MA in Computational Social Science

Faculty Directors

• James Evans, Sociology
• Marc Berman, Psychology

Executive Committee

• Luc Anselin, Sociology
• Kathleen Cagney, Sociology
• Guanglei Hong, Comparative Human Development
• Ali Hortaçsu, Economics
• Leslie M. Kay, Psychology
• Howard Nusbaum, Psychology
• John Padgett, Political Science
• Stephen W. Raudenbush, Sociology

Affiliated Faculty

• Wilma Bainbridge, Psychology
• Stéphane Bonhomme, Economics
• Magne Mogstad, Economics
• James T. Sparrow, History
• Rochelle Terman, Political Science

Assistant Director and Assistant Instructional Professor

• Benjamin Soltoff

Assistant Instructional Professors

• Jon Clindaniel
• Philip Waggoner

Preceptors

• Pedro Arroyo
• Elizabeth Huppert
• Shilin Jia
• Sanja Miklin

Managing Director

• Chad Cyrenne

Director of Career Services and Senior Program Development Officer

• Shelly Robinson

Career Preparation and Programming Manager

• Gözde Erdeniz

Student Affairs Administrator

• Vanessa Carey

Administrative and Events Coordinator, MA Programs

• Vanessa Carey

Business Administrator

• Tekeisha Yelton-Hunter

GENERAL INFORMATION

The Master of Arts in Computational Social Science is a two-year program of graduate study. It has a structured curriculum, with a total of 18 required and elective courses tailored to the disciplinary track a student
follows. Students submit an article-length MA thesis in their second year, modeled on a professional journal article.

The program aims to produce leading social scientists in each of our core social science fields – economics, sociology, political science, psychology, history, and anthropology – producing competitive PhD applicants, well-trained in computational approaches, who have mastered the research and analytical skills necessary to make important contributions. MACSS also offers a concentration in economics which is operated jointly with the Department of Economics. Students receive close mentorship from the program’s Faculty Directors, academic staff, and members of our Executive and Affiliated Faculty to design a customized curriculum, define an area of scholarly research, and write a MA Thesis.

They receive full professional support from our Director of Career Services, with biweekly workshops, career planning, and employer recruitment.

Finally, all MA students may participate in an optional summer practicum between their first and second year, with internships drawn from academic and professional organizations. International students have three years of STEM work eligibility after they graduate.

**PROGRAM REQUIREMENTS AND COURSE WORK**

Students submit an article-length MA thesis in their second year.

The courses are selected with the advice of our academic staff, and follow different disciplinary tracks, tailored to the research commitments of each student.

In their first year, all students take a three course core: Perspectives on Computational Analysis, Perspectives on Computational Modeling, and Perspectives on Computational Research.

Most students take a two-course core in Computer Science with Social Science Applications in the Fall and Winter, and Large-Scale Computing for the Social Sciences in the Spring (with more advanced courses for students with prior exposure).

The remaining three courses will typically be graduate social science electives that vary depending on the student’s prior training and intended disciplinary path.

In their second year, all students take three advanced courses in computational methods, tailored to their disciplinary interest. They complete three social science electives, in their area of research. And they take three graduate courses that the students may select, from any university department or professional school, where the student meets the minimum prerequisites. Alternatively, students may take the MA Research Commitment as one of their three courses, intended to further develop the MA Thesis.

Outside of their coursework, all MA students are expected to attend our weekly Computation Workshop, where advanced scholars and invited guests present drafts of their research for critique and discussion.

**ADMISSION**

MACSS applicants must meet the formal requirements of the Graduate Social Sciences Division.

All applicants must submit GRE scores, except for those applying for the joint BA/MA degree.

Financial aid is merit-based, and MACSS offers tuition scholarships at the time of admission. Some financial need-based grants may be available through an application process after prospective students are admitted.

Joint BA/MA applicants pay graduate tuition rates, and are eligible to receive the same aid they had in the College.

Applicants from non-English speaking countries must provide evidence of English proficiency by submitting scores from either the Test of English as a Foreign Language (TOEFL) or the International English Language Testing System (IELTS).

