EECS 440 System Design of a Search Engine
Supporting Statement

The proposed course will address the currently unfilled need for a course which replicates the experience of working on the startup team of software developers ("devs") on a complex system design project.   Over the semester, each team of six students must plan and then build a working internet search engine, completely from scratch in C++.

The project is both "close to the metal" and one that integrates a multiplicity of topics in computer science, including operating systems, networking, compilers, data structures, algorithms and information retrieval. It exposes students to the practical applications of multithreading and locks, sockets, mapped files and data structures on disk, and top-down recursive descent compiling, making this an ideal student MDE.

As on a real team, students are forced to negotiate and cooperate to make decisions on product goals, architecture, and assignments and responsibilities, often under uncertainty.

Roughly 30% the student's final grade is based on their ability to work with others and contribute to their team's success, e.g., as defined in ABET Outcome 5, and is measured both objectively, e.g., by lines of code and other contributions compared to others on the team and in the class, and subjectively, e.g., by survey of the student's teammates.

Past offerings

The course has been taught three times as an experimental EECS 398 course, in W18 as System Design in C++ and in W19 and F19 as System Design of a Search Engine, better reflecting what the course is really about.

	W18	W19	F19	Total
Enrollment	26	61	58	145
Student teams	5	10	10	25

Student evals have shown both interest and improvement as I've added content, better organized the material and responded to suggestions, e.g., to drop the discussion of the Windows API to focus on Linux. Students are clearly satisfied with the outcomes, especially, their understanding of and interest in the subject matter, and their confidence and ability to solve real world engineering problems.

Student evals

	W18 Final	W19 Midterm	W19 Final	F19 Midterm	F19 Final
Q1. Overall, this is an excellent course	2.70	3.6	4.2	4.6	4.0
Q2. Overall, Nicole Hamilton is an excellent teacher.	3.75	3.4	4.3	4.6	4.1
Q3. I learned / am learning a great deal from this course.	3.38	3.9	--	4.7	--
Q4. I had a strong desire to take this course.	4.67	4.7	4.4	4.8	4.6
Q61. Prerequisites provided adequate preparation for the course.	3.17	--	3.7	--	4.0
Q1631. This course advanced my understanding of the subject matter.	--	4.4	4.8	4.6	4.6
Q1632. My interest in the subject has increased because of this course.	--	4.8	4.8	4.6	4.6
Q1769. I developed confidence in my abilities as an engineer.	--	--	4.2	--	4.6
Q1770. I developed the ability to solve real world engineering problems.	--	--	4.7	--	4.7

The falloff in Q1 and Q2 scores between the midterm and final evals in F19 is probably deserved. I need to cover most of the essential topics beyond 281 needed to build a whole search engine by about midway through the course if the students are to have enough time left to actually do it. I'm doing well on that. But then I need to fill my remaining lectures with content that's relevant and interesting but not essential to completing the project. That still needs work.

Course development

In W18, I started with little more than a syllabus and struggled to create two hours of original lecture material every two days. And the evals reflect that. The big risk was that the task was impossible for a small team of students in one semester. But while some teams did better than others, every team succeeded in delivering a working engine.

In W19, I renamed the course to communicate more clearly what it was about and focused on organizing the material to get teams crawling the web sooner, added lecture content and, with the help of an IA, created some autograded homework and labs.

In F19, again with the help of my IAs, I've continued to add labs, homework and lecture content and to improve the sequence of topics. Responding to student feedback, I dropped the Windows discussion. I also arranged for each student to receive a free $100 student Amazon Web Service (AWS) account so they can build their engines on the cloud.

I also introduced a simple Google survey in which students were asked to rate their teammates' contributions and effectiveness as members or leaders of a team, in support the ABET objective. Actual example individual and team reports, but with fictious names, are attached.

Student engines

As the course has improved, the students' indexes have steadily grown. Their engines are typically crawling on multiple machines on the cloud with gigabit internet connections and anywhere from 1,000 to over 14,000 threads. (In W19, as I reorganized the material to get teams crawling faster, they got better faster than I expected; one team accidentally DOS'ed the Duke University Law School Registrar's site during reg week. We now spend much more time discussing politeness before unleashing such awesome power.)

