



WPI

DATA SCIENCE **OPEN CALL FOR GQP PROPOSALS!**

Data Science Graduate Qualifying Project Experience:

Students in WPI's Data Science program are trained in a rigorous 33-credit MS degree program to derive decisions and actionable insights from large data-sets. The very last phase of the MS curriculum is the Graduate Qualifying Project (GQP). The project, done in teams, provides the students with a strong capstone experience in applying their data science and analytics skills to a real-world problem. This integrates the theory and practice of data science, and includes the utilization of tools and techniques acquired in the Data Science Program.

The project may be carried out in cooperation with a sponsoring organization (henceforth referred to as sponsor) providing the problem and the data. Each project will be overseen by one or more faculty members affiliated with the Data Science Program (henceforth referred to as faculty). The project will be assigned to a team of 3-4 graduate students. The students will work on tasks including understanding the problem, preparing and analyzing the data provided by the sponsor over a 14-week semester. The GQP culminates with a written report and a presentation. The teams perform their work under the terms of a mutually-beneficial *agreement* between WPI and each sponsor which will include provisions on *confidentiality* and *work product ownership*.

Benefits of a sponsored GQP. *Benefits of these GQP Projects to the students, the sponsors, and the program include:*

- **Work with aspiring data scientists** that focus on tackling the sponsor's problem with a second set of eyes, under guidance of a faculty member
- **Access and screen a well-trained and much-sought-after talent pool** with students interested in the field of big data and data science
- **Form meaningful experiences for WPI Data Science students** as they apply their data science knowledge in a team-work based setting with real-world data analytics problems
- **Share Data Science knowledge** regarding the terminology, challenges and tool sets prevalent in the application domain particular to the sponsor with faculty and students
- **Exploration of new techniques and technologies in Data Science** by the student on behalf of the sponsor in order to create/increase value of the project for the sponsor
- **Contribute to a cutting-edge Data Science and STEM program** via this innovative project-based educational mechanism that bridges the gap between educational training and industry by blending classroom experiences with industry-supplied problems
- **Possibly create new Intellectual Property** as a sponsor through a well-designed GQP project.

CALL FOR PROJECT PROPOSALS

The Data Science program is interested in working with organizations that want to sponsor a team of students to work on a project of interest to the organization for one semester. The sponsor should outline a challenging data science problem for the students to tackle. Below are guidelines for organizations that wish to [submit a project proposal](#). Each year, we expect to conduct approximately ten projects; however this may vary year by year. We aim to select a variety of projects each year. Project proposals that are not selected by any student team can be considered for an additional round in the future.

Overview of GQP Projects

An appropriate GQP project should provide the students with a strong capstone experience in applying data science skills to a real-world problem in your organization. Projects are selected based on the project proposal's fit with the overall educational goals of the program, the learning experience afforded to the students, and interest by the students.

Examples of Projects: Examples of projects below are simple representative examples and are by no means complete or prescriptive of the scope of appropriate GQP projects.

- **Health Care Analytics:** Performing feature selection to identify the most predictive variables in a database of cancer patients, customized based on personal traits, and using them for creating predictive models that can aid general patient medical care as well as personalized healthcare.
- **Data Science for A Sustainable Planet:** Visually exploring energy usage patterns by smart homes equipped with green energy production devices as well as energy monitoring devices to optimize system-wide energy-usage, determine optimized charging, and predict future shortfalls and distribution challenges for smart grids affected by environmental characteristics such as hot summer days or long cold winters.
- **Big Data Security Analysis:** Applying machine-learning approaches for classifying and categorizing Android sources and sinks to determine malicious agents, or analyzing network big data for real-time intrusion detection.
- **Intelligent Transportation Planning:** Mining and querying data from the so called internet of things such as GPS traffic data sets from a large metropolitan area to establish insights into taxi routes, best places to relocate bus stops to maximize flow and/or save fuel for the fleet, the identification of traffic patterns that contribute to increased traffic jams and/or accidents as the region's population undergoes growth.
- **Analytics for Business and Engineering Optimization:** Development of predictive models that extract insights and enable optimization of engineering processes for avoidance for down-time under equipment failure or for monitoring of supply chain enterprise performance for agile adaption under supply change interruption due to weather or natural disasters.