Some non-native English speakers are exempt, if they have studied in an English language University. Please contact our Dean of Students Office with any questions: ssd-admissions@uchicago.edu

**HOW TO APPLY**

The Application for Admission and Financial Aid, with instructions and deadlines, is available online at: https://apply-ssd.uchicago.edu/apply/.

For additional information about our program, please contact Vanessa Carey, our Student Affairs Administrator, at 773-702-8301 or carey1@uchicago.edu.

Please also visit our website: https://macss.uchicago.edu (https://macss.uchicago.edu/)
COURSES
MACS 30000. Perspectives on Computational Analysis. 100 Units.
Massive digital traces of human behavior and ubiquitous computation have both extended and altered classical social science inquiry. This course surveys successful social science applications of computational approaches to the representation of complex data, information visualization, and model construction and estimation. We will reexamine the scientific method in the social sciences in context of both theory development and testing, exploring how computation and digital data enables new answers to classic investigations, the posing of novel questions, and new ethical challenges and opportunities. Students will review fundamental research designs such as observational studies and experiments, statistical summaries, visualization of data, and how computational opportunities can enhance them. The focus of the course is on exploring the wide range of contemporary approaches to computational social science, with practical programming assignments to train with these approaches.
Instructor(s): Benjamin Soltoff Terms Offered: Autumn

MACS 30100. Perspectives on Computational Modeling. 100 Units.
Statistical learning methods are an important tool for computational social scientists seeking to conduct inference and prediction using observable data. This course presents a generic definition of a model in the social sciences, as well as a survey of a wide range of different types of statistical models used. While there is substantial overlap with the field of machine learning, this course focuses on methods emerging from statistics applied to supervised and unsupervised modeling, in service of the goals of inference, classification, and prediction. Students develop an understanding of the theory underlying these methods, apply the methods in an open source computing environment, and interpret model results. As a survey course, students will be exposed to a wide range of modeling strategies, with the expectation that they will later pursue deeper study into specific areas we cover.
Instructor(s): Phillip Waggoner Terms Offered: Winter
Prerequisite(s): MACSS students have priority. Others admitted with instructor consent.

MACS 30120. Computing Fundamentals Boot Camp. 100 Units.
This boot camp focuses on introducing fundamental open-source tools for producing reproducible, computational research. Topics include the basics of Python programming, working on the Linux command line, as well as using Git/GitHub for version control. The course assumes no prior exposure to these topics and serves as preparation for MACS 30121 "Computer Science with Social Science Applications I."
Instructor(s): Jon Clindaniel Terms Offered: Summer

MACS 30121. Computer Science with Social Science Applications I. 100 Units.
This course is the first in a sequence that teaches computational thinking and skills to students from a variety of Social Science disciplines. The course will cover abstraction and decomposition, simple modeling, basic algorithms, and programming in Python. Applications from a wide range of fields serve both as examples in lectures and as the basis for programming assignments. In recent offerings, students have written programs to simulate a model of housing segregation, determine the number of machines needed at a polling place, and analyze tweets from presidential debates.
Instructor(s): Jon Clindaniel Terms Offered: Autumn

MACS 30124. Computational Analysis of Social Processes. 100 Units.
How does the human social and cultural world develop and change? The focus of this course is on introducing computational methods for studying the evolution of phenomena over time, alongside relevant theories for interpreting these processes from fields such as History, Anthropology, and Sociology. Students will gain hands-on experience using the Python programming language to harness a diverse set of digital data sources, ranging from satellite images to social media posts. Additionally, they will learn to employ computational approaches, such as simulation and dynamic topic modeling, to study social processes over a variety of different time scales: from the short term (changes in social media network structures over the course of the past week), to longer term (the evolution of English language discourse over the past 100 years), to deep time scales (long-term settlement pattern dynamics over the past 10,000 years).
Instructor(s): Jon Clindaniel Terms Offered: Autumn
Equivalent Course(s): MAPS 30124

