

**Project Assignment 1**  
**(due February 3<sup>rd</sup> 2016, 4:00pm, in class—hard-copy please)**

**Reminders:**

- a. Out of 100 points. Contains 2 pages.
- b. Rough time-estimates: 1~2 hours.
- c. Please type your answers. Illegible handwriting may get no points, at the discretion of the grader. Only drawings may be hand-drawn, as long as they are neat and legible.
- d. There could be more than one correct answer. We shall accept them all.
- e. Whenever you are making an assumption, please state it clearly.
- f. Lead TA for the project: Sorour Amiri.

**Project Description: BiblioVT**

The goal of the class project is to implement a database system application, based on the data in [DBLP](#), which is a database of over 1.4 million publications in the computer science database and related communities. Related such systems are a big business now (under the name of digital libraries). For example, the [ACM digital library](#) provides a beautiful searchable index and retrievable repository of nearly all of the publications of ACM (the Association for Computing Machinery). We will work on a subset of data crawled from the web (we will provide you with this data later in the semester). The project includes the following activities spread over the entire semester:

- Form groups, think about the domain, and functionality of the system (this assignment)
- Based on the project description, design the ER diagram. Model the data stored in the database (Identify the entities, roles, relationships, constraints, etc.). (Project Assignment 2)
- Design, normalize, and perfect the relational database schema. (Project Assignment 2)
- Write the SQL commands to create the database. (Project Assignment 2)
- Finally and most importantly, populate the database and write the software needed to embed the database system in the application. (Project Assignment 3)

As you can see, Project Assignment 1 is a warm-up and Project Assignments 2 and 3 are the more involved ones. The end result should be a functioning application that runs on the web and that uses your database to allow useful functionality.

Form a team of 2-3 people and decide on a group name. All groups will work on the same overall project, but each group will have its own instance of the project data and will build a separate and unique interface to its instance. Please answer the following questions (total 100 points; some of the questions may need a longer answer than others, so try to be brief but explain what you have in mind):

- Q1.1 Please write name of the project group, names and pids of the students in the group. (Note: We will use your group name to give you access and create your instance on our PostgreSQL server.)
- Q1.2 The domain of your database application. (Note: Just write a one-two sentence description of the overall context or situation that your application is a part of. For example, if had the choice to work on a books database, you may describe your domain as "the books available at a library" or "the books sold by a bookstore.").
- Q1.3 What are the application specifications i.e., what functionality will your completed system provide? (Note: To give an analogy, suppose you were designing a web browser. Your answer might say that the browser will support visiting URLs, maintain a history of visited web pages, allow users to tag bookmarks, etc. Your answer will not need to go into the details of the positions of various menu items.)
- Q1.4 What aspects of the application will your system model? What will your system not model? (Note: For example, in the books domain, you can write something like "We will model books, authors, publishers, and printers" but not "reviews, bookstores and sales figures". Feel free to look up the DBLP XML/JSON format to get a sense of the information you need to model).
- Q1.5 Write some of the other (at least 6) "value-added" facilities your system could support? You can draw inspiration from how websites like [CiteUlike](#) or [Mendeley](#) use publication data, if you want. (Note: For example, if this was a movie info system like IMDB, such a facility could be a "recommender system" that makes movie recommendations for users based on buying trends. In other words, the recommender system is a facility that will be enabled by the presence of a database system).
- Q1.6 What is the role of each project member in the project? (Note: You don't have to "commit" to anything now. We want to see if any of the group members brings special talents/experiences to bear upon the project. For example, if one of you has worked in a digital store project, (s)he may help identify design choices from this point of view. If one of you has experience

in web-based software development, then that would be a good thing to mention).