1107954  SANGHYEON CHO

ICT320 Task2 part a

Contents

[1. Identification of major implementation flaws in the existing system, and/or areas for improvement 2](#_Toc22857689)

[1.1 Removing repeated data, 2](#_Toc22857690)

[1.2 De-normalization for optimization 2](#_Toc22857691)

[1.3 changing some areas to NoSQL 2](#_Toc22857692)

[1.4 Rational for the creation of each specific constraint (Foreign Key and Unique) and/or index. 2](#_Toc22857693)

[1. Primary key 2](#_Toc22857694)

[2. Foreign Key 3](#_Toc22857695)

[3. Index 3](#_Toc22857696)

[2. Identification of the current Normal Form of the system. 3](#_Toc22857697)

[3. An ER Diagram for the current design of the system, and an ER Diagram for the re-designed system. 3](#_Toc22857698)

[3.1 Current ER Diagram 3](#_Toc22857699)

[3.2 Re-designed Diagram 5](#_Toc22857700)

[4. Your assumptions and any supplementary design requirements regarding the system. 5](#_Toc22857701)

# 1. Identification of major implementation flaws in the existing system, and/or areas for improvement

## 1.1 Removing repeated data,

* The values of ANIMAL ID are repeated in Table ‘Contact and Morph’
* The values of column ‘TAGGED’ and ‘DNA’ are repeated in table ‘Individuals’

## 1.2 De-normalization for optimization

Data redundancy is found in table ‘Contacts and morph’. When animals are observed, the animal id is recorded repeatedly. So if the values are sorted out by its ID, we can avoid the data wasting. Create a temporary denormalized table which contains date, time, type, lat, long and location sighted columns with animal id array.

1.3 changing some areas to NoSQL (if so, include what type of NoSQL database).

In ‘Individual’ table, all the rest of the columns are listed by animal id. It means that animal id contains all the information. the database can be changed to MongoDB format which is document type of NoSQL as the table below.

|  |  |
| --- | --- |
| Key | Document |
| 1 | {Aniname: “Acer”, Photo: “False”, Sex: “M”… } |
| 2 | {Aniname: “Admiral”, Photo: “True”, Sex: “M”… } |
| 3 | {Aniname: “Ajax”, Photo: “True”, Sex: “M”… } |

## 1.4 Rational for the creation of each specific constraint (Foreign Key and Unique) and/or index.

### Primary key

* ‘Animal\_id’ in Animal table
* ‘Photo\_id’ in Photo table
* ‘Tag\_id’ in Tag table
* ‘DNA\_id’ in DNA table
* ‘Cont\_id’ in Contact table
* ‘Morph\_id’ in Morph table
* ‘User\_initial’ in Users table
* ‘Record\_id’ in Record table

### Foreign Key

* In ‘Photo’ table has a foreign key ‘Animal\_id’ which is connecting to ‘Animal\_id’ in ‘Animal’ table
* In ‘Tag’ table has a foreign key ‘Animal\_id’ which is connecting to ‘Animal\_id’ in Animal table
* In ‘DNA’ table has a foreign key ‘Animal\_id’ which is connecting to ‘Animal\_id’ in Animal table
* In ‘Contact’ table has a foreign key ‘Animal\_id’ which is connecting to ‘Animal\_id’ in Animal table
* In ‘Morph’ table has a foreign key ‘cont\_id’ which is connecting to ‘cont\_id’ in Contact table
* In ‘Record’ table has a foreign key ‘cont\_id’ which is connecting to ‘cont\_id’ in Contact table

### Index

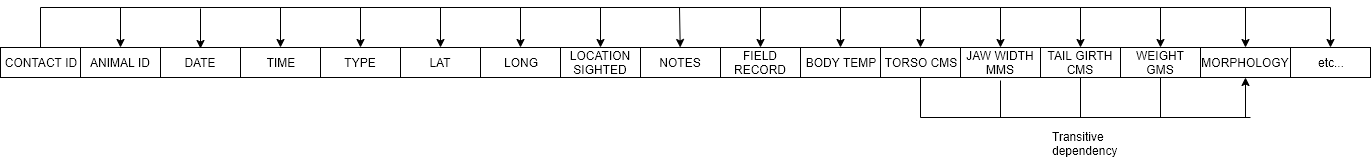
* Unique index : CREATE UNIQUE INDEX User\_NDX1 ON Users (User\_initial);
* INDEX: CREATE INDEX ANIMAL\_NDX ON Contact (Animal\_id);

CREATE INDEX STATE\_NDX ON VENDOR(V\_STATE);

In plan A the DBMS uses a full table scan of VENDOR. The SORT operation is done to order the output by vendor state. In Plan B, the DBMS uses an Index Scan Range of the STATE\_NDX index to get the VENDOR RowIDs. After the EMPLOYEE RowIDs have been retrieved, the DBMS uses those RowIDs to get the VENDOR rows. Finally, the DBMS sorts the result set by the state of vendor.