MACS 30150. Perspectives on Computational Modeling for Economics. 100 Units.
Students are often well trained in the details of specific models relevant to their respective fields. This course presents a generic definition of a model in the social sciences as well as a taxonomy of the wide range of different types of models used. We then cover principles of model building, including static versus dynamic models, linear versus nonlinear, simple versus complicated, and identification versus overfitting. Major types of models implemented in this course include linear and nonlinear regression, machine learning (e.g., parametric, Bayesian and nonparametric), agent-based and structural models. We will also explore the wide range of computational strategies used to estimate models from data and make statistical and causal inference. Students will study both good examples and bad examples of modeling and estimation and will have the opportunity to build their own model in their field of interest. This course will be specifically tailored to students concentrating in Economics.
Instructor(s): R. Evans Terms Offered: Winter
Note(s): MACSS students have priority.
MACS 30200. Perspectives on Computational Research. 100 Units.
This course focuses on applying computational methods to conducting social scientific research through a student-developed research project. Students will identify a research question of their own interest that involves a direct reference to social scientific theory, use of data, and a significant computational component. The students will collect data, develop, apply, and interpret statistical learning models, and generate a fully reproducible research paper. We will identify how computational methods can be used throughout the research process, from data collection and cleaning, to exploration, visualization and modeling, to the final communication of results. The course will include modules on theoretical and practical considerations, including topics such as epistemological questions about research design, writing and critiquing papers, and additional computational tools for analysis.
Instructor(s): Philip Waggoner Terms Offered: Spring
Prerequisite(s): MACSS students have priority. Others admitted with instructor consent.

MACS 30250. Perspectives on Computational Research for Economics. 100 Units.
This course focuses on scaling up computational approaches to social science analysis and modeling with big data in context of opportunities afforded by high performance and cloud computing. We will begin by exploring various data structures encountered in social science research, how to deal with large or complex data storage and streaming data, and how to factor considerations of computational complexity into their analyses. We will also study social science applications of parallel computing, both on stand-alone machines and in supercomputing environments, to carry out complex computations. Students will learn to carry out parallel I/O and parallel computation on their own machines and on a cluster. We will also address API construction and access, and explore cloud configurations for social science research designs. We will also help students construct web-based outward facing data, analysis and visualization portals. Students will efficiently gather, structure, perform and present analysis on large-scale social science data. This course will be specifically tailored to students concentrating in Economics.
Instructor(s): R. Evans Terms Offered: Spring
Prerequisite(s): MACSS students have priority.

MACS 30300. Civic Data & Technology Clinic. 100 Units.
The Civic Data and Technology Clinic at the University of Chicago partners with public interest organizations to leverage data science research and technology to address pressing social and environmental challenges. The Clinic also provides students with exposure to real-world projects and problems that transcend the conventional classroom experience including: (1) working with imperfect datasets, applying models and algorithms to real-world data, and navigating security and privacy issues, (2) communicating results to a diverse set of stakeholders (e.g., industry, public interest, government agencies), and translating information into actionable insights, policy briefs and software prototypes. The Clinic is an experiential project-based course where students work in teams as data scientists with real-world clients under the supervision of instructors. Students will be tasked with producing key deliverables, such as data analysis, open source software, as well as final client presentations, and reports. Find details about the current Clinic projects and course here: https://uchicago.box.com/s/v1r4ai5fgn0gmvbteoxgy1a5x4ncg0 Students must apply to participate in this class, and applications are being accepted on a rolling basis through the start of spring quarter. Apply here: https://forms.gle/KNRJ9ZymynjDg7D7
Instructor(s): STAFF Terms Offered: Autumn
Prerequisite(s): Instructor Consent Required to Enroll For the current list of potential projects see here, https://capp.uchicago.edu/sites/capp/files/uploads/Civic Data and Technology Clinic Overview.pdf
Equivalent Course(s): CAPP 30300, PPHA 30581

MACS 30501. Computational Anthropology. 100 Units.
This course exposes students to the methods and data of Computational Anthropology-the systematic, computational study of the human species, past and present. Such methods have been essential in recent years for simulating human behavior in different cultures and economic systems, uncovering ancient demographic changes that still have an influence into the present day, preserving cultural heritage, and much more. Anthropological data allows social science researchers to evaluate long term trends in the human condition, across a variety of cultures, with a unique combination of material, textual, and structured database data. Students will have the opportunity to evaluate state of the art approaches in computational anthropology and learn how to apply these methods to their own social scientific research agendas using open anthropological datasets and the Python programming language.
Instructor(s): Jonathan Clindaniel Terms Offered: Winter

