Course Syllabus:

How to Build a Bot

Course Information

Summer 2025

606 Pratt Manhattan Campus

M/Th 10:00 AM - 1:50 PM

5/27/2025 - 7/8/2025

Credits: 3

Prerequisites or other restrictions: INFO 664

Instructor Contact Information

Filipa Calado, PhD (she/her)

Office Location: Pratt Manhattan Campus, room 602

Office Hours: by appointment (summer only)

Phone: 718-687-5194

Email: fcalado@pratt.edu

Zoom: https://pratt.zoom.us/my/fcalado

Website: https://filipacalado.com

Course Description

This course offers a practical introduction to building bots in Python alongside a critical examination of algorithmic bias. Students will learn core programming skills like data collection and API interaction in order to create web crawlers and social media bots. In parallel, students will explore the ethical consequences of automated systems on social platforms, such as the amplification of misinformation and bias that perpetuate social inequalities and discrimination. This course will equip students with the critical perspectives and technical skills to analyze automated systems in a world increasingly shaped by AI and algorithmic decision-making. This course is intended for students interested in both technical development and the social impacts of automation.

Student Learning Outcomes

Upon successful completion of this course, students will be able to:

- develop programming literacies for working with popular software in Python for web

crawling and API usage.

- gain hands-on experience in computational data collection and parsing.
- identify and explain how choices in bot design, such as source selection, keyword filtering, and data processing, can introduce and/or amplify biases.
- critically assess the societal impact of bots, especially those used in social media, evaluating how bots might influence information visibility and public discourse.

Course Format

This course will be held in hybrid format, with the first four weeks of meetings (May 29 - June 23) in person, and the last two weeks of meetings (June 26 - July 7) on zoom.

Our sessions will be split evently between lecture, individual practice, and group work. This contrasts with the prerequisite course, INFO 664, which focused more heavily on lecture.

For each meeting, one 30-minute break will take place from approximately 12:00pm-12:30pm.

course materials

Having a personal laptop (not a tablet) where you can install software is essential for this class.

All assignments and readings will be provided electronically and hosted on github at https://github.com/gofilipa/botz/class_notes

Class Communication

The instructor will contact you via your pratt email (linked to Canvas). If you don't check that email frequently, please remember to do so for this class or set up mail forwarding.

The best avenue for contacting the instructor is via email, at fcalado@pratt.edu. Response time should be within 2 business days, otherwise feel free to follow up.

Portfolio

Work completed for this course may be included in your portfolio. For more information on each program's portfolio requirements, please visit the program's respective webpage:

MS Library & Information Science: Portfolio - http://bit.ly/prattmslisportfolio

MS Information Experience Design: Portfolio - http://bit.ly/prattmsixdportfolio

MS Data Analytics and Visualization: Portfolio - http://bit.ly/prattmsdavportfolio2

MS Museums and Digital Culture: Portfolio - http://bit.ly/prattmsmdcportfolio2

Your final project would be a good possibility to include in your portfolio

Course Schedule

Unit 1 web crawling bots - 2 weeks

May 29, session 1: intro to Python & web scraping with bs4

Homework (due June 2) : reflection on interests & scraping

- what kind of data are you interested in? Once you've done some refelction, find 2 websites to scrape. Make sure if they are actually scrapable with bs4.

June 2, session 2: scrapy & the scrapy shell

Homework (due June 5): Compost Engineers chapters 1 & 2, and reading response (prompt below)

- Joana Varon and Lucía Egaña Rojas. Chapters 1 & 2 from Compost Engineers and Sus Saberes Lentos: A Manifest for Regenerative Technologies. Coding Rights, 2024, https://codingrights.org/docs/compost_engineers.pdf.
- Prompt: Pick an idea from the reading that interests you (either because you agree with it, disagree with it, or are otherwise provoked by it) and explain why. 1 page.

June 5, session 3: blockers & XHR

Homework (due June 9): Compost Engineers chapters 3 & 4, and reading response (prompt below)

- Joana Varon and Lucía Egaña Rojas. Chapters 3 & 4 from Compost Engineers and Sus Saberes Lentos: A Manifest for Regenerative Technologies. Coding Rights, 2024, https://codingrights.org/docs/compost_engineers.pdf.
- Prompt: From the authors' proposals, what do you find useful or surprising, and what do you have doubts about? 1 page.

June 9, session 4: selenium

Assignment (due June 12): web scraping

- Using either scrapy or selenium, scrape some data from a website that you couldn't scrape before, and submit your python file(s) and the data on canvas.

Unit 2 chat bots - 1.5 weeks

June 12, session 5: spaCy for processing text

Homework (due June 16): ACLU tech & privacy analysis write-up

- Choose a topic from the last 6 months on the "Tech & Privacy" page on the ACLU website: https://www.aclu.org/press-releases?issue=privacy-technology

- Write up analysis of what is going on, and your opinion on the issue. How does the issue handle privacy rights and ethical uses of data? 1 page.

