

Hollister Herbarium Digital Data Policy

Specimen data quality access preservation and reuse policy

Administrative unit	Department of Biology, College of Arts and Sciences
Policy owner	Director and Curator, Hollister Herbarium
Approval authority	Department of Biology and Tennessee Tech
Effective date	June 1, 2026
Review cycle	Every three years or sooner if requirements change

Purpose and scope

This policy governs digital records associated with HTTU specimens, including catalog data, label transcriptions, georeferences, taxonomic determinations, images, annotations, collector and identifier attributions, loan and sampling records, and data shared through SERNEC, GBIF, or other services. It supports open scientific use while protecting sensitive information and maintaining a trustworthy record of the physical collection.

Authoritative records and identifiers

- Each accessioned specimen receives a stable HTTU catalog number. The number is never reassigned, even if the specimen is transferred, lost, or deaccessioned.
- The collection database is the operational record for specimen data. The physical label and annotations remain primary evidence and are not overwritten to match an interpretation.
- Verbatim label fields are separated from standardized or interpreted fields whenever practical. Corrections preserve provenance and the date or agent responsible.
- Published records should include persistent dataset identifiers and links when available.

Data standards and quality

HTTU uses Darwin Core terms and community conventions supported by SERNEC and GBIF. Data entry and review should prioritize catalog number, scientific name, collector, collection number, event date, locality, county, state or province, country, habitat, coordinates, georeference metadata, identification history, and image linkage when present.

Quality control	Expected practice
Transcription	Preserve label wording in verbatim fields; standardize capitalization, names, dates, and geography in interpreted fields.
Taxonomy	Retain original and subsequent determinations; record the accepted display name without erasing identification history.

Quality control	Expected practice
Georeferencing	Record coordinates, datum, uncertainty, protocol, sources, and georeferencer. Do not imply precision greater than the label supports.
Validation	Check impossible dates, coordinates, country-state mismatches, duplicate catalog numbers, missing identifiers, and controlled-vocabulary values.
Versioning	Document substantial batch edits, imports, exports, mappings, and corrections so changes can be reconstructed.

Access and licensing

HTTU intends specimen occurrence data to be openly discoverable and reusable through SERNEC and GBIF. The current GBIF dataset uses CC0. Before changing a license or publishing a new data class, HTTU will confirm authority, donor or permit restrictions, third-party rights, and Tennessee Tech requirements. Images may carry separate rights or attribution terms from occurrence data.

Sensitive locality data

Exact localities may be withheld, generalized, or delayed when disclosure could increase risk to rare plants, private landowners, culturally sensitive resources, active research, or protected sites. The public record should indicate that information has been generalized when doing so does not create additional risk. Full data may be considered for qualified users under a documented agreement and agency or landowner restrictions.

Personal and restricted information

- Do not publish private contact information, student records, confidential donor terms, access credentials, or information unrelated to the scientific record.
- Collector and identifier names are part of specimen provenance, but identity reconciliation must be evidence based. Do not infer a unique person from initials alone.
- Loan, visitor, permit, and sampling administration records may require access controls even when the specimen occurrence record is public.

Images and derived data

Images should include the complete specimen and labels at a useful resolution, a color or measurement reference when available, and a reliable connection to the catalog record. Cropped details, OCR, transcriptions, georeferences, and machine-generated values must be identified as derived information and reviewed before being treated as authoritative.

Preservation security and continuity

- Use university-supported storage and access controls for working files, credentials, administrative records, and backups.
- Maintain at least one managed backup independent of the primary working copy and test recovery periodically.

- Limit write permissions according to role; use individual accounts rather than shared credentials when systems allow.
- Export the complete collection database and essential media inventory on a documented schedule and after major batch changes.
- Report suspected data loss, unauthorized access, malware, or credential compromise through Tennessee Tech information-security procedures.

Corrections citation and reuse

HTTU welcomes documented corrections and will evaluate them against labels, images, annotations, and authoritative sources. Users should cite the HTTU dataset DOI or portal citation and identify specimens as HTTU plus catalog number. Publications, annotations, sequences, and other derived results should be returned to HTTU or linked to the specimen record when practical.