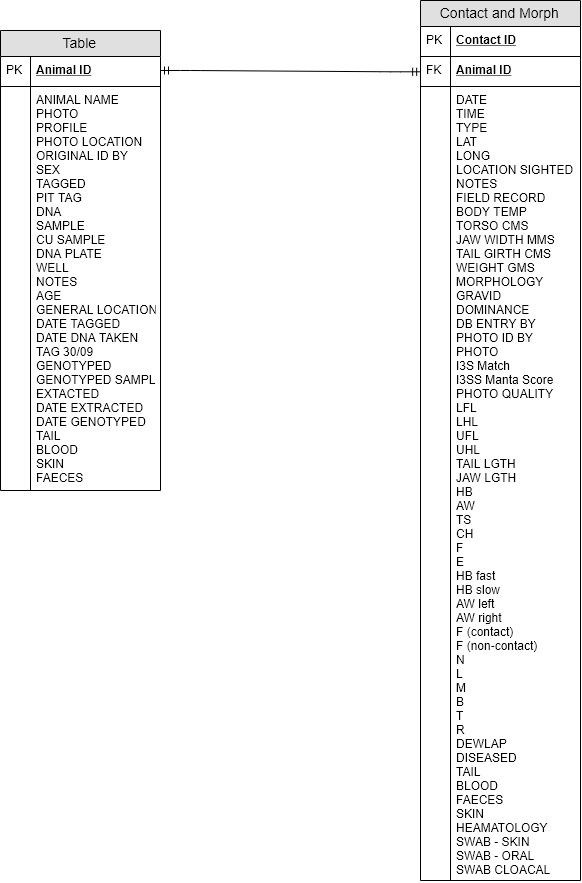
# 2. Identification of the current Normal Form of the system.

* The current system should be **second Normal Form (2NF)** since there is no partial dependencies both of tables but there is transitive dependency in table ‘Contact and Morph’. When the values are existed in columns ‘TORSO CMS’, ‘JAW WIDTH MMS’, ‘TAIL GIRTH CMS’ and ‘WEIGHT GMS’, the value of ‘Morphology’ column is ‘True’.

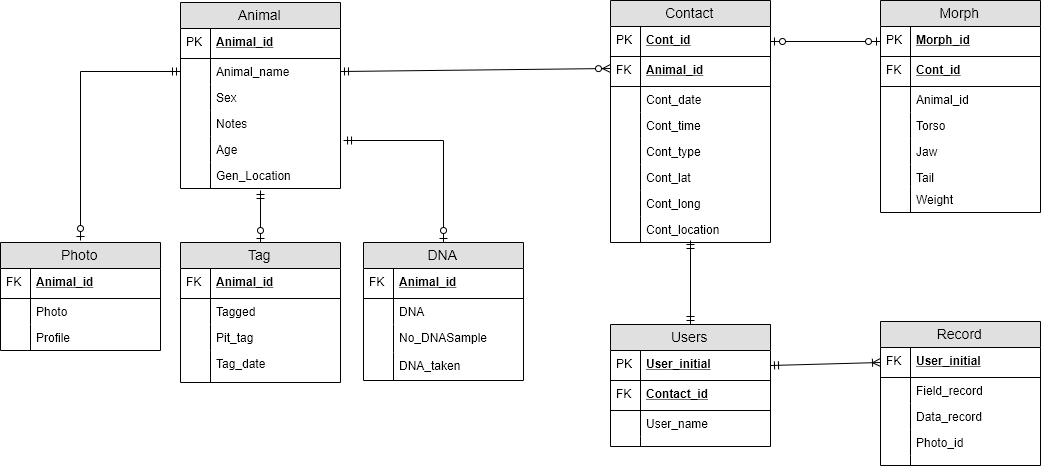


# 3. An ER Diagram for the current design of the system, and an ER Diagram for the re-designed system.

## 3.1 Current ER Diagram



## 3.2 Re-designed Diagram



# 4. Your assumptions and any supplementary design requirements regarding the system.

1. Split ‘Individuals’ table into four smaller tables.
   1. Table ‘Animal’ contains Animal\_id, Animal\_name, Sex, Notes, Age and general location.
   2. Table ‘Photo’ contains Photo\_id, Animal\_id, Photo and Profile. Photo\_id will automatically have own number.
   3. Table ‘Tag’ contains Tag\_id, Animal\_id, Tagged, Pit\_tag and Tag\_date. Tag\_id will automatically have own number.
   4. Table ‘DNA’ contains DNA\_id, Animal\_id, DNA, No\_DNASample and DNA\_taken. DNA\_id will automatically have own number.
2. Split ‘Contact and Morph’ table into three smaller tables.
   1. Table ‘Contact’ contains Con\_id, Animal\_id, Cont\_date, Cont\_time, Cont\_type, Cont\_lat, Cont\_long and Cont\_location.
   2. Table ‘Morph’ contains Morph\_id, Animal\_id, Cont\_id, Torso, Jaw, Tail, Weight and Morphology. Morph\_id will automatically have own number.
   3. Table ‘Record’ contains Record\_id, Cont\_id, Field\_record, Data\_record and Photo\_id. Record\_id will automatically have own number.

3) Create table ‘Users’ to clarify easily the person who observed animals and recorded data.

* 1. Table ‘Users’ contains User\_initial and User\_name.

1. the rest of the columns will be eliminated in both of tables since those are redundant.
2. To create view for animal with 25 or more contacts and DNA, select Animal\_id and count number of Animal id from Contact table. And then used Having clause to show the animals only counted more than 25 times.
3. To create view for morphologies per animal, select animal\_id and count the number of rows where the value equal ‘TRUE’ from Morph table.
4. To create procedure for animal with the greatest number of contacts, selected the greatest number of animal id observed by using ‘MAX(count)’ clause.
5. To create procedure for animal with the greatest number of morphologies, selected the greatest number of animal id where the value equal TRUE by using ‘MAX(count) clause.
6. TO import data from both of ‘csv’ files, used Table data import wizard function in MySQL.