CS672 Spring 2016 - DIVA: Data Interaction and Visual Analytics

Instructor: James Abello  
abello at dimacs dot rutgers dot edu

abelloj at cs dot Rutgers dot edu

Pre-requisites:  CS512 or CS513 (Algorithms), CS539(Data Bases or equivalent). A class on Computer Graphics and Vision (CS428 or equivalent) may be useful but it is not required.

Useful Languages to know:  C/C++, Java, JavaScript, Python

Goal: To become proficient on the current major techniques and systems for algorithmic data analysis, exploration, visual interaction, and summarization. Students will complete a competitive group project that will incorporate all the facets of a software product development, namely:  Conceptualization, Data collection, Algorithm identification and Implementation, User Interface, and Evaluation.

*Projects will be judged by a faculty panel and interested industry sponsors.

Guiding evaluation principles will be: the “value” of the extracted information from the chosen data set, the methods and models used, and the final application Interactivity.

*Sample Project Categories (A Non Exhaustive Guide)

a. Similarity Search
b. Recommender Systems
c. Collaborative Filtering
d. Data Retrieval and Topic Waves
e. Prediction and Verification
f. Transaction Driven Data
g. Medical Imaging
h. Computer Aided Manufacturing
e. Apps for: Data Erasure, Nameless File Systems, Password Boxes, Phonetic Lyrics Search, Sentence Completion, ...

Data Sets:  Students will choose a Data set of their interest and devise representation methods that are conducive to efficient algorithmic exploration, Interactive Visualization, Analysis, summarization and sense making. The data sets used may be real or artificial. Some typical data sets that may be considered include: data feeds from Tweeter, YouTube, news streams, stocks, financial transactions, joke collections, movies, songs, online encyclopedias (Ex: OEIS, Algorithm and Software repositories ), transportation schedules, data analytics blogs, funding agencies, startups, computer science educational materials, internet of things, ...
Reference Materials

a. VisMaster-book


b. Selected papers from the literature on Algorithmic Analytics, Visualization and Computer Human Interaction.

Class Syllabus

Topics: Visual Analytics History, building blocks and inherent scientific challenges, Data Management for Visual Analytics, Data with Spatial and Temporal Components; Infra Structural and Language Issues (Hadoop, MapReduce, R, Python); Evaluation Methodologies and Challenges.

Time Line

Phase 1 (Background - Week 1 and 2). During the first two weeks students will be exposed to the fundamental principles of Data Analytics and Visual interaction as described in the VisMaster-book.

Phase 2 (FUN Project Selection - Week 3). During week 3, each student will present to the class and faculty a project conceptualization with a feasible plan of completion. One third of the projects will be selected for continuation. Those students whose projects get selected will become the project leaders. Student Leaders will select two non-leader partners that are willing to commit to the successful project completion. Faculty will choose those Feasible, Useful, and Novel projects they are willing to supervise and sponsor.

Phase 3 (Project Prototype Progress Report - Week 4 - 5 - 6). Development and Evaluation of an operational Prototype. Faculty will evaluate project progress and assess the feasibility of project completion by the end of the semester. Only those projects judged as feasible will be allowed to move forward. Those students whose projects do not get selected will be assigned to become testers and project writers of those projects moving forward.

Phase 4 (Pre Final Defense - Week 7, 8, 9, 10) Faculty sponsors will meet alternatively every week with sponsored projects to monitor their progress. During week 10 each project will have a pre final defense presentation. One third of the projects will be chosen for a final gala presentation (on week 14) in front of the class and a faculty/industry jury.

Phase 5 (Gala Presentation - Week 14) The best five projects will be selected, project DIVA awards will be distributed, and formally recorded in the planned MSCS Wall of Fame.