# New Mexico Office of the Medical Investigator: Database Conceptual Design Document (CDD)

**Original Author:** Andrew Barndt, Phillip Rieman

**Modified by: Anmol Gill, Kunj Bhavsar**

March 07, 2022

# Contents

# 1 The Mini-world (or Universe of Discourse)

For this database, our mini-world centers around the forensic anthropologists and staff that work at the state of New Mexico's Office of the Medical Investigator (OMI) and the University of New Mexico's Maxwell Museum of Anthropology (Museum). These anthropologists work to recover, identify, and examine remains that are reported all over the state of New Mexico. Whether each case implies a crime or not, each subject is collected and classified based on physiological and biological standards as well as environmental factors like taphonomic indicators and trauma. In order to successfully perform their job, forensic anthropologists need a database they can use frequently and with efficiency when handling caseloads of several dozen or more at a time.

## 1.1 Purpose and Scope

Ultimately, the purpose of the database is to aggregate and organize the contents of all cases handled by the OMI and the Museum that require forensic anthropological analysis. This is not a database to store other information of the OMI's or Museum's information. Thus, the scope of the database is all forensic anthropology cases referred to the OMI and Museum. Not all death cases in the state of New Mexico will be stored, nor will any case from other states be contained within the database.

## 1.2 Information Requirements

All information regarding an individual case will be stored within the database. This includes information about the person, describing the person's characteristics (age, sex, bioaffinity, stature, and where the person is from), if the person is identified or not, and estimations of characteristics using the person's remains. Each case also may contain photos or other forms of visualizing a case, so photos and illustrations will also be included. Bone data is also contained - both an inventory of all bones in the human body and the bones found at the location. Additional bone-related information, such as the trauma suffered, any analysis performed using the bone (DNA or isotope sampling, pathological tests, methods used to estimate the characteristics above). Taphonomic analysis is also to be contained in the database, which includes the date and location the remains were found, what kind of circumstances the person was found in (inside an unmarked grave, for example)

and conditions that affected the remains (if there is noticeable fire damage, for example), the conditions of the remains themselves (including animal activity), any personal artifacts, and lab analysis for weathering, bleaching, or biological processes affecting the state of the remains.

### 1.3 Users

Users from both the OMI and Museum will have access to the contents of the database. Regardless of a given user's affiliation, however, the dominant user group will be forensic anthropologists. Thus, the database will be built around the needs of forensic anthropologists. Other users may access the database if needed, but such instances we anticipate being rare enough that the database need not worry about these outliers.

This document's purpose is to explain the database in higher-level conceptual terms. It will discuss the overall design decisions and structure of the database, listing the different information and types of data that will be stored and their relationships to each other. Implementation of the database is not covered in this document.

The rest of this document proceeds as follows: Section 2 describes the notation used in the descriptions of each component of the database schema. Section 3 provides a comprehensive list of all entities, relationships, and attributes, the schema providing an illustration of the proposed database, and the integrity constraints required. Section 4 covers numerous anticipated queries the database will need to satisfy.

## 2 Notation and Definitions

Notation in this document is provided by the LaTeX template as given by the instructor.

The notation used: all upper case for the entity names, lower case for the relationship names, and the first letter capitalized for attribute names.

The description of the entities starts with a sentence which explains their meaning. Then the attributes to describe the instances are included. The relationships are also described by a sentence and a list of attributes if it has any.

Each attribute has a four-letter code which describes the type of attribute

according to the four classification criteria for attributes.
The format for this code is $(xyzw)$ where:

$x$ tells that the attribute is simple (S) or composite (C),

$y$ tells that the attribute has a single value (S) or is multi-valued (M),

$z$ tells that the attribute is primitive (stored) (P) or derived (D), in case it is derived, an explanation of how to deduce it from other attributes or a formula/procedure must be specified, and

$w$ tells that the attribute is fixed (F) (i. e. it must have a value that is not null) or optional (O), i.e. the domain of the attribute allows the null value.

# 3 Conceptual Schema of the Database

This section contains, in order, a list of entities, an explanation of each primary entity and its attributes, and the proposed schema diagram.

## 3.1 Entities

The entities defined for this database are:

- METHOD
- BONE
  - HEAD
    - SKULL
    - TEETH
  - LIMB
    - ARM
      - FOREARM
      - HAND
      - UPPER ARM
    - LEG
      - CALF

        - FOOT
        - THIGH
    - TORSO
        - PELVIS
        - RIBS
        - SPINE
- CASE
- DAMAGE
    - DECOMPOSITION
    - PATHOLOGY
    - TRAUMA
- IMAGE
- RECOVERED
- GENERIC

A detailed description of each entity follows.