June 16, session 6: spacy continued, intro to transformers

Homework (due June 19): run a task on your own data

June 19, session 7: transformers continued

Assignment (due June 23): dataset proposal

- What is the dataset you'd like to create for your final project? Where would you get the data, and how would you transform it? You can consider tools from this class (like text generation, named entity recognition, pattern matching), or you can consider other possibilities for transforming your data. 1 page, double spaced, submitted on canvas.

Unit 3 social media bots - 1.5 weeks

June 23, session 8: twitter bots

Homework: make a plan for actions steps by next class

(online) June 26, session 9: group projects

Work on projects

(online) June 30, session 10: group projects continued

Homework (due July 3): project proposal due on Canvas

Unit 4 project workshops & presentations – 1 week

(online) July 3, final projects

Work on projects

(online) July 7, final project presentations

Assignments

Participation (30%)

- Includes in-class engagement, and completing and sharing homework

Unit assignments (30%)

- Average score of 3 assignments at the end of units 1-3

Final project: a bot! (40%)

- A final project that takes some data from web scraping or APIs, and uses it as the content for a bot.
- bot to be automated and published on github.

Recommended readings

On data gathering and web scraping:

- Dodge, Jesse, et al. "Documenting Large Webtext Corpora: A Case Study on the Colossal Clean Crawled Corpus." Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, edited by Marie-Francine Moens et al., Association for Computational Linguistics, 2021, pp. 1286–305. ACLWeb, https://doi.org/10.18653/v1/2021.emnlp-main.98.
- Jo, Eun Seo, and Timnit Gebru. "Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning." Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Association for Computing Machinery, 2020, pp. 306–16. ACM Digital Library, https://doi.org/10.1145/3351095.3372829.
- Chan, Anita Say. Predatory Data: Eugenics in Big Tech and Our Fight for an Independent Future. University of California Press, 2025. library.oapen.org, https://doi.org/10.1525/luminos.215.
- Métraux, Julia. "Eugenics Isn't Dead—It's Thriving in Tech." Mother Jones, https://www.motherjones.com/politics/2025/01/eugenics-isnt-dead-its-thriving-in-tech/. Accessed 14 Feb. 2025.

On machine learning:

- Alammar, Jay. The Illustrated BERT, ELMo, and Co. (How NLP Cracked Transfer Learning). https://jalammar.github.io/illustrated-bert/. Accessed 14 Apr. 2025.
- Alammar, Jay. The Illustrated DeepSeek-R1. 10 Feb. 2025, https://newsletter.languagemodels.co/p/the-illustrated-deepseek-r1.

Case studies of algorithmic bias & audits:

- Hada, Rishav, et al. "Akal Badi Ya Bias: An Exploratory Study of Gender Bias in Hindi Language Technology." The 2024 ACM Conference on Fairness, Accountability, and Transparency, ACM, 2024, pp. 1926–39. DOI.org (Crossref), https://doi.org/10.1145/3630106.3659017.
- Gajjala, Radhika, et al. "Get the Hammer out! Breaking Computational Tools for Feminist, Intersectional 'Small Data' Research." Journal of Digital Social Research, vol. 6, no. 2, 2, May 2024, pp. 9–26. jdsr.se, https://doi.org/10.33621/jdsr.v6i2.193.

- Tang, Ningjing, et al. “AI Failure Cards: Understanding and Supporting Grassroots Efforts to Mitigate AI Failures in Homeless Services.” The 2024 ACM Conference on Fairness, Accountability, and Transparency, ACM, 2024, pp. 713–32. DOI.org (Crossref), https://doi.org/10.1145/3630106.3658935.
- Groves, Lara, et al. “Auditing Work: Exploring the New York City Algorithmic Bias Audit Regime.” The 2024 ACM Conference on Fairness, Accountability, and Transparency, ACM, 2024, pp. 1107–20. DOI.org (Crossref), https://doi.org/10.1145/3630106.3658959.
- Costanza-Chock, Sasha, et al. “Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem.” Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, Association for Computing Machinery, 2022, pp. 1571–83. ACM Digital Library, https://doi.org/10.1145/3531146.3533213.

POLICIES

Grading Policy

Superior work:	A 4.0 (96-100); A- 3.7 (90-95)
Very good work:	B+ 3.3 (87-89); B 3.0 (83-86); B- (80-82)
Marginally satisfactory:	C+ 2.3 (77-79); C 2.0 (70-76)
Failed:	F 0.0 (0-69)

Policies

The following abbreviated set of policies is especially relevant to this class. Full details on policies and procedures can be found on the Pratt website or in the Office of the Vice President for Student Affairs, Main Building, Lower Level.

Community Standards

All Pratt students, faculty, and staff members are expected to value and uphold the community standards essential to the pursuit of academic excellence and social responsibility. These include expectations for social conduct, academic integrity, non-discrimination, and other policies described in the link above, and apply to all Pratt-sponsored activities, on or off campus.

Academic Integrity

Academic integrity at Pratt means using your own and original ideas in creating academic work. It also means that if you use the ideas or influence of others in your work, you must acknowledge them.

At Pratt,

- We do our own work,
- We are creative, and
- We give credit where it is due.

