**2013-2014 UC Berkeley Master of Engineering**

**Capstone Project Proposal**

**Overview:**

The Capstone Project, a 5-unit Master of Engineering course requirement, integrates core leadership coursework with a student’s engineering concentration. Capstone Project teams range from three to ﬁve students, drawn from the cross-disciplinary engineering cohort, to apply diverse knowledge and skills to actual industry problems, identiﬁed by faculty or industry partners. The Fung Institute for Engineering Leadership within the College of Engineering provides capstone cohort support and curriculum integration.

**Capstone Sponsor Information:**

Please read the following instructions and requirements before submitting your proposal. In order to be considered, this document must be completed in full. By submitting this proposal, you agree to its inclusion in the *UC Berkeley Master of Engineering Capstone Project Portfolio* for the 2013-14 Academic Year. Use of links, diagrams and images to illustrate your project is encouraged. Example projects can be found here: <http://funginstitute.berkeley.edu/programs/capstone-projects>

**Timeline for submission and important deadlines:**

|  |  |  |
| --- | --- | --- |
| **Year** | **Dates** | **Activity** |
| **2013** | **March** | **Capstone Project Call for Proposals** |
|  | **By April 1** | Submit a one-sentence description of your project idea. |
|  | **By May 1** | **Full Project Proposals due**  Please use the proposal form supplied. |
|  | **May-July** | **Proposal Review –** screening for skill set and objective fit with incoming M.Eng. class**.** |
|  | **July-August** | **Student Project Exploration**  Industry advisors should be available for questions and interview screening of students during this time. |
|  | **August 12-31** | **Capstone Team Selection Process, Sponsor and Faculty office hours** |
|  | **September 1-12** | **Capstone Final Match:** Notification no later than Sept 12 |
|  | **Early December** | **Fall Student Poster Session** |
| **2014** | **Early May** | **Spring Student Poster Session** |
|  | **May 1-17** | **Final Student Presentations and Deliverables** to Industry & Faculty Advisors |

If selected for the 2013-2014 Capstone Project Portfolio you will be responsible for sponsoring and adhering to the terms you outline below. **As the Capstone Sponsor, please *initial* the following requirements by which you are agreeing to the following:**

Top of Form

Provide a point person from your organization to advise the capstone team on a regular basis and throughout the whole duration of the project

Supply all necessary tools, software, and/or data necessary to do the project in a timely manner

Ensure the project has achievable deliverables that fit into a 9-month timeframe

Provide clear objectives for both the technical and business-related challenges of the

project

By signing below you are indicating that you completed this form to the best of your knowledge and are agreeing to all the requirements of UC Berkeley’s Capstone Project Program as listed above.

We look forward to working with you!

Name: Title:

Email: Phone:

Signature or Initials:

Date:

*Questions?* Contact Beth Hoch hoch@berkeley.edu or 510-664-4587

**Proposal Form (please complete all sections):**

|  |  |
| --- | --- |
| **Project Title** | **Big Data; Business Models and Open Innovation** |
| **Industry Partner**  Company Name, Department, and Website | Advisory Board includes leaders from IP Checkup, RPX, Former Commissioner of US Patent Office David Kappos ([Cravath, Swaine & Moore](http://en.wikipedia.org/wiki/Cravath,_Swaine_%26_Moore)) |
| **Problem**  (Describe the industry problem your project addresses in 100 words or less.) | Big Data is a popular buzzword, but business models are not clear yet; using the Fung Institute US Patent database, we will explore a variety of issues and applications, including big data business models, machine learning (disambiguation), text analysis, open innovation and tournaments, and academic technology transfer business models. |
| **Technical Challenge**  (Highlight the technical challenge of the problem in 100 words or less) | Variety of challenges including:   1. Machine learning 2. Disambiguation 3. Visualization 4. Text Analysis 5. Business Models |
| **Objective**  (In 100 words or less, use bullet format and ensure objective is practical for a 9 month project) | To integrate business model and technical aspects of big data analysis. |
| **Project Illustration (Optional)**  Include websites, videos, diagrams or images to help students understand your project | CODE AND DATA FROM THE UC BERKELEY FUNG INSTITUTE  The Fung Institute generates and distributes a variety of data software useful for innovation research. Patents and patenting figure prominently in such research. To promote transparency and reproducibility, the Fung Institute provides software source code and the raw data used for research purposes via the Github social coding site. The data and tools are free to download and use.  <https://github.com/funginstitute/downloads>  **INVENTOR MOBILITY MAPS FOR THE UNITED STATES, 1975-2010**  These movies illustrate the mobility of patented U.S. inventors, based on different addresses and their date of patent application.  The arc appears in the year that the latter patent was applied for.  Of note is the substantial increase in inventor migration to California over the observed years.  Michigan's noncompete brain drain can also be seen, in the increased emigration following the 1985 passage of the Michigan Antitrust Reform Act.  Latter years appear light due to the lag between patent application and grant.  Disambiguation of inventors based on R. Lai and D. Doolin, A. D’Amour, A. Yu, Y. Sun, V. Torvik. “Disambiguation and co-authorship networks of the U.S. Patent Inventor Database,” working paper, Fung Institute. For econometric evidence of Michigan brain drain, see M. Marx and J. Singh, “Regional Disadvantage: Non-competes and Brain Drain," working paper, Fung Institute.  Visualizations done by Guan-Cheng Li.  [Inventor Mobility Maps for the US, 1975-2010](http://funglab.berkeley.edu/mobility/)  **DISAMBIGUATION AND CO-AUTHORSHIP NETWORKS OF THE U.S. PATENT INVENTOR DATABASE**  Lai, R., D’Amour, A., Yu, A., Sun, Y., Torvik, V. and Fleming, L. "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database."  Using a Bayesian supervised learning approach, we have disambiguated all inventor names from the U.S. utility patent database, from 1975 to the end of 2010, in order to uniquely identify inventor careers. This paper provides a non-technical overview of the disambiguation methods, assesses their accuracy, characterizes the resulting dataset by basic descriptive statistics, calculates network measures based on co-authorship and collaborative variables, and provides illustrative examples of networks and potential research questions. The dataset is available at the Patent Network Dataverse (http://dvn.iq.harvard.edu/dvn/dv/patent); this website also provides an interface for sub-setting network illustration tools, an overview of the technical algorithms, and pointers to the data, code, and documentation.  Keywords: disambiguation, patents, networks, inventors, careers  [Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database.pdf](http://funginstitute.berkeley.edu/sites/default/files/Disambiguation%20and%20Co-authorship%20Networks%20of%20the%20U.S.%20Patent%20Inventor%20Database.pdf)  **PATENT DISCOVERY PROJECT**  The Fung Institute is developing several tools to aid in the discovery of patents around a common technology or application. Leveraging advances in machine learning and automated content analysis, the Patent Discovery Project scans a set of hand-coded "training patents" and then finds a range of similar patents based on the similarity of patent abstracts and patent claims. These tools can be used by researchers to identify a broad population of patents for study from a smaller set of known cases, significantly reducing the search time of researchers in finding related patents.  The first application of the Patent Discovery Project is a tool that investigates the emergence of Clean Technologies in the U.S. Patent data. After identifying patents (as described above), the exploration tool below  allows users to filter through and browse patents by cluster, technology, and type of patent assignee.  [Clean Tech Patent Discovery Tool](http://funglab.berkeley.edu/cleantechx/) |
| **Open or Closed Model – Please check one:**  Open Model (Public collaborative and may use university lab equipment) or Closed Model (Virtual internship, private, with faculty liaison)  \* Please list the necessary equipment, software or data that is needed and will be provided to the team. | Could be open or closed, the team will decide based on project scope.  **Tools and Equipment that will be provided include:**  Any needed software or hardware necessary to complete the project. Reimbursement for travel to client sites, if applicable. |
| **Ideal Team Size**  (We prefer teams of 4 students, unless otherwise specified) | 3 to 4 students |
| **Departments Accepted**  (Choose from CEE, EECS, IEOR, ME, MSE, NE. Indicate ideal team makeup and technical concentrations desired, i.e.  “1 CEE ; 1 EECS; 2 IEOR”) | EECS = Electrical Engineering & Computer Science  IEOR = Industrial Engineering & Operations Research |
| **Specific Skills Required**  (i.e. *C/C++/C#, Python ,CAD, Robot Kinematics, MATLAB, Excel Financial Modeling, etc.*)  The more detail provided here the better team match you will receive. | Each team member needs ONE of the following skills:  1. Programming background including at least one of the following: Database, SQL experience, machine learning, BASIC, Python.  2. Solid statistical background (including inference and machine learning) |
| **Coursework**  (Indicate any recommended/required prerequisite/co-requisite classes) | Recommended:  Python Boot Camp  IEOR Database Course (Ken Goldberg)  STATS/EECS Machine Learning/Visualization courses |
| **Industry Advisor(s)**  **Name, Email, Phone Number**  \*If this is a closed model an Industry Point Person from your organization is required for the duration of the project and must be available to advise the team on a regular basis and provide all necessary resources | TBD, Advisory Board includes leaders from IP Checkup, RPX, Former Commissioner of US Patent Office David Kappos ([Cravath, Swaine & Moore](http://en.wikipedia.org/wiki/Cravath,_Swaine_%26_Moore)) |
| **Faculty Advisor(s) or Academic Liaison**  **Name, Department, and Email**  \*If this is an open model the Faculty Advisor or Academic Liaison is the primary party responsible for the advising and guidance of the capstone team, including providing all the necessary resources | Professor Lee Fleming  [lfleming@ieor.berkeley.edu](mailto:lfleming@ieor.berkeley.edu)   * Faculty Director, Coleman Fung Institute of Engineering Leadership * Coleman Fung Chair of Engineering Leadership * Industrial Engineering and Operations Research, College of Engineering * Management of Organizations, Haas School of Business * Department Editor, Entrepreneurship and Innovation, *Management Science*, a journal of the Institute for Operations Research and the Management Sciences |