

## Overview

Assembling and organizing data often occurs over time. Differing approaches to data storage, organization, and metadata may be used at different stages of project development. A comparison is provided of file systems (#1) and relational databases (#2, #3) for heterogeneous field data projects.

## Three Approaches to Data Organization

**1. File system** with files named and placed logically, hierarchically for data storage and organization.

**Strength:** Change is handled with less effort for file systems than for databases; change is a property of high value at the beginning of a project.

**Weakness:** File systems can not have many too many relationships, which makes some analysis difficult.

**2. Relational Database Single Key** (1 to n relations) with a single key defining relations for 1-to-n queries so multiple files can be opened but specific information cannot be pulled out. This works well for data that can be assembled in a single table but not at the variable level.

**Strength:** More structure with some flexibility, so it can identify and access many files easily.

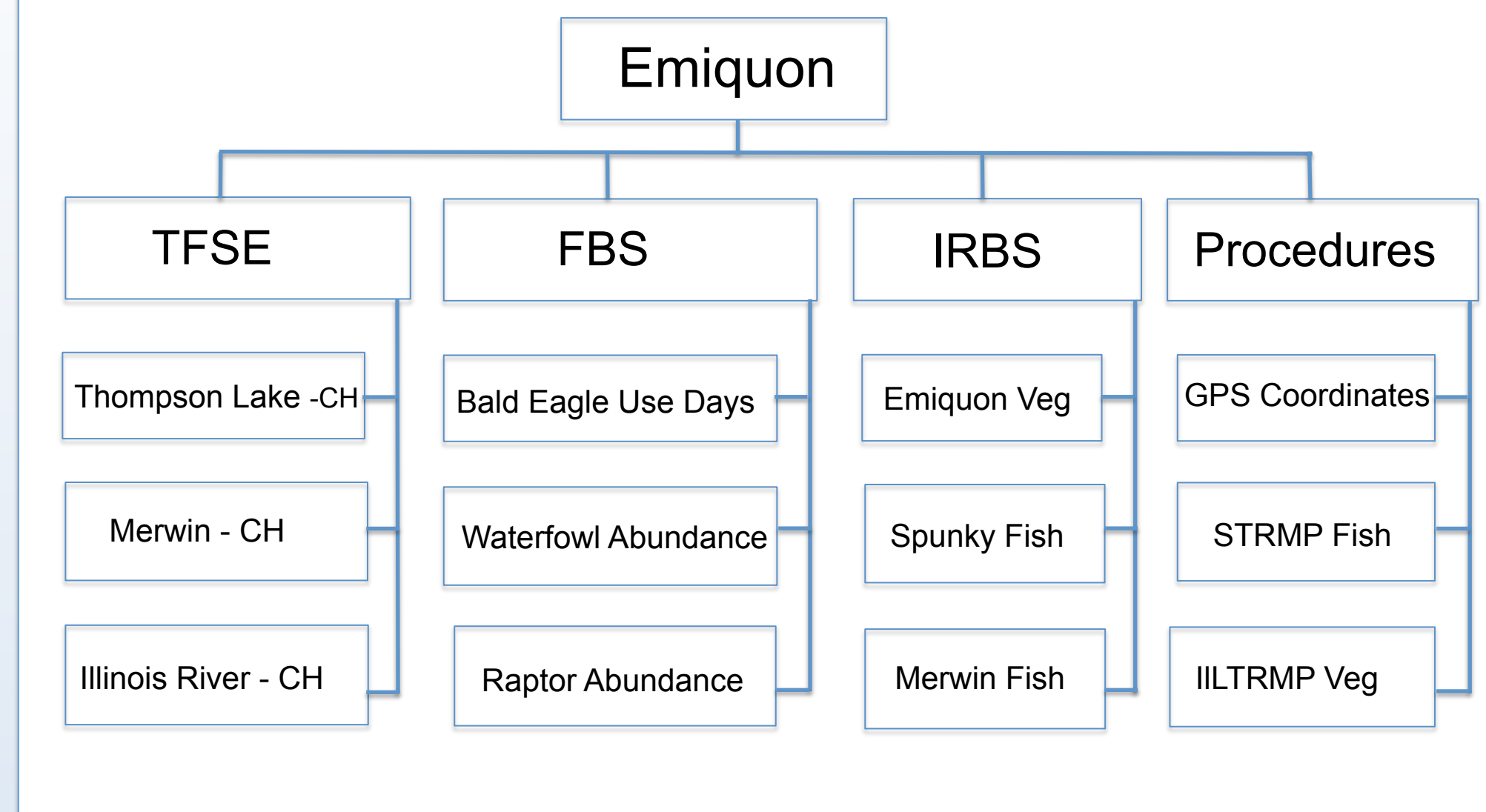
**Weakness:** There are no many to many relationships so complex analysis is difficult.

