Data

- Specimen data
- Determination history
- Imaging specimens
  - Image archive (TIF, RAW, DNG)
  - Web images (JPG)
  - Skeletal data
- Other
  - Genetic links, voucher info, references, general annotations, comments, etc.
Data Management

- Specify, KE-Emu, Brahms
- Symbiota, Arctos, CollectionSpace
- Best is relative

Needs:
- Easy to use
- Affordable
- Appropriate IT support
- Adherence to standards
- Publishing, data extraction
- Protocols for quality control and cleaning
Locally installed software

Full service

IT support needed
  ◦ Install and update
  ◦ Server
  ◦ MySQL database server
  ◦ Backups
  ◦ Image support
KE–EMu

- Comprehensive museum management
- Costs money
  - Initial data loading, training, license
- Some IT support needed
  - Server, images, software install and updates
- Data publication support?
NANSH (Symbiota Portal)

- Publishing with management options
- Data stored in cloud
- Some services not available
  - E.g. custom invoices
- Live data / Online presences
- Distributed support
  - Backups automatic
  - Image support
  - No software updates
Other Management Systems

- **Arctos**
  - Shared cost
  - Browser based system

- **Brahms**
  - Software purchase
  - Local install and updates

- **CollectionSpace**
Custom Database

- Local IT support needed
  - Future?
- Image management?
- Regular local backup system?
- Standard compliance?
- Data publishing?
- GUID support?
- Avoid Excel “databases”
Data Augmentation & Cleaning

- Authority lists
  - Taxonomic Names
    - TROPICOS, ITIS, Catalog of Life, EOL
  - Geography
  - Collectors

- Georeferencing
  - GeoLocate

- Get data online!
  - Use the data
  - Let others contribute
  - Local experts
Data Publishing

- Darwin Core Archive
- IPT – Integrated Publishing Toolkit
- Data aggregators
  - iDigBio
  - GBIF
  - BISON
  - Symbiota
  - EOL
- Why publish?
- Why are there so many options
Data Exchange Standards

- Darwin Core
- Darwin Core Extensions
- Darwin Core Archive
- Audubon Core
- Global Unique Identifiers (GUIDs)
- Enables data sharing, integration, and use
- Who should care?