data types in Specify, e.g. Agents (or people) are shared (i.e. limited or 'scoped') within a Discipline but not among them. Data entry form layouts are also often customized at the Discipline level.

5. A **Discipline** must have one or more **Collections** in Specify. **Collections**, such as the Fish Main (Alcohol) Collection, Fish Tissue Collection, Teaching Collection, etc. For institutional catalogs, a Specify **Collection** represents the physical collection of biological artifacts held and typically owned by that institution. Specify 6 also support collections-based research databases which contain specimen records from multiple institutional sources. Specify 6 also handles species occurrence data from observations without museum vouchers.

6. **Collections** in Specify 6 contain **Collection Objects** which in turn can be categorized into **Preparations and Preparation Types**. (In Specify 6, Collections correspond to Specify 5's **Series**. **Preparations** replace **Physical Collection Objects** in Specify 5. Preparations can be loaned, gifted and managed within transactions as discrete collection artifacts.

7. Specify 6 is designed to store all of the information for every Collection at an Institution within a single (MySQL, SQL Server, etc.) database installation. All data tables for all Divisions, Disciplines and Collections will be typically managed within one database installation at an Institution. When data from multiple collections are managed within a single Specify installation, data records from all collections are interspersed in the same data tables. With a few intentional exceptions, records which belong to a particular Collection will only be seen by the users of that Collection database. Alternatively, it is possible but usually not preferred, to have multiple installations of Specify each with their own database repository (DBMS) within a single institution. If an Institution's collection data are accessed from a shared server and there is good network connectivity, there would be little reason to create a separate Specify 6 installation for each collection. Backup and restore operations are performed on all data in all collections within an installation. There is no capability in Specify 6 to backup or restore a subset of collections databases when they are managed within a single installation. Specify 6.0 does not support queries across Specify collections, from within Specify, but a future release will. We designed Specify to support that capability.

8. We use the term "database" with Specify, in a broad and in a narrow sense. The broader sense is the one described above, to indicate the complete contents of a Specify database (and database management system) installation which comprise all of collection records at an Institution. The narrow sense is what a user perceives–the database of his/her collection's data. When we use the term, we are usually referring to the narrower definition, in Specify's case it is the collection-specific, filtered view of the entire institutional repository which a curator or collection manager would recognize as 'my database.'

9. In the Specify 6 model, there is usually, but not always, a one-to-one correspondence between a physical collection and a database (in the narrow sense of the term). For example, a **Discipline** like Ichthyology may have a main collection of fish

stored in ethanol in bottles and have a second collection of fish tissues in freezers. At many institutions, the two sets of specimens would be considered distinct collections, and their data would be managed in Specify 6 as two separate Specify **Collections under** a single **Discipline**. However, if tissues were managed by Ichthyology as just another specimen **Preparation Type**, then the data from both kinds of **Preparations** can also be easily managed together as a single Ichthyology Collection. Some museums and herbaria curate tissues (and other frozen collections like DNA) as a distinct administrative and collection units which cut across **Divisions** and **Disciplines** within the institution and prefer to manage all data from tissues (or fruits, or seeds, or pollen, etc.) in a separate database. How an institution chooses to delimit its individual collections depends a lot on historical management practice, but the decision also has consequences for whether various types of data records are shared among collections within Specify for a particular collection configuration. For example, if fish tissues are managed as a distinct Collection but within the Discipline of Ichthyology, then tissue and alcoholic specimen data records can share for example, Taxon Tree, locality data, collecting event, collecting trip information.

10. In Specify, if a **Division** combines catalog data from two or more **Collections** into a single database (e.g. Main Fish and Fish Tissues) queries, reporting and statistics could still be grouped on the basis of **Preparation Type**, as if there were two distinct administrative collections. In Specify 5, it was possible to combine data from multiple Disciplines into a single Specify Collection using the Specify 5 **Catalog Series** concept, permitting a single set of catalog numbers to be used across the **Institution**. In Specify 6, data from different taxonomic Disciplines must be managed in distinct Specify 6 Collections, but Specify 6 can be configured at installation time to share catalog numbers across all collections within a Discipline, Division or Institution. Specify's **Collection Object** storage tree is also shared or "scoped" across the entire Institution but linked collection objects are only visible to the users with accounts for that particular collection. Accessions data are optionally configurable to be shared across collections.

| Specify Data Table | Example #1 Univ. of Kansas | Example #2 Univ. of Kansas | Example #3 Univ. of Michigan | Example #4 Univ. of Oklahoma | Example Data Table Sample Fields in Specify6 |
|---|---|---|---|---|---|
| **Institution** | Biodiversity Institute | Biodiversity Institute | University of Michigan Herbarium | Sam Noble Oklahoma Museum of Natural History | · Name<br>· Abbreviation<br>· IPR (statement)<br>· Links one-to-many to Division |
| **Division** | Ichthyology | Herbarium | Herbarium | Biological Collections (SNOMNH also has Geological Collections) | · Abbreviation<br>· Name<br>· Icon URI<br>· Description<br>· Links one-to-many to Discipline |
| **Discipline** | Ichthyology | Botany | Botany | Paleontology | · Name<br>· Links to many tables<br>· Links one-to-many to Collection |
| **Collection** | · Main<br>· Tissue | · Vascular Plants<br>· Lichens | · Vasc Plants: Main<br>· Vasc Plants: Teaching<br>· Vasc Plants: Wood<br>· Non-Vasc Plants | · Paleobotany<br>· Invertebrate Paleontology<br>· Vertebrate Paleontology | · Description<br>· Collection name<br>· Primary Focus<br>· Scope |

| Database (not a Sp6 table) | · Main · Tissue or · Combined with Main and Tissue within one database) | · Vascular Plants · Lichens · Fungi | · Vasc: Main · Vasc: Teaching · Vasc: Wood · Non-Vascular · Fungi (to be verified) | · Paleobotany · Invert Paleo · Vert Paleo | Not a Specify 6 table, but database can defined in parallel for each collection, or what are essentially preparation types can be combined into a single database, for example if a collection maintains its own tissues. |
| --- | --- | --- | --- | --- | --- |

Table: Specify 6 hierarchical concepts for administrative collection units in biological museums and herbaria, with four examples (draft).