JOHNS HOPKINS UNIVERSITY
Whiting School of Engineering | Department of Civil and Systems Engineering

AI & DATA SCIENCE FOR SOCIETAL SYSTEMS

Fall 2026 syllabus | EN.560.315.01

COURSE EN.560.315.01 | Fall 2026	**LECTURES** Tu & Th, 9:00–10:15am
LOCATION Latrobe 320 | Homewood Campus	**CANVAS** https://jhu.instructure.com/courses/131991
INSTRUCTOR Hao Frank Yang, PhD	**EMAIL** haofrankyang@jhu.edu
OFFICE HOURS Th 10:15–11:15 am, Latrobe 320; or by appointment	**TEACHING ASSISTANT** Junyue Jiang
TA EMAIL jjiang67@jh.edu	**FORMAT** Two 75-minute sessions each week

Course purpose

AI becomes consequential when it leaves the benchmark and enters the systems people rely on. Traffic sensors, connected vehicles, electronic health records, mobile devices, and public infrastructure produce enormous streams of data—but data alone does not create better decisions. The hard work is designing what to sense, building trustworthy data pipelines, choosing models that match the problem, and knowing when their outputs are safe enough to guide action.

This course follows that full journey. Using transportation and public health as recurring case domains, we will move from sensing and databases to machine learning, sequential and spatio-temporal models, computer vision, multimodal learning, and generative AI. Along the way, we will confront uncertainty, bias, privacy, fairness, interpretability, and the practical realities of deploying AI in the real world.

You will not only study models; you will build a complete data-to-decision pipeline. By the end of the semester, you will have a documented dataset, a reproducible model repository, an evaluation story, and a final project that connects technical choices to a meaningful societal outcome.

1 SENSE	2 MANAGE	3 LEARN	4 EXPLAIN	5 ACT

Learning objectives

By the end of the course, you will be able to:

1. Translate a societal challenge into an end-to-end data-to-decision pipeline, from sensing and data management to explanation and action.
2. Design privacy-aware data collection and build reproducible, quality-controlled databases and extract-transform-load (ETL) pipelines.
3. Develop and evaluate machine learning, time-series, and spatio-temporal models using appropriate baselines, metrics, uncertainty analysis, and interpretation.
4. Combine time series, images, text, and geospatial data, and use generative AI responsibly for simulation, analysis, and decision support.
5. Identify and address risks involving bias, privacy, security, fairness, robustness, and accountability.
6. Build and communicate a reproducible AI system through clear code, documentation, visual evidence, a technical report, and a presentation.

How the course works

- **Concept sessions:** Core ideas are developed through short lectures, worked examples, and discussion. Tuesday meetings will usually establish the week's framework.
- **Studios and experiments:** Coding demonstrations, method clinics, design reviews, and interpretation exercises turn concepts into practice. Thursday meetings will often emphasize application.
- **Case-driven learning:** Transportation and public health anchor the course, while examples from energy, emergency response, manufacturing, smart buildings, and supply chains show how the methods transfer.
- **Course project:** You may work individually or in a team to propose, build, evaluate, and communicate a data-driven solution for a societal or engineered system.

Prerequisites and tools

- Working knowledge of Python, including data structures, control flow, and basic use of NumPy or pandas.
- Introductory probability, statistics, and linear algebra.
- An introductory transportation, civil systems, public health, or related course is helpful but not required.
- Students without the formal prerequisites may enroll with instructor permission if they have equivalent preparation or relevant experience.

We will use Python, Jupyter, Git/GitHub, NumPy, pandas, scikit-learn, PyTorch, and selected geospatial and generative-AI tools. Setup instructions and starter repositories will be posted on Canvas.

Learning materials

There is no required textbook. Course notes, research papers, case studies, datasets, and technical documentation will be posted on Canvas. Please review the assigned material before class so that our time together can focus on questions, experiments, and design choices.

Recommended references:

- Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*, MIT Press, 2016.
- Roger D. Peng and Elizabeth Matsui, *The Art of Data Science*.
- Jake VanderPlas, *Python Data Science Handbook*, O'Reilly Media.

Selected readings will also come from leading AI, data science, transportation, public health, and systems venues. Reading lists may evolve as the field moves; any updates will be announced on Canvas.

Assessment and grading

The grade is designed to reward both technical mastery and the ability to turn methods into a coherent, responsible system solution.