**3. Relational Database Multiple Relations** (n-to-n queries) with multiple keys that facilitate complex queries and allow subsets of data from multiple tables to be assembled into a single product.

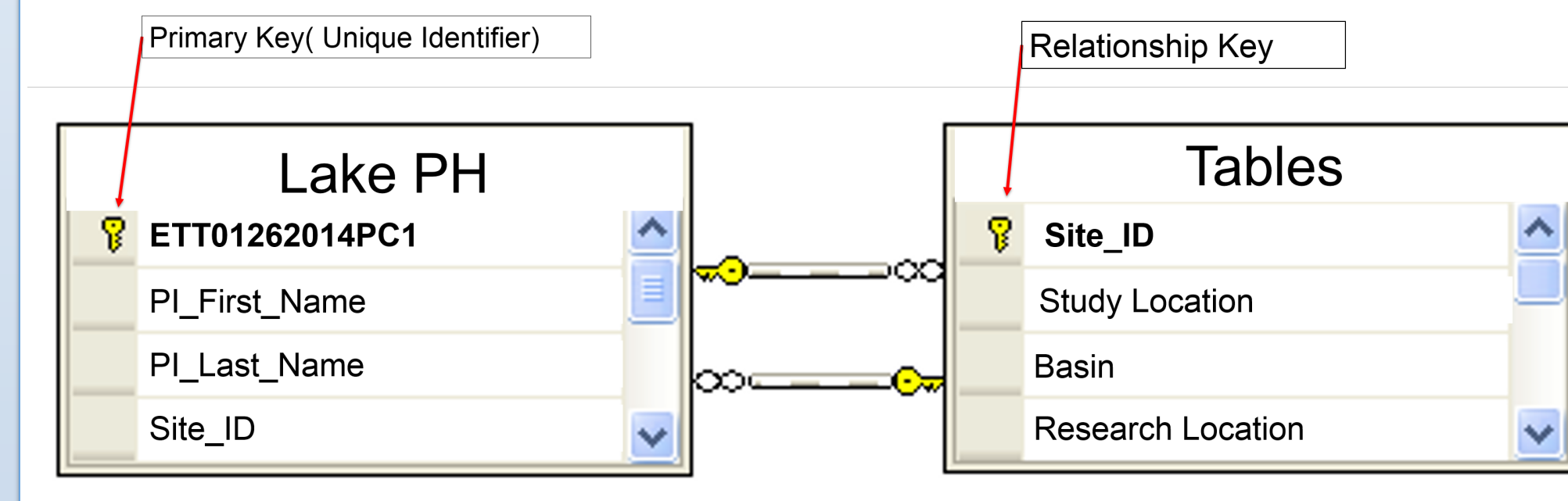
**Strength:** Databases can query across many tables to support complex, efficient analysis.

**Weakness:** Databases are rigid designs with set rules and programmatic constraints can make changes and redesign options difficult.

