# CHALLENGES IN DATABASE SELECTION AND DESIGN

Laura Brenskelle

**Graduate Student** 

Jackson School of Geosciences, The University of Texas at Austin

#### Introduction

 Evolution of data management – written catalogs, punch cards, databases...



- Internal use → sharing data with the public & researchers
- Collections managers in the 21<sup>st</sup> century not only need to know about the discipline they work in, but they are also essentially information managers.

## **Understanding the Trajectory**

- How do you choose a database management system
- How do you customize your choice to match your needs

## Research Questions - #1

- To what degree do natural history collections staff whom interact with a database understand its structure and how to effectively manage it?
  - Informational content is accurate.
  - Database is functional.
  - Could be exported to another platform with relative ease.

## Research Questions - #2

 How do the structures of different natural history collections database schemas vary, and how well do these schemas match the information of the science they are mapping?

#### Specify Database Schema Stodyrochy K, FK1 Stodyrochy Geologic Trine Siper Circup Oxop Formation Bed Porcarks Text1 Text2 Number TimesomyCles Underfore Marchary TimesomyCles Underfore TimesomyCles Underfore TimesomyCles FRASE Proceedings of the Control of PACY | Second Pacy | Pa 4 Gascapping Control of the Control Geography Geography Geography CurrentO ContinentOrOcean ContinentOrOcean ContinentOrOcean ContinentOrOcean ContinentOrOcean ContinentOrOcean State ContinentOrOcean State ContinentOrOcean State ContinentOrOcean Continent Contin Cofection CofectionS CollectionName CollectionName CollectionName CollectionName CollectionName CollectionName CollectionName Remarks Teet Teet Namber Nam Address D Address D Ony State Country Postatode Remarks Timestamph Timestamph Lastif cladili International Control of the Control PK ImageAget PK2,U1,I2 ImageED PK1,U1,I1 ImageED Revision Timestampl Timestampl Timestampl Timestampl Legend Loan PKV LeastD PKVUI Spermett Ust Collection Ust Conduction Use C Objects & Things Taxa & Determinations People & Agents PK Accession PKU1,UJ Accession UJA4 UJA4 Observation Property Control Actions & Transactions Closervation Qbancation Use Relagication Court Description Activity Renaria Text2 Number2 Timesampilia Time Time & Places Specify Software Project: http://specifysoftware.org Supported by U.S. National Science Foundation, Award BIO/BDI 0446544 and earlier grants.

## Research Questions - #3

 Because databases play a key role in organizing and maintaining information within a collection, how will natural history collections databases and their schemas change to accommodate the emergence of various forms of derivative analytical data? How well are the caretakers of collections preparing for these changes?

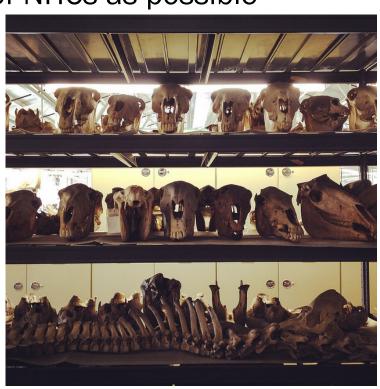


## Why?

- What training could be made available to natural history museum staff that work with a database?
- How could usability be improved to accurately capture information NHCs want to preserve?
- What would YOU want?

#### How to find the answers?

- Research protocol approved by The University of Texas at Austin's Office of Research Support Institutional Review Board (IRB)
- Visiting as many diverse kinds of NHCs as possible
  - Discipline
  - Database system used
  - Institution size
  - Geographically varied ideally



#### How?

- Participants collections staff that interact with and maintain the database as part of their job responsibilities
  - Audio-recorded
  - Anonymous and confidential
- Estimated time = approximately2 hours per institution
  - 30 minutes to 2 hours per participant



## Step 1 – Observation

- Participants are asked to perform specific, routine database tasks that fall within their job responsibilities
  - Specialization of tasks within a single institution

## Step 2 - Replication

- Researcher repeats these tasks for herself
  - Eliminate bias from an individual's experience level

## Step 3 – Interview

- Participants' opinions about the software in use at their institution
- Any experience with available technical support
- Any training experience
- How database tasks are divided among staff, if at all
- Operational questions of how certain things are done at the particular institution
- Future collections plans

## **Step 4 – Questionnaire**

- Educational experience and relevant training
- Employment information
- Collection-specific information

## **Currently...**

- Funding for travel?
- 2 site visits, 1 trial visit so far
- Recruitment and scheduling of future visits ongoing



### **Observations**

- Very different database knowledge-levels
- Similar future goals
- A bit "Specify-heavy"

## **Questions for You**

- As collections staff, what information would be helpful to learn from this study?
- Do you think Step 2 Replication is a viable way to eliminate the bias of experience?
- Email Ibrensk@utexas.edu

## Acknowledgements

- Committee: Chris Bell, Matt Brown, Unmil Karadkar, Tim Rowe
- Ann Molineux and Angie Thompson
- Fellow paleontology graduate students at UT
- Thank you for listening!