Assessment	Weight
Homework 1: Sensing, data quality, and databases	15%
Homework 2: Machine learning and sequential modeling	15%
Homework 3: Data-centric and generative AI for systems	15%
Course project: proposal, checkpoint, presentation, report, and repository	20%
Final exam	30%
Discussion, low-stakes checks, and course participation	5%
Total	**100%**

Any curve applied at the end of the semester will only improve, never reduce, a student's letter grade. Detailed rubrics will be posted before each major submission.

Homework

Three homework assignments build the course in stages: first the data foundation, then predictive and sequential modeling, and finally data-centric and generative AI. Unless an assignment explicitly allows collaboration, submitted code, analysis, and writing must be your own. You are encouraged to discuss concepts and debugging strategies, but you must acknowledge substantive help.

Course project

The project is your opportunity to follow a question that matters to you. Examples include crash-risk forecasting, hospital demand, transit reliability, exposure assessment, emergency response, infrastructure inspection, energy management, or human-AI decision support. Strong projects do four things well: frame a meaningful system problem, build a defensible data pipeline, evaluate a suitable model honestly, and communicate what the results can—and cannot— support.

Project milestones:

- **Proposal and data plan:** problem, users, data, risks, baselines, and evaluation strategy.
- **Checkpoint demo:** working pipeline, early results, and a clear account of what is not yet working.
- **Final presentation:** concise problem-to-evidence narrative for a technical audience.
- **Final report and repository:** reproducible code, a documented dataset or data sheet, model card, results, limitations, and next steps.

Invited guests may provide feedback during presentations, but the instructor assigns all grades using the published rubric.

Project report format

The report should read like a short conference paper, not a lab notebook. Use a clear title, abstract, problem statement, related work, data and methods, results, limitations, responsible-AI discussion, and references. The minimum length is 1,500 words for undergraduate students and 2,500 words for graduate students, excluding figures and references.

The Transportation Research Board paper format is recommended as a professional model. The TRB author instructions and TRB Overleaf template are available online. Another standard conference format may be used with instructor approval.

Late work

If you need an extension for a legitimate reason, contact the instructor before the deadline whenever possible. Approved extensions carry no penalty. Without an approved extension, work may be submitted up to four calendar days late with a 50% deduction; after four days, it will normally receive no credit. If serious or unexpected circumstances arise, communicate as soon as you are able so that we can identify a fair path forward.

Communication

Canvas is the source of truth for announcements, files, assignment instructions, and schedule changes. Use email for individual questions or circumstances that should remain private. Questions that may help the whole class should be posted in the course discussion space. Please allow up to two business days for an email response.

Classroom participation and technology

Bring a laptop when a coding activity is announced. During class, use devices for course work and keep phones silent and out of the way. Food and drinks are welcome when they do not distract others or create a safety issue. Participation means arriving prepared, contributing thoughtfully, asking questions, and helping create a classroom where people can test ideas without embarrassment.

Attendance and absences

Regular attendance matters because many activities are interactive and cannot be fully recreated from slides. If health, family, religious observance, varsity athletics, or another significant circumstance requires you to miss class, notify the instructor as early as possible and arrange to make up essential work. Excused absences do not reduce the participation grade. Repeated unexcused absences or disengagement may affect it. JHU's attendance and religious-holiday guidance applies.

Responsible use of generative AI

Because this course studies AI, we will use AI tools—but we will use them transparently and critically.

Permitted uses, unless an assignment says otherwise:

- Brainstorming, outlining, explaining concepts, debugging, language editing, and generating small test cases.
- Exploring alternative methods or interpretations, provided you verify all technical claims and results.

Your responsibilities:

- Disclose meaningful AI assistance in a short statement that names the tool and explains how it was used.
- Verify code, equations, citations, factual claims, and data. You remain responsible for every submitted element.
- Protect privacy: never upload confidential, proprietary, sensitive, or personally identifiable course or project data to an external AI service.

Not permitted:

- Using AI during quizzes or the final exam unless explicitly authorized.
- Fabricating data, results, citations, experiments, or sources.
- Submitting unedited AI output as your own work or using AI in a way that prevents you from explaining and defending the submission.

