



JOHNS HOPKINS

WHITING SCHOOL  
of ENGINEERING

**Computer Science EN.601.465 / EN.601.665**  
**Natural Language Processing**  
**Fall 2025 (4 credits)**

**Instructor**

Professor **Jason Eisner** (he/him) <[jason@cs.jhu.edu](mailto:jason@cs.jhu.edu)>

Office hours: After class in the classroom, or by appointment

Office location: Hackerman 324C, or <https://wse.zoom.us/my/jeisner1>

**Teaching Assistants**

Head CA: Edmund Sumpena <[esumpen1@jh.edu](mailto:esumpen1@jh.edu)>

CA: Shaobo Liang <[sliang24@jh.edu](mailto:sliang24@jh.edu)>

CA: Rishi More <[rmore2@jh.edu](mailto:rmore2@jh.edu)>

CA: Jiru Xu <[jxu188@jh.edu](mailto:jxu188@jh.edu)>

CA: Arda Uzunoglu <[auzunog1@jh.edu](mailto:auzunog1@jh.edu)>

CA: Hanxiang Qin <[hqin14@jh.edu](mailto:hqin14@jh.edu)>

CA: Alvin Zhang <[bzhang90@jh.edu](mailto:bzhang90@jh.edu)>

You can reach all the course staff at once via a private post on Piazza, or via [cs465-staff@cs.jhu.edu](mailto:cs465-staff@cs.jhu.edu).

Office hours: TBA. See the class homepage, <http://cs.jhu.edu/~jason/465>.

**Meetings**

Mon/Wed/Fri 3:00–4:15 pm, Remsen 101

Tue 6:00–7:30 pm, Remsen 101 (recitation)

This class is in the “flexible time slot” MWF 3–4:30. **Please keep the entire slot open.** Class will usually end at 4, followed by office hours in the classroom from 4–4:30 (stick around to get your money’s worth!). However, class will sometimes run till 4:15 in order to keep up with the syllabus. I’ll try to give advance notice of these “long classes,” which among other things make up for no-class days when I’m out of town.

We also run a once-per-week recitation led by the prof or the TAs. This session is normally held in the Tuesday slot, and will usually focus on solving problems together. It is meant as an efficient and cooperative way to study for an hour: it reinforces the past week’s class material without adding to your homework load. Thanks to recitation, you won’t be startled by the exam style—the discussion problems are taken from past exams and are generally interesting.

Occasionally we may swap lecture and recitation slots.

We’ll occasionally ask you to watch some video lessons outside of class. (Similar to a required reading. These video lessons were created during a previous attempt to “flip the classroom” for Covid virtual learning.)

If you miss a lecture, please email [<eisner@jhu.edu>](mailto:eisner@jhu.edu) to get access to a video recording. Watch it promptly so you can follow subsequent lectures. Lecture recordings are also available for review upon request, and will be made available before each exam.

If you click the CC button, automatically generated closed captions will show up on the video. You can also view the full transcript and use it to navigate to a specific part of the video. (The video lessons have manual captions. The lecture recordings have automatic captions; these aren't perfect, but they might help you understand my words better.)

Recitations will generally not be recorded, but the materials from recitation are available [on the class website](#).

## Textbooks

This course does not follow any textbook. The material is covered by the lectures and videos and the reading handouts that are associated with the homework projects, supplemented by in-class discussion.

However, the following textbooks are at an appropriate level. They are usually a good reference to get another written perspective on the material. The website will suggest optional readings from time to time.

- Dan Jurafsky & James Martin (2025), *Speech and Language Processing (3rd ed.)*. A draft of the 3rd edition is online [here](#).
- Jay Alamar & Maarten Grootendorst (2024), *Hands-On Large Language Models*. Code for the book's examples is [here](#).
- Jacob Eisenstein (2019), *Introduction to Natural Language Processing*.
- Brian Roark & Richard Sproat (2007), *Computational Approaches to Morphology and Syntax*.
- Chris Manning & Hinrich Schütze (1999), *Foundations of Statistical Natural Language Processing*. PDFs of the full book and the individual chapters are available via the JHU library at [this link](#).

## Online Resources

Everything you need will be linked from the class homepage, <http://cs.jhu.edu/~jason/465>.

Go there now! Make sure to sign up for the Piazza site, and make sure to watch lecture videos when they are assigned later. (And of course, come to class, come to recitations, do the homeworks, take the exams.)