Index sizes in documents	W18	W19	F19
High	13.4M	150M	586M
Median	8,000	4.65M	32.12M
Low	4,816	1.9M	1.2M

Engines in lines of code	W18	W19	F19
High	14,271	26,887	17,079
Median	5,300	13,000	11,973
Low	4,170	9,414	4,366

Individual lines of code	W18	W19	F19
High	4,096	8,646	7,826
Median	1,006	1,750	1,571
Low	0	444	120

Reflections on F19

I've now run 25 teams through this class. I know the class is getting better, but a surprising result in F19 was that for the first time, I had 3 teams decide to take incompletes and finish in January because they weren't ready to demo.

In my final meetings with each team, students frequently cited fall semester distractions like recruiting and football, but it was also apparent that many seemed to have falsely regarded the Thanksgiving break as the halfway point in the semester, which it's not, especially this year, when it came late. (Thanksgiving recess was 85 days into F19; winter vacation is only 52 days into W20.) It's possible the course is better scheduled for the winter.

But it's also possible that I contributed to the problem by making the course "better". As I added content and structure, especially, labs and homework, I intended for these to be helpful with important

concepts and students consistently remarked that they were. But the result may have been to distract energy and attention from the project.

In my final reviews with the individual teams, then later in their evals, many students agreed with a change I propose to make next time, which is to announce the homework and lab assignments all at once, early in the semester, and make them all group assignments. I can do that now that they all exist.

In previous semesters, students suggested I prereq EECS 482 Operating Systems. But I wanted the course to be accessible to sophomores, so I resolved to improve my treatment of these topics. I was gratified that the course evals this time indicated agreement that the course is fully accessible to students with only EECS 281.

The peer evaluation surveys worked well. I got a high response rate and the results were helpful, especially in the insights it gave me into the team dynamics. For example, on one team, one student contributed a huge share of the total lines of code in their project which looked pretty good on the numbers alone. But his teammates rated him poorly on quality, reliability and cooperativeness; this led to discovery in discussion with the group that he had habit of dumping big lumps of buggy code on them and then disappearing from contact.

I also had one team that tried to game the system by giving each other 5 out 5 on every attribute. This was one of the teams that took an incomplete, so they were definitely not all 5 out 5 on everything. Going forward, I intend to instruct students that they are on their honor not to collude or discuss their peer evals with each other and that if I detect gaming, I may ignore their surveys in determining their grades.

Several students remarked in the evals and in discussions that if they didn't come in with friends in the course, that it was difficult to find a group. Not wanting to be rude by walking away to look for a better match, they teamed with the first people they talked to. This sometimes led to dysfunctional teams, the most common problem being that they couldn't get everyone in one room at the same time to get stuff done. At the same time, students were overwhelmingly in favor of being able to choose their own teams; they did not want to be assigned.

I hope to somewhat ameliorate this problem of finding a compatible group next time by creating a shared Google spreadsheet where students may fill in contact info, best times and places (e.g., North vs. Central) for meetings, and any other relevant info they care. I may also introduce some "speed dating" exercises into the first couple of lectures.

Looking ahead

My hopes for this class are:

1. To continue developing it toward what I hope will be a complete and largely turnkey course with content and staff that other instructors might also enjoy teaching or working on. I think I'm getting close.

2. For the students, to deliver an experience of success and satisfaction working on a small team building a complex system that they can explain to family and friends and to recruiters and hiring managers.

Example individual report.

```
Hello, Kara Danvers (kara).

Below is an individual report of your EECS 398 peer evaluations.
Results range from 1 (Strongly Disagree) to 5 (Strongly Agree).

                                                      Your Rating   Team Average   Class Average
-------------------------------------------------     ------------  -------------  --------------
Sets High Standards And Delivers High Quality.            2.6           4.2             4.2
Treats Others With Respect And Respects Their Boundaries. 4.6           4.4             4.6
Listens To And Encourages Others' Ideas.                  4.2           4.3             4.5
Is Cooperative.                                           4.2           4.4             4.5
Is Reliable, Shows Up, Answers Emails, Delivers On Time.  1.6           4.1             4.2
Contributed Lots Of Helpful Ideas.                        3.4           4.2             4.2
Responds Constructively To Suggestions Or Criticism.      4.4           4.3             4.4
Really Stepped Up To Do Their Share Of The Work.          2.8           4.3             4.2
Is An Excellent Teammate.                                 2.8           4.3             4.3
```