**BONE:** Skeletal information regarding the case. Some example instances are: specific names of bones like Femur, Clavicle, and Radius.

| | | |
|---|---|---|
| Attributes: | Measurement | (CSPF) |
| | - Consists of a Value and a Unit, such as 3 kilograms | |
| | Color | (SSPF) |
| | Integrity | (SSPF) |
| | DNA_Sent | (SSPF) |
| | DNA_Result | (SSPO) |
| | Isotope_Sent | (SSPF) |
| | Isotope_Result | (CSPO) |
| | - Consists of the Sample_Type and a Result | |

**HEAD:** Subset of BONE that pertains to the cranial portion of the skeleton.

**SKULL:** Subset of HEAD that relates to the parts of cranial region not including the teeth.

**TEETH:** Subset of HEAD that relates only to the teeth.

**LIMB:** Subset of BONE that pertains to the arms and legs of the skeleton.

**ARM:** Subset of LIMB that relates to the upper limbs of the skeleton.

**FOREARM:** Subset of ARM that relates to the portion of the arm between the wrist and the elbow.

**HAND:** Subset of ARM that relates to bones including and below the wrist.

**UPPER ARM:** Subset of ARM that relates to the portion of the arm from the elbow to the shoulder.

**LEG:** Subset of LIMB that relates to the lower limbs of the skeleton.

**CALF:** Subset of LEG that relates to the portion of the leg between the ankle and the knee.

**FOOT:** Subset of LEG that relates to the portion of the leg including and below the ankle.

**THIGH:** Subset of LEG that relates to the portion of the leg between the knee and the pelvis.

**TORSO:** Subset of BONE that pertains to the middle section of the skeleton from the neck to the hip.

**PELVIS:** Subset of TORSO that relates to the portion of the skeleton below the rib cage and above the leg.

**RIBS:** Subset of TORSO that relates to the portion of the skeleton below the neck and above the pelvis.

**SPINE:** Subset of TORSO that relates to the collection of bones that span the length of the torso behind the rib cage from the pelvis up to and including the neck.

**RECOVERED:** Subset of BONE that was recovered at the scene.

**GENERIC:** Subset of BONE used for reference of what an ideal bone is. These would be stock or example bones.

Attributes: Belongs_To_Adult (SSPF)

**CASE:** Identification classifier assigned by the OMI used to track and organize different cases under investigation. Additionally contains several attributes holding taphonomic information.

| | | |
|---|---|---|
| Attributes: | Case_Number | (SSPF) |
| | Date_Opened (MM/DD/YYYY) | (CSPF) |
| | - Consists of Month, Day, and Year | |
| | Date_Found (MM/DD/YYYY) | (CSPF) |
| | - Consists of Month, Day, and Year | |
| | Person_Name | (CSPO) |
| | - Consists of First, Middle, and Last | |
| | Person_DOB (MM/DD/YYYY) | (CSPO) |
| | - Consists of Month, Day, and Year | |
| | Person_SSN | (SSPO) |
| | Person_Stature | (CSPO) |
| | - Consists of Feet and Inches | |
| | Person_Sex | (SSPO) |
| | Person_Ancestry | (SMPO) |
| | Location | (CSPF) |
| | - Consists of Number, Street_Name, Zip_Code, City, and County | |
| | Environment | (SMPF) |
| | Indoors | (SSPF) |
| | Outdoors | (SSPF) |
| | Identified | (SSPF) |
| | Resolved | (SSPF) |
| | Era | (SSPF) |
| | Clothes_Found | (SMPO) |
| | Body_Orientation | (SSPF) |
| | Overall_Decomposition | (SSPF) |

**DAMAGE:** Contains information regarding any kind of malformity. Instances can include examples like Fracture or Charring.

Attributes: Severity (SSPF)

**DECOMPOSITION:** Subset of DAMAGE that reports on the amount of decomposition present. Part of taphonomic analysis.

| | | |
|---|---|---|
| Attributes: | Rodent | (SSPF) |
| | Insect | (SSPF) |
| | Carnivore | (SSPF) |
| | Weathering | (SMPF) |

**IMAGE:** All visual components of a case that is retained in a physical and/or digital version. Example instances include CT scans and Graphs.

| | | |
|---|---|---|
| Attributes: | Name | (SSPF) |
| | Body_Area | (SMPF) |
| | Type | (SSPF) |

**METHOD:** Analysis used to determine a person's characteristics such as age, sex, stature, and estimated postmortem interval. Examples include the aging methods provided by OMI, such as Brooks & Suchey (1990), a method for estimating age.