## Course Information

- **Catalog description:** An in-depth introduction to core techniques for analyzing, transforming, and generating human language. The course spans linguistics, modeling, algorithms, and applications. (1) How should linguistic structure and meaning be represented (e.g., trees, morphemes,  $\lambda$ -terms, vectors)? (2) How can we formally model the legal structures and their probabilities (e.g., grammars, automata, features, log-linear models, recurrent neural nets, Transformers)? (3) What algorithms can estimate the parameters of these models (e.g., gradient descent, EM) and efficiently identify probable structures (e.g., dynamic programming, beam search)? (4) Finally, what kinds of systems can be built with these techniques and how are they constructed and evaluated in practice? Detailed assignments guide students through many details of implementing core NLP methods. The course proceeds from first principles, although prior exposure to AI, statistics, ML, or linguistics can be helpful.
- **Prerequisites**
  - Data Structures (601.226)
  - Python
  - Basic familiarity with partial derivatives, matrix multiplication, and probabilities

The class aims to be fairly self-contained and teach you everything else you need, presenting it from an NLP perspective. That includes relevant aspects of formal languages (601.431/631), probability

(553.420/620 or 553.310/311), and machine learning (601.475/675, 601.482/682, ...). So those courses are not formal prerequisites. That said, it may be helpful to have had prior exposure to those concepts.

- **Elective** (Applications)

## Course Goals

This course is designed to introduce you to some of the problems and methods of natural language processing, and their relation to linguistics and statistics. At the end you should agree (I hope!) that language is subtle and interesting; feel some ownership over some of NLP's formal and statistical techniques; and be able to understand research papers in the field. (Caveat: Understanding recent research papers may sometimes require additional background in machine learning.)

In the end, I hope that NLP, like all good courses, stretches your mind and leaves you with new ways of thinking. Not to mention [all this](#).

Specific Outcomes for this course:

- Show sensitivity to linguistic phenomena and an ability to model them with formal grammars.
- Understand and carry out proper experimental methodology for training and evaluating empirical NLP systems.
- Be able to manipulate probabilities, construct statistical models (including neural networks) over strings and trees, and estimate parameters using supervised and unsupervised training methods.
- Be able to design, implement, and analyze NLP algorithms.

This course will address the following [CSAB Criterion 3 Student Outcomes](#). Graduates of the program will have an ability to:

1. Analyze a complex computing problem and to apply principles of computing and other relevant disciplines to identify solutions.
2. Design, implement, and evaluate a computing-based solution to meet a given set of computing requirements in the context of the program's discipline.
5. Function effectively as a member or leader of a team engaged in activities appropriate to the program's discipline.

## Course Topics

Please see the class website for a list of topics and resources: <http://cs.jhu.edu/~jason/465>.

## Course Expectations & Grading

Your work will be weighted as follows (although we are considering a fancier scheme):

- **50% homeworks**, equally weighted (see [lateness policy](#))
- **15% midterm exam**
- **30% final exam**
- **≈ 5% participation**

Participation includes your contributions to an interesting and useful class discussion, whether synchronously during class sessions, or via public posts or replies on our Piazza discussion website. This includes asking questions, of course. We've found that participation has a substantial effect on the final grade.

Homeworks will be submitted via Gradescope. We expect to have 8 homework projects, which focus on different skills you'll need to do NLP:

- (1) writing linguistic grammars
- (2) manipulating probability formulas
- (3) implementing and properly evaluating supervised models
- (4) designing and implementing combinatorial algorithms
- (5) working with formal meaning representations
- (6) unsupervised learning of models for structured prediction
- (7) discriminative learning of neural models for structured prediction

## (8) controlling and evaluating large language models

We plan to have two exams—a midterm exam and a final exam. These will have some challenging questions. In the recitations, you'll get to work in small groups to solve interesting problems from past exams.

Since some questions are hard or are graded harshly, we use a curve to pull the grades back up. If your grade was already high in absolute terms, it will stay high (so if everyone gets 100, everyone gets an A+, not a C).

Extra credit is added only after the curve is determined. Thus, extra credit is not required to do well, although it can help make up for low grades elsewhere.

### Late Homework Policy

Since emergencies sometimes arise, I'll grant extensions on request, in multiples of 24 hours, up to a *total* of up to 10 "late days" during the term. The full policy and advice on how to use it are at <http://cs.jhu.edu/~jason/465/late-policy.html>.

### Audit Policy

If you register for the course as an auditor (or switch to this status), please let me know that you are auditing. You're still expected to attend the lectures and recitations and to participate in them like any other student. (That's what it means for the course to appear on your transcript with an AU grade.) If you are not able to sustain this level of commitment, you should drop the class.

You are also welcome to do any homeworks or exams that are useful to your learning. We will grade them and give you the usual feedback.

### Key Dates

Please see the class website: <http://cs.jhu.edu/~jason/465>.

### Assignments & Readings

Please see the class website: <http://cs.jhu.edu/~jason/465>.

### Assistance and Collaboration Policy

*Understanding course content:* We really want you to understand the course material! Pretty much anything is fair game for this. We already provide lectures, reading handouts, recitation puzzles, office hours, and optional recommended readings. But you might also look for other readings, talk to classmates or CAs in person or on Piazza, or discuss in depth with an AI chatbot.

