Overview of Course Projects

Abdu Alawini
University of Illinois at Urbana-Champaign
CS411: Database Systems
January 30, 2019
Project Tutorial Schedule

• Overview of Project Track 1 and Track 2 (today!)
  • Timeline & Requirements
  • Tips
  • Project Showcasing

• Web Programming Basic (2/4 7-8:20PM, SC1404)
  • Development process
  • Frontend technologies

• Backend Development Tutorial (2/8 5-6:20, SC1404)
Project Tracks

**Track 1**

Build a relational database-centric application
- Web or mobile app
- Basic & advanced functions
- Both 3- and 4-credit students

**Track 2**

Research survey on DBMS
- Paper (8-10 pages)
- Presentation (~ 10 slides)
- 4-credit students only
What You Get Out of Project Track 1

• Use what you learn in class
• Come up with good ideas
  • Brainstorm
  • Survey
  • Pick the best idea
• Collaborate and coordinate with teammates
  • Task Assignment
  • Take on challenges
  • To grow is to do what you could not.
## Track 1 Timeline

<table>
<thead>
<tr>
<th>Stage</th>
<th>Due Date</th>
<th>Milestone</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/31</td>
<td>Team Formation</td>
</tr>
<tr>
<td>1 (4%)</td>
<td>2/8</td>
<td>Project Functions and ER Design</td>
</tr>
<tr>
<td>2 (3%)</td>
<td>2/19</td>
<td>Development Plan and DB Schema</td>
</tr>
<tr>
<td>3 (3%)</td>
<td>3/5</td>
<td>Development Environment Setup</td>
</tr>
<tr>
<td>4 (30%)</td>
<td>3/28 - 30/3</td>
<td>Midterm Demo</td>
</tr>
<tr>
<td>5 (60%)</td>
<td>4/17 – 4/19</td>
<td>Final Demo</td>
</tr>
</tbody>
</table>
0. Team Formation

• Size constraint: 3-4 People
• No constraint: undergrad/grad
• Tips
  • Diversity
    • UI, frontend
    • Server, backend
    • Data acquisition & pre-processing
    • Data analytics & visualization
  • Leadership
1. Project Functions and ER Design

Design Principles

• Purposeful
  • Does it solve an actual problem, meet a real need?
• Unique
  • Is it a mere clone of an existing product?
• Innovation:
  • UX/UI (think about what you often complain as a user)
  • Engineering: performance, scalability, reliability
• Real
  • Data source: web crawling, user contribution
• (Realistic)
1. (cont.) Advanced Functions

Questions to ask yourselves:

• Do you actually implement the advanced functions?
  • Do you get the result via straightforward SQL queries or API calls to libraries/services? No!
  • You don't have to build it from scratch, but if you are to use an existing library/service, your effort has to involve a substantial data preprocessing and functional extension to achieve your goal.

• Is it challenging?
  • "We implemented a recommendation engine, using cosine distance." No (unless you have advanced feature vectors).
  • Document the difficulties and trickiness in detail.
    • Even if you tried something ambitious, but it didn't work great, still document the process.
    • There is an "effort" component in your grade!
2. Development Plan and DB Schema

- Project Timeline with Milestones
- Division of Responsibilities
- Source of Data
- Development Stack
  - Client
  - Server
  - DBMS
  - (Optional) Frameworks/libraries for advanced functions
- DB Schema (from ER Design)
  - Constraints, Views,...
3. Development Environment Setup

• Set up your source code repository for version control and collaboration
• Set up your stack locally
• Set up your stack remotely
  • Engineering Web Hosting via cPanel, OR
  • Commercial Options
4. Midterm Demo

Via your own Admin UI

• CRUD
• Two Advanced Queries
  • Subqueries
  • GROUP BY, Aggregate functions, HAVING
  • If conditions (CASE)
  • Set operations (e.g. EXCEPT)
5. Final Demo

• Meet minimum data requirement
• Complete demo video
• Project report
• Live demo of advanced functions
5% Bonus Points for the 4 Best Projects

https://wiki.illinois.edu/wiki/display/cs411sp19/Hall+of+Fame
Project Track 2 -- 4-credit

- 4-credit students do both:
  - project 1 (DB application)
  - project 2 (survey on one of cutting edge DB research topics)
Goal of Writing Survey Paper

- Expert knowledge of the **state-of-the-art**
  - Don’t reinvent the wheel
  - Really understand the connections between methods
- Build a special angle on a topic and identify prospective research problems
- Improve your knowledge on **critical** thinking skills and analytical work with a lot of scientific literature
List of Topics (must choose one of them)

• Main-memory database systems
• Interactive visualization systems
• Columnar databases
• Auto-tuning of database systems
• Approximate query processing
• Stream processing systems
• Graph database systems
• Semi-structured database systems
• OLAP Systems
• Multi-query optimization
• Parallel query optimization
How to get started?

- Read 5-10 surveys from ACM Computing Surveys to get the idea of material organization
  - What are the typical sections?
  - How long it should be?
  - How to present ideas?
- Select a few interesting topics and:
  - Search for existing surveys and assess their quality
  - Estimate the amount of papers published
  - Is topic trendy, popular, important, impactful?
  - Pick the one that you will write about!
Have Materials and a Plan

• Find top researchers and research groups, conferences, journals, demo systems on selected topics
• Skim through highly cited, recent articles of prominent researchers to get a good approximation of other good works
• Write a plan with the answers to questions:
  • Why your survey will be important? What’s new? What perspective you want to open?
  • Set deadlines for yourself and follow them
  • Add key citations at the end of the proposal
Paper Sources

• Top Database Conferences
  • VLDB, SIGMOD, SIGKDD, etc.

• Top Database Journals
  • TODS, TOIS, TKDE, etc.

• Search Engines
  • Google Scholar
    • https://scholar.google.com/
  • DBLP
    • http://dblp.uni-trier.de/
Typical Timeline

• 1-2 weeks for paper selection, initial reading, and survey idea formalization
• 2-4 weeks writing an initial draft
• 2-4 weeks for the final editing before submission

• Think Big! You really can submit it to the journal!
Track 2 Deadlines

• Stage 0 (due: 12:00 pm, 1/31) **Survey teams formation**
  • In addition to project 1 team
  • Create Team page in Wiki

• Stage 1 (due: 12:00 pm, 2/14) **Survey proposal submission**
  • Title, Author, Abstract, Proposal (most important part).
  • Upload a PDF file to your team page

• Stage 2 (due: 12:00 pm, 4/9) **Survey final draft submission**
  • 10 slides presentation
  • Final draft: 8 - 10 pages
More Information & Resources

- **Project Track 1**
  - [https://wiki.illinois.edu/wiki/display/cs411sp19/Project+Track+1](https://wiki.illinois.edu/wiki/display/cs411sp19/Project+Track+1)

- **Project Track 2**
  - [https://wiki.illinois.edu/wiki/display/cs411sp19/Project+Track+2](https://wiki.illinois.edu/wiki/display/cs411sp19/Project+Track+2)