Example team report.

```
SUMMARY FOR SUPERHEROS

This team member sets high standards and delivers high quality.

Name                      Uniqname     Rating    Team Average    Class Average
------------------------  ----------   --------  --------------  ----------------
Arthur Curry              arthur         4.2           3.8             4.2
Barry Allen               barry          4.8           3.8             4.2
Bruce Wayne               bruce          4.4           3.8             4.2
Clark Kent                clark          4.6           3.8             4.2
Hal Jordan                hal            4.6           3.8             4.2
Kara Danvers              kara           2.6           3.8             4.2


This team member treats others with respect and respects their boundaries.

Name                      Uniqname     Rating    Team Average    Class Average
------------------------  ----------   --------  --------------  ----------------
Arthur Curry              arthur         4.6           4.6             4.6
Barry Allen               barry          4.6           4.6             4.6
Bruce Wayne               bruce          3.2           4.6             4.6
Clark Kent                clark          4.8           4.6             4.6
Hal Jordan                hal            4.6           4.6             4.6
Kara Danvers              kara           4.6           4.6             4.6


This team member listens to and encourages others' ideas.

Name                      Uniqname     Rating    Team Average    Class Average
------------------------  ----------   --------  --------------  ----------------
Arthur Curry              arthur         4.6            4              4.5
Barry Allen               barry          4.6            4              4.5
Bruce Wayne               bruce          3              4              4.5
Clark Kent                clark          4.8            4              4.5
Hal Jordan                hal            4.4            4              4.5
Kara Danvers              kara           4.2            4              4.5
```

This team member is cooperative.

Name	Uniqname	Rating	Team Average	Class Average
Arthur Curry	arthur	4.6	4.1	4.5
Barry Allen	barry	4.8	4.1	4.5
Bruce Wayne	bruce	3.6	4.1	4.5
Clark Kent	clark	4.6	4.1	4.5
Hal Jordan	hal	4.4	4.1	4.5
Kara Danvers	kara	4.2	4.1	4.5

This team member is reliable, shows up, answers emails, delivers on time.

Name	Uniqname	Rating	Team Average	Class Average
Arthur Curry	arthur	4.8	4.1	4.2
Barry Allen	barry	4.6	4.1	4.2
Bruce Wayne	bruce	4.8	4.1	4.2
Clark Kent	clark	4.6	4.1	4.2
Hal Jordan	hal	4.2	4.1	4.2
Kara Danvers	kara	1.6	4.1	4.2

This team member contributed lots of helpful ideas.

Name	Uniqname	Rating	Team Average	Class Average
Arthur Curry	arthur	4.2	3.8	4.2
Barry Allen	barry	4.6	3.8	4.2
Bruce Wayne	bruce	4.4	3.8	4.2
Clark Kent	clark	4.2	3.8	4.2
Hal Jordan	hal	4.4	3.8	4.2
Kara Danvers	kara	3.4	3.8	4.2

This team member responds constructively to suggestions or criticism.

Name	Uniqname	Rating	Team Average	Class Average
Arthur Curry	arthur	4.4	4.2	4.4
Barry Allen	barry	4.6	4.2	4.4
Bruce Wayne	bruce	4	4.2	4.4
Clark Kent	clark	4.4	4.2	4.4
Hal Jordan	hal	4.2	4.2	4.4
Kara Danvers	kara	4.4	4.2	4.4

This team member really stepped up to do their share of the work.

Name	Uniqname	Rating	Team Average	Class Average
Arthur Curry	arthur	4.6	3.7	4.2
Barry Allen	barry	4.8	3.7	4.2
Bruce Wayne	bruce	4.6	3.7	4.2
Clark Kent	clark	4.6	3.7	4.2
Hal Jordan	hal	4.6	3.7	4.2
Kara Danvers	kara	2.8	3.7	4.2

Overall, this team member is an excellent teammate.

Name	Uniqname	Rating	Team Average	Class Average
Arthur Curry	arthur	4.4	3.9	4.3
Barry Allen	barry	4.8	3.9	4.3
Bruce Wayne	bruce	4	3.9	4.3
Clark Kent	clark	5	3.9	4.3
Hal Jordan	hal	4.6	3.9	4.3
Kara Danvers	kara	2.8	3.9	4.3