MACS 30519. Spatial Cluster Analysis. 100 Units.
This course provides an overview of methods to identify interesting patterns in geographic data, so-called spatial clusters. Cluster concepts come in many different forms and can generally be differentiated between the search for interesting locations and the grouping of similar locations. The first category consists of the identification of extreme concentrations of locations (events), such as hot spots of crime events, and the location of geographical concentrations of observations with similar values for one or more variables, such as areas with elevated disease incidence. The second group consists of the combination of spatial observations into larger (aggregate) areas such that internal similarity is maximized (regionalization). The methods covered come from the fields of spatial statistics as well as machine learning (unsupervised learning) and operations research. Topics include point pattern analysis, spatial scan statistics, local spatial autocorrelation, dimension reduction, as well as spatially explicit hierarchical, agglomerative and density-based clustering. Applications range from criminology and
public health to politics and marketing. An important aspect of the course is the analysis of actual data sets by means of open source software, such as GeoDa, R or Python.

Instructor(s): L. Anselin Terms Offered: Spring
Prerequisite(s): STAT 22000 or equivalent; SOCI 20253/30253 (or equivalent) Introduction to Spatial Data Science required. Equivalent Course(s): ENST 20519, GEOG 30519, SOCI 20519, SOCI 30519, GEOG 20519

MACS 31300. AI Applications in the Social Sciences. 100 Units.
Artificial Intelligence (AI) describes algorithms constructed to reason in uncertain environments. This course provides an introduction to AI applications in the social sciences. Driven by the rapid increase in accessible big data documenting social behavior, AI has been applied to: increase effective diagnosis and prediction under different conditions, improve our understanding of human interaction, and increase the effectiveness of data management in different social and human services. Random forests and neural networks are among the most frequent AI methods used for prediction, while natural language processing and computer vision contribute to understanding decision-making and improving service provision. We begin with careful consideration for what AI can achieve and where current limitations exist by looking at a variety of real-world applications. We will focus on three core sections: search, representation, and uncertainty. In each section, we will explore major approaches, representational techniques and core algorithms. We will examine the trade-offs between model structure and the algorithmic constraints that this structure implies. The course is driven by hands-on exercises with AI algorithms written in Python. At the end of the term, you should be able to apply and tweak these algorithms to accommodate your own data and research interests.

Instructor(s): Brooke Luetgert Terms Offered: Spring
Prerequisite(s): One course in introductory data science as well as basic familiarity with Python are prerequisites for the course.

MACS 33000. Computational Math Camp. 000 Units.
MACS 35000. MA Research Commitment. 100 Units.
Student Initiated research and writing for the MA research component.
Instructor(s): James Evans Terms Offered: Autumn Spring Winter

MACS 35001. Structured MA Research Commitment. 100 Units.
Student initiated research and writing for the MA research component.
Instructor(s): James Evans Terms Offered: Autumn

MACS 37000. Thinking with Deep Learning for Complex Social & Cultural Data Analysis. 100 Units.
A deluge of digital content is generated daily by web-based platforms and sensors that capture digital traces of human communication and connection, and complex states of society, culture, economy, and the world. Emerging deep learning methods enable the integration of these complex data into unified social and cultural “spaces” that enable new answers to classic social and cultural questions, and also pose novel questions. From the perspective of deep learning, everything can be viewed as data-novels, field notes, photographs, lists of transactions, networks of interaction, theories, epistemic styles and our treatment examines how to configure deep learning architectures and multi-modal data pipelines to improve the capacity of representations, the accuracy of complex predictions, and the relevance of insights to substantial social and cultural questions. This class is for anyone wishing to analyse textual, network, image or arbitrary structured and unstructured data, especially in concert with one another to solve complex social and cultural analysis problems (e.g., characterize a culture; predict next year’s ideology).