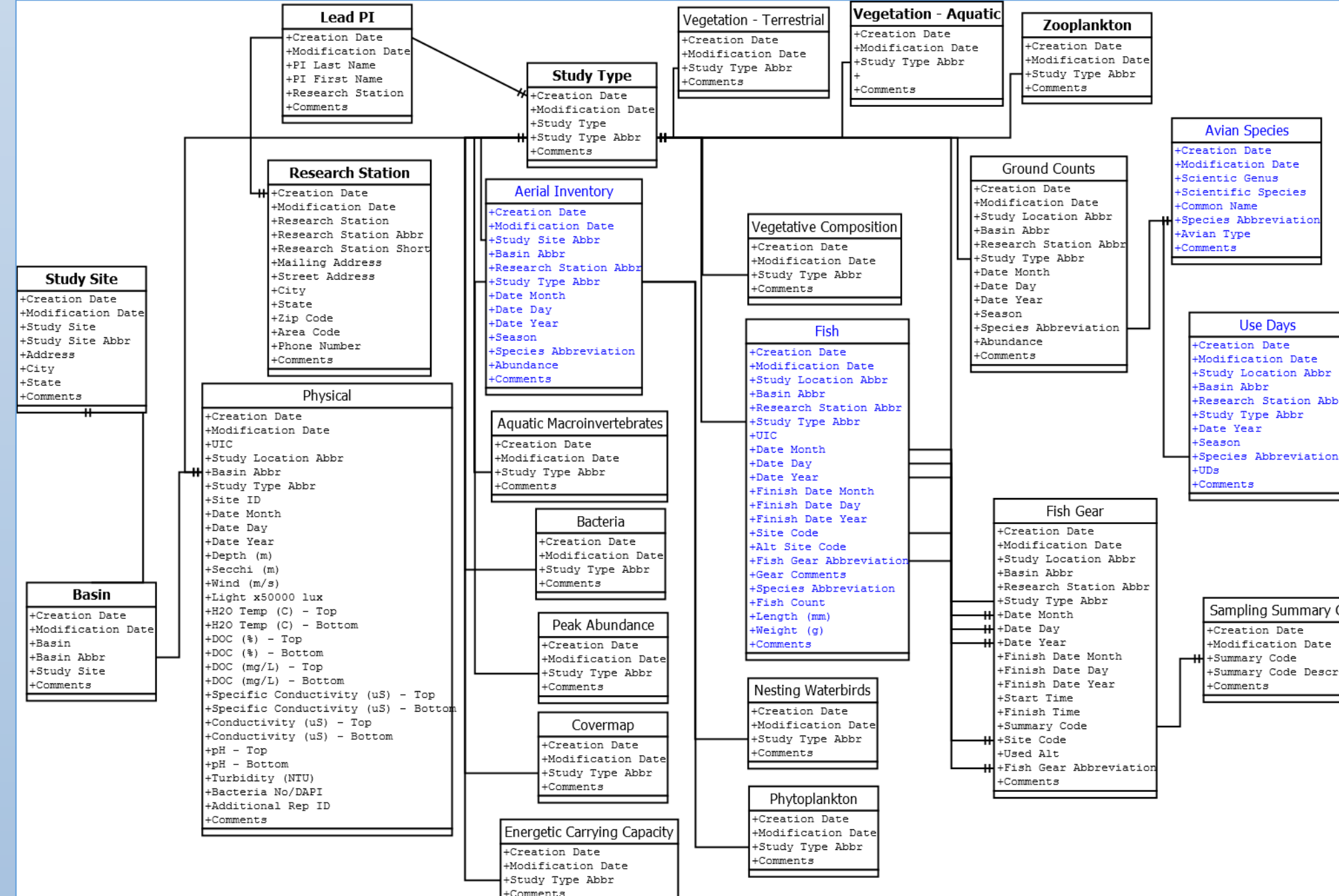
## 1. File System



## 2. Relational Database Single Key



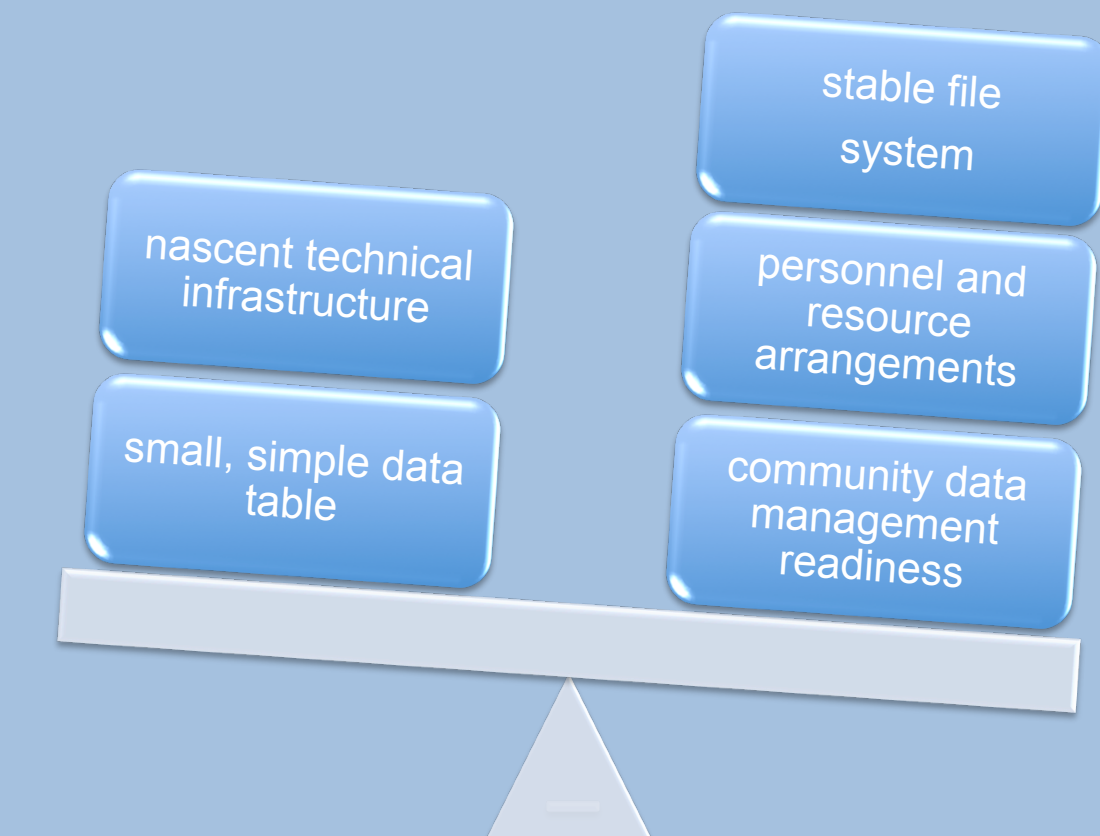
## 3. Relational Database Multiple Relations



## Kinds of Metadata

1. File System: Readme file, file names & headers (e.g. Box)
2. Relational Database Single Key: One key (e.g. FileMaker Pro)
3. Relational Database Multiple Relations: Multiple keys, data dictionaries & machine readable form (e.g. Access)

### Factors for making a transition



## Examples of Kinds of Databases

By content type:

- Catalog
- Document-oriented
- Full-text
- Graphic
- Photographic
- Knowledge
- Platform stream
- Real-time

By subject:

- Spatial (Geographical)
- Temporal (Time period)
- Project
- Theme/Phenomenon
- Domain
- Botany
- Chemical
- Ecological
- Rivers (hydro)

## Acknowledgement

Supported by National Science Foundation (NSF DEB, Rapid Grant# 1347077) and the Institute of Museum and Library Services (IMLS) Data Curation Education in Research Centers (DCERC, Award# RE-02-10-0004-10).