Academic integrity

The strength of Johns Hopkins University depends on academic and personal integrity. Be honest and truthful in your data, code, analyses, writing, collaboration, and use of AI. Academic misconduct includes cheating, plagiarism, unauthorized collaboration, reuse of work without permission, falsification, improper use of electronic tools, and helping another person violate course rules.

The current Whiting School teaching and academic-integrity guidance and Homewood academic policies apply to this course.

Accessibility and accommodations

Johns Hopkins University is committed to welcoming, equitable, and accessible educational experiences. If you need disability-related accommodations, request them through Student Disability Services as early as possible so that arrangements can be made effectively. You are welcome to speak with the instructor about access needs, but formal accommodations must be coordinated through SDS.

Homewood Student Disability Services is located in Shaffer 101. Phone: 410-516-4720. Homewood SDS

Inclusive learning environment

Our strongest work comes from a classroom where people with different backgrounds, identities, disciplines, and levels of experience can participate fully. Treat classmates with respect, critique ideas rather than people, and make room for voices that have not yet been heard. Discrimination, harassment, and retaliation are inconsistent with University policy and with the culture of this course.

Concerns may be raised with the instructor or with the Johns Hopkins Office of Institutional Equity.

Syllabus changes

This syllabus is a working roadmap. The instructor may adjust topics, readings, or due dates to respond to the pace of the class, new developments in the field, or University requirements. Material changes will be announced on Canvas with reasonable notice.

Fall 2026 course schedule

The schedule below is aligned with the Johns Hopkins University Fall 2026 calendar. Canvas will contain the most current version and detailed readings, datasets, and assignment instructions.

Week	Dates	Focus	Milestone
1	Sep 1 & 3	Why societal systems need AI; the data-to-decision pipeline; course launch; transportation and public health case framing.	Course launch
2	Sep 8 & 10	Sensing landscapes, sensor capabilities, failure modes, observability, and the design of reliable field data collection.	Sensing tech survey
3	Sep 15 & 17	Data management I: quality, validation, cleaning, missingness, bias, reproducible ETL, provenance, and drift monitoring.	Data clinic
4	Sep 22 & 24	Data management II: Relational, time-series, geospatial, and document databases; querying; visualization that communicates uncertainty.	HW 1 due Oct 1
5	Sep 29 & Oct 1	Machine learning foundations: problem formulation, baselines, train/validation/test design, metrics, uncertainty, and interpretability.	Model evaluation lab
6	Oct 6 & 8	Sequential and spatio-temporal learning: forecasting, anomaly detection, RNNs, attention, and Transformers.	HW 2 released Oct 15
7	Oct 13 & 15	Computer vision for systems automation; multimodal alignment and fusion across images, text, time series, and maps.	Project topics released Oct 22
8	Oct 20 & 22	Representation learning: encoders, decoders, embeddings, and latent spaces; designing multimodal AI systems for societal applications by integrating images, text, and time series.	Project topics released Oct 22
9	Oct 27 & 29	Coding studio: sequential modeling experiments using Codex; Vibe Coding, project design clinic; reproducibility and error analysis.	Proposal due Oct 30
10	Nov 3 & 5	Generative AI foundations: language models, prompting, retrieval-augmented generation, adapters, and system integration.	GenAI studio I
11	Nov 10 & 12	Generative AI foundations II: hallucination, grounding, robustness, scenario generation, and agentic workflows.	HW 2 due Nov 12
12	Nov 17 & 19	Generative AI foundations III: Human-AI teaming, guardrails, fairness, model cards, and accountable deployment.	HW 3 released Nov 19
13	Nov 24 & 26	Fall Recess—no class meetings.	No class
14	Dec 1 & 3	Coding studio: generative and multimodal AI for social systems using Codex; project checkpoint demos and peer review.	Checkpoint Dec 3
15	Dec 8 & 10	Course synthesis: connecting sensing, data stewardship, learning, explanation, and action; final project presentations.	Presentations Dec 10
After	Dec 11–23	HW 3 due Dec 11. Project report and repository due Dec 15. Final exam during the University exam period; exact date and time will follow the Registrar's schedule.	Exam period Dec 15–23

Key University dates

August 31	First day of Fall 2026 classes
September 7	Labor Day—University closed; no class
November 23–27	Fall Recess; no class
December 10	Last meeting for this Tuesday/Thursday course
December 11	University last day of semester classes
December 14	Reading Day
December 15–23	Final exam period

Official calendar: Johns Hopkins University Registrar.