Project Agreements & Contracts

WPI's Corporate Engagement office will work with each sponsor to establish the appropriate agreement for the proposed project. Typically there is a negotiated Sponsorship Fee for these projects and there may also be materials or travel related fees, depending on the project needs. Our Contract Administrator will coordinate with each sponsor on confidentiality and intellectual property ownership provisions and cover requirements regarding the publication of project results. Provisions that deal with IP can fall into a wide spectrum, ranging from IP developed through the project with the sponsor's confidential/proprietary information; IP independently developed when working on the project; to IP jointly developed by the student team, the faculty mentor, and the sponsor while working on the project. The specific provisions covering your project will be addressed in the project agreement.

Sponsor Commitments

To assure the success of our Graduate Qualifying Projects for both our students and the sponsoring organization, each sponsor is required to:

- Assign a primary Point of Contact who will be the technical mentor acting as liaison between the student team and the sponsor's organization throughout the project
- Offer a clear project description up front and needed clarifications throughout the project period to ensure the project is in line with the sponsor's expectations
- Provide timely and secure access to relevant data to assure the project can be undertaken in the given time period in a safe manner
- Provide a Point of Contact in the sponsoring organization for coordination of project documentation

For a successful attainment of the project's objectives, it is desirable that the student team and the sponsor interact regularly during the course of the project, typically weekly. However, this may vary depending on the project and the liaison's availability.

Logistics for Project Proposals

The proposal consists of three components: (1) a concise description (1-2 paragraphs) of the problem to be addressed; (2) a description of the data to be analyzed; and (3) the key contact(s) within your organization championing the project. Our Data Science students are trained in a variety of technical and professional skills. They will collectively have skills selected to match the proposed project. These may include statistics, data cleaning, data management, computer programming, visualization, big data analytics, data mining and machine learning, business intelligence, and familiarity with a variety of analytical software tools. The team will also have professional development skills, such as communication, teamwork, leadership, and collaboration, along with storytelling.

Project Proposal Submission

Project proposals are [submitted online](#) and can be refined by interacting with the Data Science program. The timeline below outlines the typical annual cycle of proposal submission and review as well as of the project execution. Project proposals can be submitted year round and we are ready to begin a dialogue with a prospective sponsor at any time. However, project proposals submitted early in the annual cycle have the best chance of acceptance. Note that the DS program cannot guarantee that a team can be assigned to each proposed project, as assignments will be based on team availability, fit of project and interest by students.

Project Timeline

The expected time line for GQP project solicitation, proposal, review and execution is outlined in the table below. Although most projects will take place during the spring semester (mid-January to end of April), there may be a limited number of student teams available for the fall semester (end of August to mid-December). For these projects, the timeline above is shifted by 8 months. Below we summarize the general timeline in a table format.

	Spring Semester Projects Most projects will be completed in the spring Semester (Late Jan-Late April)	Fall Semester Projects A small number of projects may be done during fall semester, based on availability of student teams (Late Aug-Late Dec)
Project proposal submission and refinements	Throughout the year, generally by October 15	Throughout the year, generally by June 15
Project proposals review and selection	Throughout the year with deadline of November 1	Throughout the year by July 1
Announcement of sponsored projects selected for current year	Around November 15	Around July 15
All project documentation to be completed and signed	Completed by December 15	Completed by August 15
Project launch meetings held on the first week of the semester	Typically around mid-January	Typically around end of August
Project period: Full semester (14 weeks)	Typically mid-January to end of April	Typically end of August to mid-December
Midpoint review meeting	Typically beginning of March	Typically mid-October
Final project delivery and presentation	Typically last week of April	Typically mid-December

More Information

For any questions about becoming a Graduate Qualifying Project (GQP) project sponsor, please contact Elke Rundensteiner, Director of WPI's Data Science Program, via rundenst@cs.wpi.edu or by calling 508-831-5815, or Fatemeh Emdad, GQP Project Coordinator, via femdad@wpi.edu or by calling 508-831-4883. If you are ready to [submit a project proposal online](#), please review the form and submit the form with as much information as you are ready to share.

Updated: February, 2016