Please share useful materials and insights on Piazza so that others can also benefit. As NLP folks, we are especially curious about whether discussion with AI chatbots can help your understanding. Note that ChatGPT offers a "[Study Mode](#)" that is designed for this ([direct link](#)).

*Doing graded work:* The point of the homeworks and exams is for you to wrestle with problems. That's how you learn, and also how you confirm your understanding. It's also how we confirm your understanding (for grading). We want you to succeed! However, if you don't solve the problems yourself, then you're not the one who is succeeding. So, it's ok to get help, but not in a way that solves the problems for you. Here are key principles:

- It's always fine to get help from the course staff. They will use good judgment about how much help is appropriate.
- Some homeworks allow collaboration with 1–2 partners. (In most cases, this means making a single submission as a group.) You are expected to do the work *together* with your partners, not divide it up: if you didn't work on a question, you don't deserve credit for it! Your solutions should emerge from collaborative discussions with the *whole group* participating.

- If you get help in your graded work from anyone else—that is, help beyond course staff and your partners—then you need to disclose this in detail on your homework submission. This includes help from
  - other humans
  - AI chatbots
  - documents that give the answers
- You only need to disclose help on actually solving the graded problems. You don’t need to disclose help that merely gets you up to speed so that you are prepared to do the problems, such as
  - “how do I read and write JSON files in Python?”
  - “why isn’t matrix multiplication commutative?”
  - “give me 5 matrix expressions in PyTorch and their mathematical notation equivalents in T<sub>E</sub>X”
  - “I’m suddenly confused about lecture 12”
- You need to do substantive work to get credit:
  - Even if you get some help, **your written solutions should always be in your own words**. This ensures that you actually have the ideas in your head. Your English will be good enough for us to follow; there’s no need to have an AI rewrite it for you.
  - **You should also write your own code**; don’t use an AI-powered code editor. This ensures that you are competent to write or review the core code for NLP algorithms. (If you’re thinking that it might be appropriate to have AI fill in non-core code—boring parts that have little to do with NLP—there’s actually no need for that in this course. We have already filled those parts in for you, by providing well-crafted “starter code” that you should read.)
  - Don’t provide an AI with the actual homework questions or starter code. That’s too close to having it solve the problems for you.
- If you’re not sure what’s allowed, or you’d like permission to do something that seems reasonable to you, please ask us! And err on the side of disclosure. Better safe than sorry.

---

*The material below is standard for CS Department syllabi as of Fall 2024, except for portions in this color.*

## Integrity

The strength of the university depends on academic and personal integrity. You must be honest and truthful in this course, abiding by the CS department’s [Academic Integrity Code](#):

Cheating is wrong. Cheating hurts our community by undermining academic integrity, creating mistrust, and fostering unfair competition. The university will punish cheaters with failure on an assignment, failure in a course, permanent transcript notation, suspension, and/or expulsion. Offenses may be reported to medical, law or other professional or graduate schools when a cheater applies.

Violations can include cheating on exams, plagiarism, reuse of assignments without permission, improper use of the Internet and electronic devices, unauthorized collaboration, alteration of graded assignments, forgery and falsification, lying, facilitating academic dishonesty, and unfair competition. Ignorance of these rules is not an excuse.

Academic honesty is required in all work you submit to be graded. Except where the instructor specifies group work, you must solve all homework and programming assignments without the help of others. For example, you must not look at anyone else’s solutions (including program code) to your homework problems. However, you may discuss assignment specifications (not solutions) with others to be sure you understand what is required by the assignment.

*If* your instructor permits using fragments of source code from outside sources, such as your textbook or on-line resources, you must properly cite the source. Not citing it constitutes plagiarism. Similarly, your group projects must list everyone who participated.

Falsifying program output or results is prohibited.

Your instructor is free to override parts of this policy for particular assignments. To protect yourself: (1) Ask the instructor if you are not sure what is permissible. (2) Seek help from the instructor, TA or CAs, as you are always encouraged to do, rather than from other students. (3) Cite any questionable sources of help you may have received.

On every exam, you will sign the following pledge: "I agree to complete this exam without unauthorized assistance from any person, materials or device. [Signed and dated]". Your course instructors will let you know where to find copies of old exams, if they are available.

In this class (NLP), practice exam problems will be provided for you. You should not otherwise make use of homeworks or exams from previous years.