When students submit any work for academic credit, they make an implicit claim that the work is wholly their own, completed without the assistance of any unauthorized person. These works include, but are not limited to exams, quizzes, presentations, papers, projects, studio work, and other assignments and assessments. In addition, no student shall prevent another student from making their work. Students may study, collaborate, and work together on assignments at the discretion of the instructor.

Examples of infractions include but are not limited to:
The following examples are drawn from the Academic Integrity Code and should be curated or supplemented based on assignments for your course.

1. Plagiarism, defined as using the exact language or a close paraphrase of someone else's ideas without citation.
2. Violations of fair use, including the unauthorized and uncited use of another's artworks, images, designs, etc.
3. The supplying or receiving of completed work including papers, projects, outlines, artworks, designs, prototypes, models, or research for submission by any person other than the author.
4. The unauthorized submission of the same or essentially the same piece of work for credit in two different classes.
5. The unauthorized supplying or receiving of information about the form or content of an examination.
6. The supplying or receiving of partial or complete answers, or suggestions for answers; or the supplying or receiving of assistance in interpretation of questions on any examination from any source not explicitly authorized. (This includes copying or reading of another student's work or consultation of notes or other sources during an examination.)
7. The use of generative artificial intelligence (AI) to produce or to improve work, whether visual or textual, except when called for by an assignment or instructor and acknowledged transparently as one tool among others in the creative process.

The Academic Integrity Standing Committee (AISC) is charged with educating faculty, staff, and students about academic integrity practices. Whenever possible, we strive to resolve alleged infractions at the most local level possible, such as between student and professor, or within a department or school. When necessary, members of this committee will form an Academic Integrity Hearing Board to hear cases regarding cheating, plagiarism, and other infractions described below; these infractions can be grounds for citation, sanction, or dismissal. Detailed procedures are explained in the full version of the Academic Integrity Policy.

Attendance Policy

Consistent attendance is essential for the completion of any course or program. Attending class is part of participation grade, and excessive unexcused absences will lower the grade. You will not be penalized for excused absences.

Please let me know if you are going to miss class. For full details, see Pratt's Attendance Policy.

Academic Support

Multiple academic support resources are available to students across campus:

- For assistance with time management and/or studio, subject, and software tutoring, contact the Student Success Center at success@pratt.edu.
- For assistance with writing assignments, contact the Writing and Tutorial Center at wtc@pratt.edu. The Pratt Libraries can also help with research and citations.
- Academic advisors are also a great resource; students can find their advisor's contact information or schedule an appointment through Starfish.

Accessibility

The Learning/Access Center (L/AC) coordinates access for students with disabilities. Students who identify as having any type of disability are entitled and encouraged to enroll with the L/AC in order to determine and implement reasonable accommodations. Contact the Learning/Access Center at lac@pratt.edu or 718.802.3123 for information or to schedule an appointment.

Title IX and Sexual Misconduct Policy

Pratt Institute is committed to fostering an environment that is safe, secure and free from sex discrimination and sexual harassment, sexual violence, dating and domestic violence, and stalking among all forms of sexual misconduct. The Institute takes prompt and appropriate action to address misconduct, end a hostile environment if one has been created, and prevent the recurrence of a hostile environment. To submit a concern, please use the Title IX and Sexual Misconduct Disclosure Form. For full details, see the Title IX and Sexual Misconduct Policy.

If you inform me of an issue of sex discrimination or sexual misconduct, I will keep the information as private as I can. However, as a faculty member, I am a mandatory reporter of sexual misconduct and required to bring it to the attention of the Institute's Title IX Coordinator, who can also be contacted at titleix@pratt.edu. You can also speak to someone confidentially by contacting our confidential resources in Health Services or the Counseling Center.

Nondiscrimination and Anti-Harassment Policy

Pratt is committed to fostering a welcoming, safe, non-discriminatory, and harassment-free educational, living, and working environment for its community. To submit a concern, please use the Discrimination and Harassment Disclosure Form, which also allows anonymous reports. For full details, see the Nondiscrimination and Anti-Harassment Policy.

Wellbeing

Pratt is dedicated to creating a culture where the entire community can flourish and thrive. Taking time to care for yourself and seeking appropriate support is important to achieving your academic and professional goals. Several resources are available through Starfish, including our Student Advocate and Care Coordinator, who can also be reached at 718.399.4546 or studentadvocate@pratt.edu. If you or anyone you know experiences overwhelming academic stress, persistent difficult feelings, or challenging life events, the Counseling Center can be reached 24/7 by calling 718.687.5356. To schedule a consultation, please call or email therapy@pratt.edu

Starfish

Starfish provides students with contact information for advisors, instructors, department chairs, and support services across the campus. Faculty can also use Starfish to inform students of their progress in class and connect them with resources. Students are contacted by support offices whenever a flag is raised.

Public Safety & Emergency Contacts

The Department of Public Safety provides 24-hour-a-day protection to the campus, including an emergency response guide. Contact Public Safety at security@pratt.edu or 718.636.3540. The Pratt Emergency Alert System is used to send urgent messages to registered mobile devices and emails (faculty, staff, and students have the option to opt-out, which must be renewed each year). You can update your emergency contact information in the Safety section of OnePratt.