| | | |
|---|---|---|
| Attributes: | Name | (SSPF) |
| | Paper | (SSPF) |
| | Measures | (SSPF) |

**PATHOLOGY:** Subset of DAMAGE that indicates what, if any, diseases are present. Example instances include Osteoporosis or Arthritis.

| | | |
|---|---|---|
| Attributes: | Name | (SSPF) |
| | Chronic | (SSPO) |
| | Disease_Length | (SSPO) |

**TRAUMA:** Subset of DAMAGE that describes, if present, forcibly inflicted damage.

| | | |
|---|---|---|
| Attributes: | Type | (SSPF) |
| | Time | (SSPF) |

### 3.2 Relationships

The relationships in this schema are listed and described below.

**apply**

Attributes: Result (CSPF)
- consists of Age, Sex, Stature, or Ancestry

**collect**

No attributes.

**observe**

No attributes.

**refers_to**

No attributes.

**visualize**

No attributes.

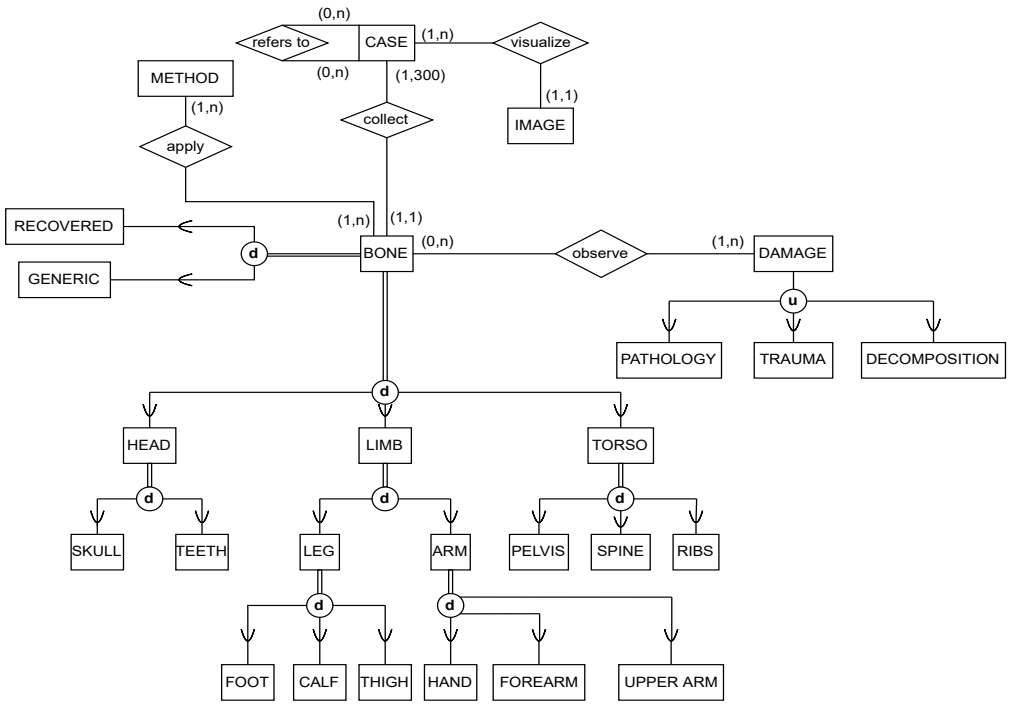


Figure 1: Proposed OMI Database

### 3.3 Semantic Integrity Constraints

1. When describing a composite attribute's individual atomic attributes (such as Person Name consisting of a First, Middle, and Last), the composite attribute's name will come first followed by the atomic attribute. e.g. Person Name Last refers to a person's last name.

2. DOB must be older than all other dates in a case.

3. Attributes with a domain of {1,2,3,4,5,6,7,8,9,10}, with 1 indicating little to no presence and 10 indicating very strong presence:

4. For the Result attribute in the apply relationship, the result is about one of four things: age, sex, stature, or ancestry.
   Age will consist of one result with three numerical components: Min Age, Max Age, and Confidence. All three must be a number using only 0-9.

5. Sex will consist of one result with two components: Sex Estimation and Confidence. Sex Estimation has the domain {"Male", "Female"} and Confidence must be a number using only 0-9.

6. Stature will consist of five numerical components: Min Ft, Min In, Max Ft, Max In, Confidence. All five must be a number using only 0-9.

7. Ancestry will consist of one result with two components: Bioaffinity and Confidence. Bioaffinity has the domain {"White", "Asian", "Hispanic", "African", "Native American"} and Confidence must be a number using 0-9 only.