Report any violations you witness to the instructor. You can also contact:

- For undergraduates:
- the Office of the Dean of Student Life at 410-516-8208 or [<studentconduct@jhu.edu>](mailto:studentconduct@jhu.edu)
- For KSAS Graduate Students: [<rseitz5@jh.edu>](mailto:rseitz5@jh.edu)
- For WSE Graduate Students: [<christinekavanagh@jhu.edu>](mailto:christinekavanagh@jhu.edu)

### **Students with Disabilities – Accommodations and Accessibility**

Johns Hopkins University is committed to providing welcoming, equitable, and accessible educational experiences for all students. If disability accommodations are needed for this course, students should request accommodations through Student Disability Services (SDS) as early as possible to provide time for effective communication and arrangements.

For further information about this process, please refer to the [SDS Website](#) or email SDS Homewood: [<studentdisabilityservices@jhu.edu>](mailto:studentdisabilityservices@jhu.edu).

### **Classroom Climate**

The following statement was provided by the Homewood Council on Inclusive Excellence and is incorporated into all WSE course syllabi.

Johns Hopkins University is committed to creating a classroom environment that values the diversity of experiences and perspectives that all students bring. Everyone here has the right to be treated with dignity and respect. Fostering an inclusive climate is important because research and experience show that students who interact with peers who are different from themselves learn new things and experience tangible educational outcomes. Please participate in creating a welcoming and vibrant classroom climate. Note that you should expect to be challenged intellectually by the instructor, the TAs, and your peers, and at times this may feel uncomfortable. Indeed, it can be helpful to be pushed sometimes in order to learn and grow. But at no time in this learning process should someone be singled out or treated unequally on the basis of any seen or unseen part of their identity.

If you ever have concerns in this course about harassment, discrimination, or any unequal treatment, or if you seek accommodations or resources, please reach out to your instructor or the TAs, who will take your communication seriously and seek mutually acceptable resolutions and accommodations. Reporting will never impact your course grade. You may also share concerns with the Department Head (Randal Burns, [<randal@cs.jhu.edu>](mailto:randal@cs.jhu.edu)), the

Director of Undergraduate Studies (Joanne Selinski, <[joanne@cs.jhu.edu](mailto:joanne@cs.jhu.edu)>), the Director of Graduate Studies (Scott Smith, <[scott@cs.jhu.edu](mailto:scott@cs.jhu.edu)>), the WSE Associate Dean of Outreach and Belonging (Darlene Saporu, <[dsaporu@jhu.edu](mailto:dsaporu@jhu.edu)>), the KSAS Assistant Dean for Diversity and Inclusion (Araceli Frias, <[afrias3@jhu.edu](mailto:afrias3@jhu.edu)>), or the Office of Institutional Equity (<[oie@jhu.edu](mailto:oie@jhu.edu)>).

In handling reports, people will protect your privacy as much as possible, but faculty and staff are required to officially report information for some cases (e.g., sexual harassment).

## Personal Wellbeing

**Illness.** If you are sick, in particular with an illness that may be contagious, please notify me by email but do not come to class. I will provide you with a live Zoom link and/or a video recording of the lecture. We can also make arrangements to take exams remotely if necessary.

You may also contact **Student Outreach and Support** at <[studentoutreach@jhu.edu](mailto:studentoutreach@jhu.edu)> for help managing your coursework during an illness, with **medical leave of absence** being an option for a longer illness. For your medical needs, see the **Student Health and Wellness Center**.

**Other Reasons to Miss Class.** If you have to miss class for another reason, such as a religious holiday or a visa problem, please notify me by email in advance. I will try to accommodate you as above (within reason). In general, of course, you are expected to come to class.

**Mental Health.** JHU has several resources to support students. Many students struggle at times with stress, anxiety, and depression. The Counseling Center has many resources available to you:

- **Johns Hopkins University Student Well-Being**

In addition, The Johns Hopkins University Behavioral Health Crisis Support Team (BHCST) pairs experienced, compassionate crisis clinicians with specially trained public safety officers on every shift on and around the Homewood campus, seven days a week. The BHCST will provide immediate assistance to those who need it and, just as importantly, link individuals in crisis to ongoing support services in the days and weeks that follow. Call Public Safety, 410-516-5600, and ask for a BHCST clinician.

If you have concerns about a specific student, please contact:

- For **emergencies** (threat to self or others): 410-516-4600 or 911
- For on-scene mental health support: BHCST at 410-516-4600
- For undergraduates: Student Outreach & Support at 410-516-7857 or <[studentoutreach@jhu.edu](mailto:studentoutreach@jhu.edu)>
- For KSAS Graduate Students: **Renee Eastwood**, Assistant Dean for Graduate and Postdoctoral Academic and Student Affairs
- For WSE Graduate Students: **Megan Barrett**, Assistant Dean for Engineering Student Affairs

**Family Accommodations Policy.** You are welcome to bring a family member to class on occasional days when your responsibilities require it (for example, if emergency childcare is unavailable, or for health needs of a relative). Please be sensitive to the classroom environment, and if your family member becomes uncomfortably disruptive, you may leave the classroom and return as needed.