Instructor(s): James Evans Terms Offered: Spring
Prerequisite(s): The course uses Python and the widely popular PyData ecosystem to demonstrate all motivating examples and includes working code, accompanying exercises, relevant datasets and additional analytics and visualization that facilitate social and cultural interpretation and communication. Familiarity with Python is required. Equivalent Course(s): SOCI 30332

MACS 40100. Big Data and Society. 100 Units.
The massive explosion of information produced by computers and sophisticated computational methods capable of harnessing this data to generate inferences has led to an increasingly data-driven society. Businesses, governments, and individuals seek to leverage this data to develop and market products, formulate policy, and improve the human condition. Computational approaches to decision making have become increasingly prevalent in domains such as criminal justice, education, employment, finance, and politics. While decision making based on data mining and algorithms has the capacity to improve society, critics argue that these approaches strengthen socioeconomic class divisions, constitute an invasion of privacy, or violate the civil rights of minority groups. This course will survey some of the major uses of big data in society and assess the potential ethical, moral, and legal implications of these models.

Instructor(s): I. B. Soloff Terms Offered: Autumn

MACS 40400. Computation and the Identification of Cultural Patterns. 100 Units.
Culture is increasingly becoming digital, making it more and more necessary for those in both academia and industry to use computational strategies to effectively identify, understand, and (in the case of industry) capitalize on emerging cultural patterns. In this course, students will explore interdisciplinary approaches for...
MA in Computational Social Science
defining and mobilizing the concept of “culture” in their computational analyses, drawing on relevant literature from the fields of Anthropology, Psychology and Sociology. Additionally, they will receive hands-on experience applying computational approaches to identify and analyze a wide range of cultural patterns using the Python programming language. For instance, students will learn to identify emerging social movements using social media data, predict the next fashion trends, and even decipher ancient symbols using archaeological databases.

Instructor(s): Jonathan Clindaniel Terms Offered: Autumn
Prerequisite(s): No previous coding experience required. A Python boot camp will be held at the beginning of the quarter to teach the coding skills necessary to succeed in the course. Open to Advanced Undergraduates with Instructor Permission.
Equivalent Course(s): CHDV 40404, MAPS 40401, PSYC 40460

MACS 40700. Data Visualization. 100 Units.
Social scientists frequently wish to convey information to a broader audience in a cohesive and interpretable manner. Visualizations are an excellent method to summarize information and report analysis and conclusions in a compelling format. This course introduces the theory and applications of data visualization. Students will learn techniques and methods for developing rich, informative and interactive, web-facing visualizations based on principles from graphic design and perceptual psychology. Students will practice these techniques on many types of social science data, including multivariate, temporal, geospatial, text, hierarchical, and network data. These techniques will be developed using a variety of software implementations such as R, ggplot2, D3, and Tableau.

Instructor(s): Benjamin Soltoff Terms Offered: Spring

MACS 41500. MA Research Methods. 000 Units.
This in-person course will foster the development of the students’ scholarship through regular interaction with their preceptors. In this course, students will work with preceptors to both synthesize the individualized coursework into a cohesive curriculum and to plan and execute the MA thesis, from choosing research questions, selecting an appropriate research design, elaborating their chosen methodology, conducting research, and writing up their results.

Instructor(s): John Hansen, Michael Albertus, James Evans Terms Offered: Autumn Spring Winter
Equivalent Course(s): CMES 41500, MAPS 41500, INRE 41500

MACS 51000. Introduction to Causal Inference. 100 Units.
This course is designed for graduate students and advanced undergraduate students from the social sciences, education, public health science, public policy, social service administration, and statistics who are involved in quantitative research and are interested in studying causality. The goal of this course is to equip students with basic knowledge of and analytic skills in causal inference. Topics for the course will include the potential outcomes framework for causal inference; experimental and observational studies; identification assumptions for causal parameters; potential pitfalls of using ANCOVA to estimate a causal effect; propensity score based methods including matching, stratification, inverse-probability-of-treatment-weighting (IPTW), marginal mean weighting through stratification (MMWS), and doubly robust estimation; the instrumental variable (IV) method; regression discontinuity design (RDD) including sharp RDD and fuzzy RDD; difference in difference (DID) and generalized DID methods for cross-section and panel data, and fixed effects model. Intermediate Statistics or equivalent such as STAT 224/PBHS 324, PP 31301, BUS 41100, or SOC 30005 is a prerequisite. This course is a prerequisite for “Advanced Topics in Causal Inference” and “Mediation, moderation, and spillover effects.”

