CS 698F: Advanced Data Management

1. **Instructor:** Medha Atre
2. **Lectures:** Wed. Fri, 15:30-17:00 at KD 314
3. **Office hours:** by appointment
4. **Description:** With the growth of the web, problems surrounding "big data" have become central to many of the "industrial strength" solutions (which tackle the scale of the data that goes beyond the capabilities of prototypical solutions, e.g., billions of data items, such as graphs, images, videos, documents etc). A lot of this data is *semi-structured* (graphs, e.g., Facebook, Twitter, LinkedIn networks) or *unstructured* (videos, images, text documents etc), or a mixture of the semi-structured and unstructured data. Hence the challenges of storage and query processing over this data are different from the traditional relational database systems which focused on strictly structured data, even though many of the robust database features of storage and indexing are utilized as the core base of the new solutions.
5. **Components:** Following are the broad components of this course:
   a. This course will first take an overview of the traditional data management and query optimization techniques.
   b. Then it will focus on methods used for query processing over mainly "graph shaped" data, including centralized and distributed solutions (Hadoop, SPARK, and others).
   c. We will read research papers from the top conferences.
   d. The course will carry a large project component.
   e. The instructor will also introduce open (challenging) problems in both theoretical as well as practical domains of the "graph data management", to give directions of further research -- with a purely theoretical, or practical, or a combination focus -- especially useful for developing a postgraduate research plan.
6. **Pre-requisites:** UG level course in DBMS, knowledge of UG level data structures and algorithms. Good knowledge of programming.
7. **Grading:** TBA
8. **References:** No dedicated text book. Instructor covers course contents through lectures and supplementary material.