Requirements Analysis

Project Phase #1

Data And Applications

—

Kannav Mehta	2019101044
Triansh Sharma	2019101006
Raj Maheshwari	2019101039

Overview

The Project aims to manage the data of the miniworld - **The National Park System**.

The design of the database aims to provide an information system about the different aspects of the wildlife in India. The construction of the database is done in such a way so as to support a variety of activities such as conservation of flora and fauna, scientific collaboration with third-party entities and also recreational and tourism activities of National Parks like treks, boating, hiking, safaris etc.

The prospective users of this database are the National Park authorities to help them maintain and monitor the ecological systems of the National Park, researchers engaging in studies and experiments on park wildlife, tourists that visit the National Parks and apply for permits, volunteer duties etc. Thus the functionalities are broken down into three primary views:

- Administrative
- User
- Scientific

Database Requirements

NATIONAL PARK

❖ The database system describes various NATIONAL PARKS, and each National Park shall be described by:
- ➢ *name*   **unique**
- ➢ *unit-code*   **unique id** number
- ➢ *region-code*   regional code
- ➢ *boundary*   In form of geospatial data
- ➢ *location*

ZONE

❖ Each National Park is further divided into ZONES each of which has:
- ➢ *zone_number*   **unique** identification number
- ➢ *contact*   phone number (must be **valid**)
- ➢ *terrain*   **multivalued** (e.g. rainforest, grassland etc.)

EMPLOYEES AND DEPARTMENTS

- ❖ The National Park Administration is itself divided into DEPARTMENTS related to multiple EMPLOYEES such as medical staff, Park Rangers etc.
- ❖ Each Employee can be described by
  - ➢ *emp_id*   **unique**
  - ➢ *emp_name*
  - ➢ *emp_email*
  - ➢ *gender*   (Male / Female / Others)
  - ➢ *date_of_joining*
  - ➢ *date_of_birth*
  - ➢ *contact-number*   (**valid**)
  - ➢ role   (officer, ranger, medic etc.)

- ❖ Each Department can be described by
  - ➢ *dep_number*   **unique**
  - ➢ *dep_name*
- ❖ Each department is chaired by exactly one employee and is related to other employees by the virtue of relationship WORKS_FOR.
- ❖ The employees shall be related to exactly one DEPARTMENT (WORKS_IN).

FEATURES

- ❖ Each zone contains several features, and each entity has a:
  - ➢ *feature_id*   **unique** *id-number*
  - ➢ *feature_name*
  - ➢ *location*
  - ➢ *availability*   days/timings
  - ➢ *images*
  - ➢ *average_rating*   **derived** (must be between 1 and 5)
- ❖ Each zone shall contain several features, which are **sub-classed** into
  - ➢ LODGING
    - ■ *name*   name of the lodging
    - ■ *capacity*   total number of available rooms
    - ■ *occupancy*   occupied rooms
    - ■ *person_in_charge*
    - ■ *contact_number*   phone number of Reception (**valid**)
  - ➢ PUBLIC FACILITIES:
    - ■ na*me*
    - ■ *person_in_charge*
  - ➢ TRAILS
    - ■ *name*
    - ■ *length*
  - ➢ VIEWPOINTS
    - ■ name
- ❖ For each trail we shall keep track of the various viewpoints/facilities that lie on it, by the relationship crosses.

SPECIES

- ❖ stores various species of flora and fauna with attributes:
    - ➢ *scientific-name*   **unique**, **composite key**
        - ■ *genus*
        - ■ *specific_name*
    - ➢ *taxonomy*
        - ■ *class*
        - ■ *order*
        - ■ *family*
        - ■ *taxonomy-code*   **unique**
    - ➢ *name*
    - ➢ *average_lifespan*
    - ➢ *common_names*   **multivalued**
    - ➢ *vulnerability*   (e.g. endangered, threatened, common)
    - ➢ *general_habitat*   (e.g. rainforests, mangroves etc.)

- ❖ A species shall also be related to other species with the relation FEEDS_ON. This relationship shall a **cardinality ratio** of **M : N**

- ❖ A species is related to a National Park by the PRESENCE relationship. This relationship shall store data specific to the National Park:
    - ➢ *record_status*   whether presence is verified or not
    - ➢ *record_time*   time when the last record was made
    - ➢ *occurrence*   is the species currently present
    - ➢ *nativeness*
    - ➢ *is_attraction*   boolean
    - ➢ *abundance*   (e.g. rare, common, uncommon)
    - ➢ *habitat*

- ❖ This relationship shall be a 3-way relationship identifying the **weak entity DEMOGRAPHY** that contains data about the *percent_of_population* under a certain *percent_of_lifespan,* along with a *time-stamp* of the data.

