MongoDB Cloud

Observer App profile Config collection document* Field observation collection

* Common set for all observers. It's essentially the menu entries for the app.

PostgreSQL / PostGIS Backroom

Extract schema

Copy documents from each collection to target tables, as unnormalized data. Retrieve per date range.

- App config
- Observer profiles
- Field observations

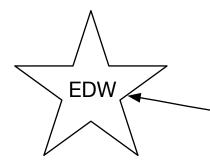
Transform schema

- 1. Normalize the app configuration data; merge with historic data tables and compute SCDs, which feed SCDs in the EDW. Examples:
- species
- calling codes, noise levels, etc
- observer profiles
- 2. Normalize and standardize the "sites."
- spatialize the sites; eliminate "suspect" sites.
- smooth the sites (to avoid excessive duplication)
- enhance sites with spatial attributes (e.g., watershed)
- 3. Normalize and standardize the field observations.
- flag suspect or thin observations.

Load schema

Views corresponding to the SCD dimensions in the EDW. View corresponding to the Fact table

Work table of new date records to add to date dim.



Data transformations, in PostgreSQL

1. App Data Synchronization

The Cloud collection of App Data contains one document per value picker (e.g., calling codes, wind scale, land use type, etc.). It's the current, official dataset, used by the application. For the ETL and EDW, it's been "downloaded" to the Extract schema in PostgreSQL as an unnormalized table. The "previous copy" is in the Transform schema. Sync it with the Cloud copy just downloaded to the Extract schema.

Extract Schema

Unnormalized Dataset
Collection name
Key
Value Description

Transform Schema

Species	Calling Codes	Land Use	Wind Scale	
Key	Code	Code	Code	et
Family	Description	Description	Description	
Genus	Date Added	Date Added	Date Added	
Species Common Name Date Added	(Type 2 SCD)	(Type 2 SCD)	(Type 2 SCD)	•

(EDW is Type 3 SCD)

2. Observer Profile Synchronization

The cloud collection contains one document per unique observer (named or anonymous). Each observer document also contains a sub-document of the geographic sites that have been observed by this observer.

Extract Schema

Observer Documents
Observer Unique ID
Affiliation
Contact data if provided
Date Added / Updated

Site Sub-Documents
Observer Unique ID
Site Unique ID
Map coodinates
Land use
Date Added / Updated
Habitat Sub-document
Water level
Date/Time Added/Updated

Transform Schema

Observer Documents
Observer Unique ID
Affiliation
Date Added / Updated

(EDW is a Type 3 SCD. Affiliation only.)

Site per observer
Observer Unique ID (foreign key)
Site Unique ID
Map Coordinates
Land Use code
Water shed ID
Geo Admin ID (s)
Date Added / Updated

Habitats within Observer Site
Site Unique ID
Habitat label
Water Level
Date Added / Updated

(EDW is a Type 4 SCD that combines the two)

3. Observations Synchronization

Extract Schema

Observation Documents
Observer
Site
Species
Calling Code
tba
tba

Transform Schema

Observation Records
Observer
tba
Species
Calling Code
tba
tba
Date Added / Updated

(This table becomes the basis of the fact table.)

Landing New Sesson Define New Site **Record Observation** FrogTalk Pick Habitat (within site Name Pick Site (Instructions: The site is the general location. It (Examples: Damn side of lake. Outlet end of Menu might contain multiple habitats, or multiple 0 -sign in observer points. Example: Robinson Lake. 0 -profile -setting Started Water level selection time Notable changes at habitat **Observation History** Weather Conditions Site A Species observed at this habitat -current condition selection date ... Species selection Calling Code -wind scale selection date ... (map view) -rain last 48 hours selection Species selection Calling Code -current temperature (C/F) Species selection Calling Code Weather Station ID Site B Species selection Calling Code Centroid of the map view is date ... **Observations Collected** stored. 0 date ... 0 Fire trigger if location services indicates a move -Habitat 1, species list -Habitat 2, species list -land use code selection Save: clear, -water source selection Save and Change Habitat habitat list New Session Done Save Done Back to landing