

Computational & Data Sciences (Interdisciplinary PhD)

Degree Requirements

PhD in Computational & Data Sciences

Upon joining the PhD program, each student is assigned an initial advisor from the DCDS faculty. This advisor meets with the student to assess their background and to advise them on course selection.

All students complete a common core curriculum as well as a domain depth requirement in a social science area. The focus of the first year is on acquiring a common set of tools and an understanding of the ranges and types of problems students may work on as they progress through the program. The entire incoming cohort takes a unique two-semester seminar sequence solely for DCDS students, which includes both general topics and a series of data-driven dives into the types of research questions that may be encountered in each of the domain areas.

In addition, students will be exposed to research in different areas through “rotations” that start in the fall semester of their first year. By the end of the summer following their first year, each student will put together an advisory committee of at least two DCDS faculty members (preferably from different tracks) and identify the specific track in which they plan to do research and pursue their degree.

Curriculum

Required Core Courses (24 credit units)

- **CSE 502N: Data Structures and Algorithms (3 credits):** This is an existing fundamental course in algorithms and data structures, including significant implementation in an object-oriented programming language (currently Java). We expect that many students will already have this background; the course is intended as a pathway for students with very little computational training.
- **Quantitative Methods I and II (6 credits):** This two-semester sequence covers essential probability and statistics, including hypothesis testing, inference and experimental methodology using a modern statistical computing language like R. The introductory courses offered by the departments of Psychological & Brain Sciences (Psych 5066) and Political Science (Pol Sci 581) will count for Quantitative Methods I credit. Their follow-up courses (Psych 5067 and Pol Sci 582) will count for Quantitative Methods II credit.

- **DCDS 510: Data Wrangling (3 credits):** We are in a new era in terms of the volume and modalities of data generated by efforts to measure human behavior. This will be a new cross-listed course that introduces students to the tools and techniques used to collect, maintain and process large-scale data sets of the kind generated in the course of studying people and social systems.
- **CSE 417T and CSE 517A: Machine Learning I and II (6 credits):** This is a two-semester sequence in machine learning. Together, the two courses cover the fundamental principles of supervised learning, including generalization, overfitting, regularization, cross-validation, model selection, and core machine learning techniques and algorithms, including linear models like logistic regression, gradient descent, tree-based and ensemble methods, kernel methods, deep neural networks and topics in unsupervised learning.
- **Computational and Data Sciences Seminar Series (6 credits):** This two-semester seminar sequence is cross-listed across participating departments and team-taught by participating faculty.
 - DCDS 499: Introduction to Graduate Research in Computational & Data Sciences will be structured around topics and ideas that do not need detailed specific-content background. The topics covered will include ethics, the nature of research, robustness and reproducibility of research, and presentations from across the different areas of interest to give students an understanding of research in human and social data analytics across the university.
 - DCDS 500: Computational and Data Sciences Research Exploration will be structured as a series of deep dives into data-driven approaches in each of the domain areas, including a module on computational methodologies. In each of these modules, the students will either be given a specific dataset to investigate or a specific hands-on task to complete (e.g., developing a visualization, assessing how easy a computational tool is for social scientists to use). Students will work in teams on these projects.

Domain Depth Tracks

Students will choose one of four focus tracks: Political Science, Psychological & Brain Sciences, Social Work & Public Health, or Computational Methodologies. Depending on the track, students must complete the following domain depth requirements:

1. **Political Science track:** Students must complete three substantive courses in one subfield (e.g., American politics, comparative politics, international relations) from a specified list for each subfield as well as a research design course (Pol Sci 540).
2. **Psychological & Brain Sciences track:** Students must complete three substantive courses in one subfield (e.g., brain, behavior and cognition, clinical science, social/personality, development and aging). With permission, students may substitute the Psychological & Brain Sciences Research Methods Course (Psych 5011) for one of the substantive courses, depending on their background in psychological science.

- Social Work & Public Health track:** Students must complete a three-course core doctoral seminar series, including conceptual foundations of social science, advanced research methods, and a theory seminar, either in public health or social work. Students will also be required to take an advanced substantive course from an approved list in their area of interest.
- Computational Methodologies track:** Students must take CSE 541T Advanced Algorithms and either CSE 412A Introduction to Artificial Intelligence or CSE 515T Bayesian Methods in Machine Learning. In addition, students must take two substantive courses in their area of interest (i.e., political science, psychological & brain sciences, or social work & public health) from among the courses acceptable for students in that track as noted above.

Sample Curriculum

A typical progression of courses is described below, with separate examples for students who enter with and without more extensive computational backgrounds.

Students Without Much Computer Science Background

Course	Fall Units	Spring Units
First Year		
Data Structures and Algorithms (CSE 502N)	3	—
Quantitative Methods I	3	—
CDS Seminar I	3	—
Introduction to Machine Learning (CSE 417T)	—	3
Data Wrangling	—	3
CDS Seminar II	—	3
	9	9
Second Year		
Quantitative Methods II	3	—
Domain course	3	3
Domain course or elective	3	3
Machine Learning (CSE 517A)	—	3
	9	9

Students With More Computer Science Background

Course	Fall Units	Spring Units
First Year		
Introduction to Machine Learning (CSE 417T) or domain course	3	—
Quantitative Methods I	3	—
CDS Seminar I	3	—
Introduction to Machine Learning (CSE 417T) or domain course	—	3

Data Wrangling	—	3
CDS Seminar II	—	3
	9	9
Second Year		
Quantitative Methods II	3	—
Domain course	3	3
Domain course or elective	3	3
Machine Learning (CSE 517A)	—	3
	9	9

Further Requirements

Additional requirements for this program are as follows:

- A minimum of 72 credit units beyond the bachelor's level, with a minimum of 37 being course credits (including the core curriculum)
- A minimum of 24 credit units of doctoral dissertation research
- Students must maintain a cumulative average grade of B (3.0 grade-point average) for all 72 credit units.
- Required courses must be completed with no more than one grade below a B-.
- Up to 24 graduate credit units may be transferred with the approval of the Graduate Studies Committee, which is chaired by the director of graduate studies.

In addition to fulfilling the course and research credit requirements, students must do the following:

- Complete at least two three-month-long research rotations.
- Pass a qualifying exam.
- Successfully defend a thesis proposal.
- Present and successfully defend a dissertation.
- Complete a teaching requirement consisting of two semesters of mentored teaching experience.