SERVICES

- ❖ A National Park also provides services that include tourist activities, events, field trips etc. Each service has attributes:
    - ➢ *service-id*   **unique**
    - ➢ *name*   e.g. safari, trekking, boating etc.
    - ➢ *availability*
    - ➢ *timings*
    - ➢ *capacity*
    - ➢ *description*
    - ➢ *price*
    - ➢ *average_rating*   **derived** (must be between 1 and 5)
- ❖ Each service can also be related to (INCLUDES) various *FEATURES*. Any service shall be related to a number of features like accomodation, viewpoints etc. This will have a **N : M cardinality**.

USER

- ❖ The database shall keep the following data about the users of these services (e.g. tourists).
  - ➤ *user-id*   **unique**
  - ➤ *user-name*
  - ➤ *contact-number*
  - ➤ *email*
  - ➤ *password*

- ❖ Each user shall be able to book one or more services, represented by the **identifying** relationship (BOOKS). Each of which has:
  - ➤ *booking-id*
  - ➤ *time*
  - ➤ *date*
  - ➤ *people_count*
    - ■ *number_of_adults*
    - ■ *number_of_children*

- ❖ A **weak entity TRANSACTION** identifiable by the *booking-id* and secondary **partial key** *transaction-id*. It has attributes
  - ➤ *time*
  - ➤ *date*
  - ➤ *price*
  - ➤ *pay_method*   e.g. credit card, net banking

- ❖ Any user-feedback on the services availed shall be stored in a **weak entity SERVICE_FEEDBACK** identified by the *user-id* and *service-id* and described by:
  - ➤ *date*
  - ➤ *rating*
  - ➤ *remarks*

- ❖ Any user-feedback on a feature shall be stored in a **weak entity FEATURE_FEEDBACK** identified by the *user-id* and *feature-id* and described by:
  - ➤ *date*
  - ➤ *rating*
  - ➤ *remarks*

- ❖ In case of volunteer activities, A shall can be related to a *DEPARTMENT* and *EMPLOYEE* by the relationship VOLUNTEER identified by *emp_id* and *dept_number* with attributes:
  - ➤ *volunteer-permit-id*
  - ➤ *date/time*
  - ➤ *job-description*

RESEARCHERS

- ❖ Researcher entity has following attributes:
  - ➢ *researcher-id*   **unique**
  - ➢ *name*
  - ➢ *contact-number*
  - ➢ *email*
  - ➢ *permit-id*
  - ➢ *permit-expiration-date*

- ❖ Each entity shall store information about researchers who work in affiliation with the National Parks.
- ❖ Now such a study would result in resources (DATA) in the form of actual data from the experiment. This shall have a:
  - ➢ *data_id*   **unique**
  - ➢ *description*
  - ➢ *report_data*
    - ■ *publishing_status*   whether published
    - ■ *ISBN-number*
    - ■ *report-title*
    - ■ *publishing_journal*
    - ■ *publishing_date*
    - ■ *co_authors*
  - ➢ *data_type*   (e.g. geospatial, numeric etc)
  - ➢ *data_link*   Additional Data accumulated, stored as a link to corresponding files

- ❖ Thus, the entities researcher, national park, report and species would constitute a **4-degree relationship**, which shall have the attributes: *duration* and *type* (like botanical, zoological etc.). The relationship (**STUDY**) shall be identified by a researcher (*unique-id*), having accumulated data (*data-id)* on some species (*scientific-name)* at some National Parks (*unit-code*). Each entity type will have **partial participation**.

Functional Requirements

Administrative

- ❖ The Park Administration must be able to retrieve data about the presence and abundance of various species present in the National Park to identify population changes and maintain census. This shall be in the form of a **report** detailing the populations of various species in a National Park.
- ❖ The Park Administration must also be able to **update demography values** periodically.

❖ Ability to query about the facilities along a particular trail or near a route using the geospatial data stored. The data-set shall be geo-hashed, so the method of query must be **partial match**.

❖ Ability to get the total population of a set of one or more species across a set of one or more national parks.

❖ To be able to get a list of a subtype of facilities for which the **average** rating is **above the desired value**, in a specific National Park.

❖ Ability to search for National Parks, where Animal Species is found. Querying can be implemented through **partial text match**.

❖ We shall be able to require for the **highest-earning service**s, or most popular services in each zone of some national park.

User

❖ The User Interface can be extended to be part of an online portal for tourists to manage their trips. Provision of the ability to query for the various services offered, check availability and timings which can be used to plan trips.

❖ The users shall be able to **provide feedback** on the various services and facilities in the National Park.

❖ The functionality to be able to **sum** up the costs of the services chosen and be able to **provide invoice** to the User.

❖ The users shall also be able to **cancel** their bookings.

Scientific

❖ Under this Interface Users shall be able to query for existing data or findings based on different parameters such as some subset of species, a particular National Park, Author, Study-type. They must also be able **retrieve** demographic information of a particular year/period etc.

Future prospects

❖ Applying Data Analysis techniques to form connections between different data-sets.

❖ Accommodating data created by IOT Devices in Real Time.

❖ This way of data storage can be extended to store geological data-sets as well.