Lab 2: Database requirements analysis and design

**Due date:** Monday, April 15, midnight.

**Note:** This lab involves a number of deliverables that are due before April 15. The due date above is for the final set of the deliverables.

This is a **team lab**. Each team submits one set of deliverables. Each team member receives the same grade for the assignment.

**Submission notes.** I will need to grade your submissions on Tuesday, April 16, so that I could communicate my feedback to you during Wednesday, April 17 class - when Lab 3 comes out. Because of this, the deadline is **hard**. See more submission instructions at the end of the document.

## Course Project Information

This quarter, CPE 366 comes with a team course project. General information about the project is provided below.

**Customer.** Our customer for this quarter is Dr. Heather Liwanag, head of Cal Poly's Vertebrate Integrative Physiology Lab, and a faculty member at the Department of Biology.

**Project.** The customer asks you to design a prototype web application for collecting and managing information about the observed populations of elephant seals along the California coast. Your goal during the quarter is to design the database to properly represent and store elephant seal observations (and other supplementary information necessary for the proper operation of the database), and implement a number of key use cases for data entry, access, retrieval, and manipulation of the data.

## Lab Overview

In the course of this lab, you will

- receive an overview of the customer requirements via direct communication with the customer (April 5);
- elicit additional requirements for the database via direct communication with the customer (April 8);
- prepare an initial design of the database for the project (April 10);
- extend the database design to model the customer requirements in full (April 11 - 15).

The lab assignment spans five lab periods, and has four distinct steps, each of which serves a different goal and yields a set of deliverables. This document specifies your assignment for each of the periods as well as final deliverable.

The overview of the lab structure is below.

**April 5, 2019: 12:00 - 1:00pm:** Groups are announced. Heather Liwanag joins us for an initial presentation. She will discuss the work of her lab, and the need for the elephant seal database application.

**April 8, 2019: 12:00 - 1:00pm:** Q&A session with the customer. Each team will have an allocated portion of time for questions. Questions/answers must be recorded and posted later.

**April 10, 2019: 12:00 - 1:00pm:** Initial database design. We will work on identifying major components of the database model and on discovering the key relationships between them.

**April 11 - April 15, 2019:** teams work on the full database design.

## Assignment

### April 5: Group Creation and Initial Presentation

We will spend the first few minutes of the lab period on April 5 on group creation. Group lineup will be announced by the instructor at the beginning of the lab. We will ensure that the group lineups are feasible (e.g., that we are not missing three people on a single 4-person team). Teams will contain three to four people.

Following team formation, our customer, Heather Liwanag will present an overview of your project, followed by a short, informal Q&A session. The main purpose of the presentation is to give you a good understanding of what the project entails. After the presentation, we will release the customer-supplied documentation.

These two documents, combined with the data, which will be released upon your completion of Lab 2 will serve as the basis for your project.

**Assignment.** Lab attendance is required for everyone - we will be organizing teams based on who comes to the lab. Please don't forget your nametags. During the presentation, you are welcome to take notes - those will come in handy later. Upon receiving customer documentation, study it in preparation for the April 9 lab session.

**Deliverables.** Each team shall come up with a name. Over the weekend, each team shall go over the documentation received from the customer, and will prepare, **in writing** a list of questions for the Q& A session with the customer.

### April 8: Question and Answer Session

The customer documentation released to you contains initial information about the database and the software application you will be building this quarter. The documentation reflects the information our customer chose to provide. This documentation, however, may be both (a) unclear and (b) incomplete. Your next step is to analyze the documentation, formulate additional questions to the customer, and solicit and record the resoponses.

**Preparation.** Before the *April 8* lab, each team shall study the materials provided by the customer. The goal of each team is to identify any aspects of the application domain, information about objects to be stored in the database, customer needs (w.r.t. the database), use cases to be implemented, and so on **that require additional clarification**. Each team shall prepare a list of questions that the team wants to ask the customer during the Q&A session. Note, that in addition to the customer, you may direct your questions at the instructor.

Please note: initial list of questions **must be prepared** by each team independently of other teams. However, once lists are prepared, I allow for teams to compare them (if you can manage this in the time between the April 5 and April 8 meetings). I expect that many questions will coincide or be similar (it is ok if multiple groups have similar/same questions). It is also ok for one group to come up with a question, no other group has asked — there is no need for other groups to add it to their lists, since each question only needs to be asked once, and all teams benefit from the answer.

Each team shall record a list of the questions as a googledoc.

**Q&A session.** Each team will be given about 10 minutes of time to ask their questions and receive answers from the customer and the instructor. We will set the order of questioning by a simple lottery at the beginning of the lab. I strongly suggest that each team sits and works together during the Q&A. Please note the following:

- Listen carefully to the questions other groups are asking. It is very likely, that some of your team's questions will be posed by teams who get to ask questions earlier.
- If your group runs out of questions yield your time to the next group.
- If your group has unanswered questions left after its Q&A period, wait until all groups ask their questions. If the question has still not been asked, you will have a chance to ask it at the end of the lab.
- You are certainly allowed to ask questions that are not on your list - some questions might occur to you as a result of things said during the Q&A session.

Each team is responsible for recording the answers to all questions its members have asked (whether those questions were prepared or were asked on the spot).

After each team had a chance to ask questions, if any additional questions remain, a free-form Q&A period will begin and will last until the end of the lab period. Any questions that went unasked, or unanswered during the Q&A session should be submitted in writing (via email) to the instructor. The questions will be forwarded to the customer, who, in turn, will provide answers prior to the Friday, April 12 lab.