# 4 Example Queries

A list of the most important queries.

1. Provide total cases by day/month/year.
2. Provide total cases by race/ethnicity.
3. Provide total cases by age and sex.
4. Provide total cases by distribution (accident, homicide, natural, etc.).
5. List all cases where a certain percentage of bones were recovered.
6. List all cases where bones exhibited certain signs of taphonomic activity.
7. List all cases where general keywords were included.
8. List all cases where craniometrics were taken.
9. List all cases with cranial trauma.
10. List all cases with ballistic trauma.
11. List all cases with a certain ancestry/bioaffinity.
12. List all cases with a complete skeleton.

### 4.1 Possible extensions and additional comments

It was not clear whether cases that were resolved or closed would be deleted from the database, therefore to err on the side of caution, a "resolved" attribute was added to CASE in order to record this fact.

# 5 The Logical Relational Schema (LRS) for the NM-OMI

The conceptual schema described for the NM-OMI Database is mapped into the Relational Schema presented in this section. All the attributes underlined in the same Relation belong to the primary key. By default, all the attributes do not belong to the primary key may be null, unless explicitly specified that they cannot be null.

**CASE**(Case_Number, Date_Opened, Date_Found, Person_Name, Person_DOB, Person_SSN, Person_Stature, Person_Sex, Person_Ancestory, Location, Environment, Indoors, Outdoors, Identified, Resolved, Era, Clothes_Found, Body_Orientation, Overall_Decompostiton)

{ Case_Number, Date_Opened, Date_Found, Location, Environment, Indoors, Outdoors, Identified, Resolved, Era, Body_Orientation, Overall_Decompostiton } may not be null

**BONE**(Case_Number, Measurement, Integrity, DNA_Sent, DNA_Result, Isotope_Sent, Isotope_Result, Bone_Name)

Case_Number is foreign key, references CASE

{ Case_Number, Measurement, Integrity, DNA_Sent, Isotope_Sent, Bone_Name } may not be null

**METHOD**(Bone_Name, Name, Paper, Measure)

Bone_Name is foreign key, references BONE

{ Bone_Name, Name, Paper, Measure } may not be null

**DAMAGE**(Bone_Name, Severity)

Bone_Name is foreign key, references BONE

{ Bone_Name, Name, Severity } may not be null

**PATHOLOGY**(Name, Chronic, Disease_Length)

Name may not be null

**TRAUMA**(Type, Time)

{ Type, Time } may not be null

**DECOMPOSITION**(Rodent, Insect, Carnivore, Weathering)

{ Rodent, Insect, Carnivore, Weathering } may not be null

**IMAGE**(Case_Number, Name, Body_Area, Type)

Case_Number is foreign key, references CASE

{ Case_Number, Name, Body_Area, Type } may not be null

**GENERIC**(Bone_Name, Belongs_To_Adult)

Bone_Name is foreign key, references BONE

{ Bone_Name, Belongs_To_Adult } may not be null

**apply**(Result)

Result may not be null

### 5.1 Additional Integrity Constraints for the relational schema

The integrity constraints that must hold for the OMI database and that are not guaranteed by the relation schemas described above are listed in this subsection.

1. For each tuple in **CASE** the value of Date_Opened, Date_Found must be greater than Person_DOB if Person_DOB is not null.
2. In **CASE**, if one of Indoors or Outdoors has a value 'Yes' the second one should have the value 'No'.
3. For every tuple in **BONE,** there has to be a subset referring to the origin of the bone. Example: HEAD, LIMB, TORSO etc.
4. In **BONE** DNA_Result and Isotope_Result cannot have a value with DNA_Sent and Isotope_Sent without any value.
5. Every Bone_Name value in the tuples of **METHOD, DAMAGE** and **GENERIC** must appear in at least one tuple of **BONE.**
6. Every Case_Number value in the tuples of **BONE** and **IMAGE** must appear in at least one tuple of **CASE.**
7. Every tuple in **BONE** must belong to only one of its subsets (HEAD, LIMB, TORSO) and cannot be present in two or more.
8. Every tuple in **HEAD** must belong to only one of its subsets (SKULL, TEETH) and cannot be present in two or more.

9. Every tuple in **LIMB** must belong to only one of its subsets (**LEG**; FOOT or CALF or THIGH, **ARM;** HAND or FOREARM or UPPERARM) and cannot be present in two or more.

10. Every tuple in **TORSO** must belong to only one of its subsets (PELVIS, SPINE, RIBS) and cannot be present in two or more.

11. Every tuple in **BONE** must be either RECOVERED or GENERIC.