Data Management Wrap-up
Wed. May 4 - Wrapup

Joshua Hurt - Carnot Provider, see https://seelio.com/w/2cmu/carnot-engine?student=joshuahurt

Selected Project Presentations:

- Option F:
  Present together:
  Falik Roy Kloc Michael E4
  https://github.com/RoyFalik/S16E4
  Evans Caelan Milligan Seth Kasukhela Akshay F2
  https://github.com/caelanevans/S16F2_FinalProject

- Inference:
  Present together:
  Ellis Brian Crous Deidre E6
  https://github.com/kitchWWW/DMproject
  Ababao Rainier Ginther Nick Yen Michael B3
  https://github.com/nginth/carnot-owl

- StarDog:
  Estes Martin Yap Shannon Lim Jeremy C6
  https://github.com/phalax4/CarnotKE/tree/stardog

Wrap-up:

- Class Notes
CS411: Intro to Upper Division Electives

Fall, 2015

Dr. Philip E. Cannata - Oracle and UT Adjunct Professor
CS347 – Data Management

CS347 will prepare you for this.
Growth of Data vs. Growth of Data Analysts

Stored Data accumulating at 28% annual growth rate
Data Analysts in workforce growing at 5.7% growth rate

CS347 will prepare you for this.
Talent Remains Key Big Data Challenge
by BOB VIOLINO
MAR 18, 2015 5:37pm ET

Sixty-one percent of IT leaders expect spending on big data initiatives to increase, while only 5% expect decreases. The challenge: Finding the right big data talent to fulfill those initiatives, according to a recent survey by TEKsystems, a provider of IT staffing and talent management services.

“Nearly 60% of the respondents are confident that their IT department can satisfy big data demands of the business and 14% are not confident,” the report states. “Despite IT leaders planning to increase, the confidence level in their department’s ability to meet big data demands in comparison to broader IT initiatives is lower.”

About two-thirds of the IT executives rank big data architects as the most difficult role to fill. Data scientists (48%) and data modelers (43%) round out the top three most difficult positions to fill. More technical big data positions are ranked less difficult to fill.

CS347 will prepare you for these.
CS347 – Data Management

You will learn – SQL, Data Modeling, Web/Database Application Development, the Oracle DBMS Architecture, Transaction Processing, Indexing, NoSQL and Graph Databases, Normalization, and SQL Query Optimization.
Possible Opportunities:

- Continue Option F Apex Development
- Continue SIM Development on top of NoSQL and/or RDF/OWL
- Continue Inference Development with OWL or standalone reasoners.
- Continue OpenCyc Integration
- Continue NoSQL Storage Development
- Investigate integrating Carnot with the Swift Programming Language.

All of these would make very good cs370 projects but this is not required.

See: https://www.cs.utexas.edu/undergraduate-program/undergraduate-research/course-credit-your-research

May work with (SIM Architect)
Hey Dr. Cannata,

I hope you're doing well. I've been wanting to write for some time now to thank you for all I learned in your class and for the skills I developed while doing research under you. I'm several months into my job, and I've certainly made good use of the SQL skills; in fact, many of my co-workers, who joined the company at the same time as me, seem to think I'm quite good with it and come to me with there SQL inquiries. I have you to thank for that. :) Perhaps even more importantly, doing the research under you taught me how building software can be difficult, and how managing people can be just as difficult.

The UTBC work gave me an appreciation for process and methodology in creating software, and it challenged me to learn how to organize, manage, and motivate people. I also learned self-discipline and time management skills to keep up with the responsibilities. All this has been invaluable since I graduated. I wanted to let you know you've made a large impact on me and that I feel I started my career with a lot of knowledge that would otherwise have taken me some time to learn and perhaps even longer to fully appreciate. I didn't realize the journey I'd go on when registering for your Data Management class fall of 2014, but it definitely ended up being one of the best decisions I made in my college career :) Thank you again for the tremendous positive impact you've had on my life and my career.

See: https://www.cs.utexas.edu/undergraduate-program/undergraduate-research/course-credit-your-research
Final Review

- Review the Midterm, Homework (especially homework 7, 8, 9 and 10) and Quizzes (especially Quizzes 5, 6, 7, and 8)
- SQL on the class website
- Normalization
- Query Execution Plans

~ 30%
Chia-Chen Hsu and Brandon Hollowell
The current grades were calculated out of 80 points as follows:

\[(20*(\text{Midterm}/150) + (15*(\text{HW Average without Lowest HW}/10)) + (20*(\text{Quiz Average without Lowest Quiz}/10)) + (15*(\text{Project}/10)) + (10*(\text{Active Participation}/95))\].

Notice - the Active Participation grade was divided by 95 instead of 100 to compensate for hiccups in TopHat.
Final Grade = max(Current Grade, Current Grade + Final Exam)

(i.e., the Final is **optional**, it can’t hurt you, it can only help – Thank you for your effort this semester especially on the **projects**.)

Dr. Philip Cannata

Have a Great Summer!