**Deliverables.** After the Q&A session, all teams, in collaboration, shall create the project knowledge base. The project knowledge base will be represented this year as a single Googledoc file created by the instructor and shared (with eidting permissions granted) to all students in the class. The structure of the knowledge base is left up to the teams (generally speaking, you want to break the entire knowledge base by topic: e.g., group information about one type of data in one place, and information about another type of data — in another). A single question can appear in multiple places in the knowledge base.

The knowledge base is one of the two **main deliverables** for this lab. The initial state of the knowledge base shall be completed by the main Lab 2 due date. The intent is to have a useful and accessible resource for your further development.

## April 10: Initial Design

We will build the database model in two stages. First, *on April 10*, you will work out the overall database model, and will determine the most important components of the model. This will be done via a set of instructor-guided activities during the *April 10* lab period. After that, you will refine the model to incorporate specific requirements that the customer placed on the identified components.

**Assignment:** Each team shall fully participate in the model-building activities that occur during the *April 10* lab period. As the result of these activities, each team should produce a draft initial (simplified) Entity-Relationship model for the customer database.

**Deliverables:** The key deliverable of this stage is the simplified E-R model of the customer database. Each team will produce a draft of the model by the end of the *April 10* lab period.

The model description shall consist of the following:

1. List of entity sets. For each entity set specify its attributes, and the primary key. Identify weak (if any) entity sets, specify their owners.
2. List of relationship sets. For each relationship set, identify participating entity sets, list any attributes, identify its type (one-to-one, one-to-many, many-to-many), and whether it is the defining relationship set for a weak entity set.
3. An E-R diagram of the model.

After the lab period is over, create a text document representing your initial E-R model. The first two parts of the E-R model description should be plain (or formatted) text. The E-R diagram shall be drawn using some graphics editor or presentation program. Eventually it will be embedded in your document, but at this stage (since it is an early draft that you will modify many times) you can keep the E-R diagram in a separate file. I use `xfig` to create diagrams that are embedded into latex/PostScript files. I use MS PowerPoint to create diagrams taht are embedded into MS Word documents.

## April 11- April 15: Database Design

**Assignment.** The main task of this lab is for each team to prepare an initial full database design document. Teams shall use April 12 and April 15 lab periods, as well as any time outside the class necessary to complete the assignment to revise and enhance the initial E-R model.

During April 12 and Lab 15 lab sessions, each team will work with the instructor on the full database model. Most of the time will be devoted to individual team meetings, with the instructor joining each team for 5-10 minutes at a time to provide on-going feedback on the design and to answer any questions.

**Deliverable.** Each team is given time until **Monday, April 15, midnight** to finalize and submit the database model document.

The design document prepared and submitted by each team shall be a Googledoc **shared with the instructor** (see instructions below). At a minimum, it shall contain the following information:

- Team name, list of all team members (names, Cal Poly email addresses)
- List of entity sets for the proposed database.
- List of attributes for each proposed entity set.
- Identification of primary keys for each proposed entity set.
- Identification of all weak entity sets in the proposed database, and of their discriminating attributes.
- Identification of any other entity set constraints for the proposed entity sets.
- List of relationship sets for the proposed database. For each relationship set, the following must be indicated:
  - All participating enitity sets.
  - All clarifying relationship set attributes.
  - Type of the relationship set (one-to-one, one-to-many, many-to-many).
  - If the relationship set is an identifying one for a weak entity set.
- List of other relationship set constraints.
- List of any class hierarchies and/or aggregates in the database model.
- List of relationship sets (with all the information as above) associated with the aggregates.

- E-R diagram of the proposed database design embedded into your document.

The design document must be typeset (handwritten submissions will not be accepted). See comments above about the software to use to draw the E-R diagram. The E-R diagram shall contain all entity sets and relationship sets. It shall also contain all key attributes for all entity sets and all relationship set attributes. Other attributes can be omitted from the diagram for clarity. All constraints that can be shown on the diagram, shall be shown.

## Deliverables Overview and Submission Instructions

Overall, the lab has the following formal deliverables:

| No. | Deliverable | Due: | Softcopy | Hardcopy |
|---|---|---|---|---|
| 1 | List of questions for Q&A session | April 8, 12:00pm | Yes | Yes |
| 3 | Project knowledge base | April 15, midnight | Yes | No |
| 4. | Database Design document | April 15, midnight | Yes | No |

**Submission.** Bring a hardcopy of your team's question list to class on April 8 (make sure team name and student names are on it), and submit it at the end of the lab period.

The project knowledge base file is the responsibility of everyone. I will examine it on Tuesday, April 16. Because I create the Google document file you will be populating, no submission activities are needed (just do the work!)

Submit your database design document by sharing a Googledoc with dekhtyar@gmail.com (note: DO NOT share with my Cal Poly email). Name your database design document `Lab2-<TeamName>-design`, where `<TeamName>` is the name of your team.

**Please, keep the soft copies of all submitted documents. When you edit your Lab 2 submissions during Lab 3, make sure that you create separate documents for your Lab 3 deliverables, while still maintaining a copy of your original Lab 2 submission. This is IMPORTANT!**

**Grading.** I will read your design documents on **Tuesday, April 16**. I will provide feedback to each team during the lab period on **Wednesday, April 17**. Fixing your design will be part of your **Lab 3 assignment**.