# Workflow Detail: Quality Assurance of Data (Pinned Things)

| Module 0 Generic tasks | Module 1 Pre-digitization curation | Module 2A Specimen imaging | Module 2B Whole-drawer imaging | Module 2C Label imaging |
|---|---|---|---|---|
| Module 3 Image processing | Module 4A Data capture from image | Module 4B Data capture from specimen | Module 4 Event data capture | **Module 5 Quality assurance** |
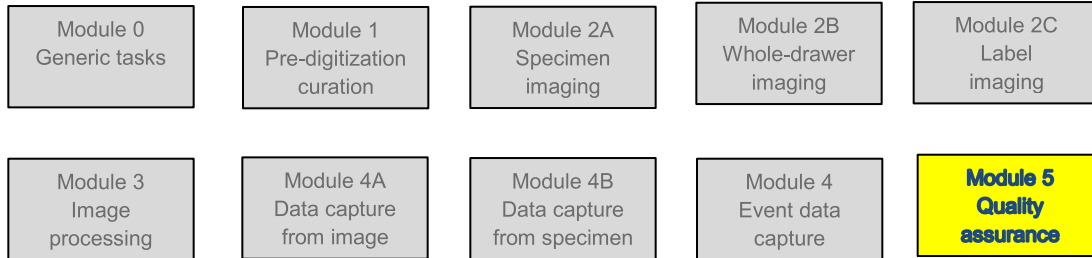
## Module 5: Quality Assurance of Data

| Task ID | Task Name | Explanations and Comments | Resources |
|---|---|---|---|
| **T1** | Expert or manager reviews and approves keyed data. | Since resources are seldom available for a manager to review all entered records, an informal or formal sampling scheme is usually required. Sampling schemes are often in the form of: initially review 100 percent of records by a new data capture technician, then reduce to review to 50%. If quality remains high, reduce sampling frequency to as low as about 10%. If problems are found for a given technician, increase sampling frequency as necessary. | |
| **T2** | Use automated scripts to check against authority files for existence of misspellings of taxon, geographical, or collector names, and for orthographic and data type inconsistencies. | | |
| **T3** | Resolve problems identified in T2. | This may involve merging duplicated collecting events or localities, correcting or | |

FLORIDA MUSEUM OF NATURAL HISTORY    THE FLORIDA STATE UNIVERSITY    NSF    UF | UNIVERSITY of FLORIDA

| | | | |
|---|---|---|---|
| | | deleting incorrect data, etc. | |
| **T4** | Confirm that records have been checked, when they were checked, and with what results. | | |
| **T5** | Compare data with that of likely duplicates of the same occurrence or collecting event. | | |
| **T6** | Cluster data and look for outliers. | | |
| **T7** | Compare data with related data. | For example, compare taxon name and georeference with known or predicted geographic distribution or dates of collection. | |
| **T8** | Expose data to community for expert feedback. | | e.g., through Email, FilteredPush (http://wiki.filteredpush.org), institutional website, etc. |