Instructor(s): G. Hong Terms Offered: This course is not offered in 2021-22.
Prerequisite(s): Intermediate Statistics or equivalent such as STAT 224/PBHS 324, PP 31301, BUS 41100, or SOC 30005
Note(s): CHDV Distribution: M; M This course will not be offered in the academic year 2021-22.
Equivalent Course(s): STAT 31900, PLSC 30102, CHDV 30102, SOCI 30315, PBHS 43201

MACS 52000. Advanced Topics in Causal Inference. 100 Units.
This course provides an in-depth discussion of selected topics in causal inference that are beyond what are covered in the introduction to causal inference course. The course is intended for graduate students and advanced undergraduate students who have taken the intro course and want to extend their knowledge in causal inference. Topics include (1) alternative matching methods, randomization inference for testing hypothesis and sensitivity analysis; (2) marginal structural models and structural nested models for time-varying treatment; (3) Rubin Causal Model (RCM) and Heckman’s scientific model of causality; (4) latent class treatment variable; (5) measurement error in the covariates; (6) the M-estimation for the standard error of the treatment effect for the use of IPW; (7) the local average treatment effect (LATE) and its problems, sensitivity analysis to examine the impact of plausible departure from the IV assumptions, and identification issues of multiple IVs for multiple/one treatments; (8) Multi-level data for treatment evaluation for multilevel experimental designs and observational designs, and spilt-over effect; (9) Nonignorable missingness and informative censoring issues.

Instructor(s): G. Hong Terms Offered: Spring. Not being offered in 2021/2022
Prerequisite(s): Intermediate Statistics such as STAT 224/PBHS 324, PP 31301, BUS 41100, or SOC 30005 and Introduction to causal inference or their equivalent are prerequisites.
Note(s): CHDV Distribution: M*
Equivalent Course(s): CHDV 40102, SOCI 40202
MACS 54000. Introduction to Spatial Data Science. 100 Units.
Spatial data science consists of a collection of concepts and methods drawn from both statistics and computer science that deal with accessing, manipulating, visualizing, exploring and reasoning about geographical data. The course introduces the types of spatial data relevant in social science inquiry and reviews a range of methods to explore these data. Topics covered include formal spatial data structures, geovisualization and visual analytics, rate smoothing, spatial autocorrelation, cluster detection and spatial data mining. An important aspect of the course is to learn and apply open source software tools, including R and GeoDa.
Instructor(s): L. Anselin Terms Offered: Autumn
Prerequisite(s): STAT 22000 (or equivalent), familiarity with GIS is helpful, but not necessary
Equivalent Course(s): SOCI 20253, SOCI 30253, ENST 20510, GEOG 20500, GEOG 30500

MACS 60000. Computational Content Analysis. 100 Units.
A vast expanse of information about what people do, know, think, and feel lies embedded in text, and more of the contemporary social world lives natively within electronic text than ever before. These textual traces range from collective activity on the web, social media, instant messaging and automatically transcribed YouTube videos to online transactions, medical records, digitized libraries and government intelligence. This supply of text has elicited demand for natural language processing and machine learning tools to filter, search, and translate text into valuable data. The course will survey and practically apply many of the most exciting computational approaches to text analysis, highlighting both supervised methods that extend old theories to new data and unsupervised techniques that discover hidden regularities worth theorizing. These will be examined and evaluated on their own merits, and relative to the validity and reliability concerns of classical content analysis, the interpretive concerns of qualitative content analysis, and the interactional concerns of conversation analysis. We will also consider how these approaches can be adapted to content beyond text, including audio, images, and video. We will simultaneously review recent research that uses these approaches to develop social insight by exploring (a) collective attention and reasoning through the content of communication; (b) social relationships through the process of communication; and (c) social state
Instructor(s): James Evans Terms Offered: Winter
Equivalent Course(s): CHDV 30510, SOCI 40133

MACS 95000. Computation MA Internship. 000 Units.
All MACSS students who have completed three academic quarters of full-time course work in our MA program are eligible to participate in the Computational Social Science Internship Program. Any interested persons must speak with Career Services, have an approved external employer, complete a petition from our Student Affairs Administrator, and enroll in this non-credit field research course. The course will appear on your transcript, and will be evaluated on a pass/fail basis, in consultation with the employer. Note that MACS 95000 does not count against your other curricular requirements.
Instructor(s): James Evans, Marc Berman Terms Offered: Autumn Spring Summer Winter