**Syllabus**

**Generative AI in the Public Sector: Use, Responsibility, and Regulation**

**PADM-GP 2166**

9/3/2024 - 12/12/2024

4:55 PM - 6:35 PM W
181 Mercer St (Paulson Center) Room 253 Loc: Washington Square

Office hours by appointment, ideally just before class

Art Kleiner (apk1@nyu.edu)

**Class Description:**

During the next few years, generative AI and other forms of artificial intelligence will transform the public sector. They will rapidly increase the productivity of knowledge work, they will expand the types of services governments can offer their citizens, and they will present broad regulatory and auditing challenges. This course covers the opportunities inherent in this technology and the challenges associated with it.

The goal of the course is to equip students entering government and related work to adopt AI responsibly, choosing and implementing tools in effective ways. It offers hands-on practice with prompt generation, direct information on the use of generative AI in government, and then focuses on 7 principles for the responsible use of AI: 1) Risk assessment and management 2) Explainable AI and open systems 3) Reclaiming data rights for people 4) Confronting and questioning the bias inherent in data 5) Accountability in the private and public sectors 6) Organizational systems and structures 7) Creative friction: the organizational culture and practices that favor better outcomes

We look at each of these issues in light the four logics of power affecting the future of AI: business, engineering, government, and social justice. We model personal and group processes to bring these issues safely to the surface, and learn a set of standards and guardrails (a “calculus of intentional risk”) that students can apply to their own work to help assess and avoid harm. This course is set up as a seminar, conducted through dialogue.

**Class Objectives:**

This course is focused on decision-making and public communications skills related to automated, algorithmic, autonomous systems. This class meets at a pivotal moment when these systems are changing knowledge work dramatically, when conventional risk assessment is being reevaluated, and when the temptation is strong to delegate action to automated systems.

Those who complete this course will gain a level of proficiency in the following skills:

1. Gain basic skills in prompt engineering and use of Generative AI for public sector purposes.
2. Evaluate appropriate and inappropriate use of AI tools
3. Conduct original research on technology-related cases, risks and opportunities.
4. Design approaches for AI system explainability and risk management
5. Consider regulatory issues for this fast-moving technological field
6. Propose and evaluate strategic solutions for problems generated by the new technologies.
7. Participate in “creative friction”-style discussions which bring together diverse perspectives toward a common objective.

The tone of class discussion is open and congenial, focused on complex causes and robust solutions.

**Course Structure**

**This course is structured around regular Brightspace discussions and four assignments:**

1. A prototype for a use case of Generative AI in government
2. A case study of a real-world AI dilemma in the public or not-for-profit sector.
3. A proposal for standards or guidelines.
4. A final project building on an earlier assignment.

In addition, class discussion will include written discussion posts in Brightspace as well as face-to-face lecture and discussion.

**Course Schedule**

1.     Weds, Sept. 4: Introduction

**Objective:** To introduce the course, provide two key concepts, and develop a collective sense of the opportunities for learning this topic.

**Class introductions:** What is your background relevant to AI in government? What do you hope to gain from this class? What aspects of AI use have you observed first-hand?

Introduction to the two major frameworks:

1. The four logics of power:

|  |  |  |
| --- | --- | --- |
|   | **Private** | **Public** |
| **Individual** | Engineers & data scientists / Labor | Society and social activists |
| **Institutional** | Corporations | Government regulators |

**Seven principles:**

1.Do No Harm.

2.  Create Code that Speaks for Itself, Explainable AI.

3.  Reclaim Data Rights for People.

4.  Use Machine Learning for Human Understanding.

5.  Question and Confront Negative Biases.

6.  Hold All Stakeholders Accountable.

7.  Embrace Creative Friction, Involve the Community of Interest.

2.     Weds, Sept 11: Issues of AI use

**Objective:** To explore the current state of affairs in AI for the public sector

**Reading:** *The Genius Makers,* Chapters 2, 3, 13 and 14.

**Brightspace Discussion:** AI in the News. Find a case related to responsibility in technology, from news sources. Why is this story relevant to AI in government? What logics are in play – public, private, social justice and engineering?

3.     Weds, Sept 18: Af Malhotra discussion

**Objective:**Introduction to AI, DEI and the use of tools

4.   Weds, Sept 25: Using Library and AI tools

**Objective:**Introduction to library work and prompt engineering

**Guest:**Carol Choi, NYU library: re research opportunities for Wagner

**Due in Brightspace Discussion:** What is one potential use case of Generative AI (ChatGPT, Claude, Dall-E, Google Gemini, etc.) in government that could be put in place today?

Practicum on Prompt engineering

5.     Weds, Oct 2: Prototypes

**Due: Assignment A – Prototype proposal**

**Objective:**To discuss our blog posts about government prototypes.

**Group exercise:** In working groups, take up one application and consider:

1. What is the benefit or purpose of this application?
2. What questions or concerns would you have about it?
3. How would you develop it?
4. What outcomes (productivity, new services, etc.) would you want to track?
5. What risks do you see?
6. Whose feedback would you need and how would you produce it?
7. How would you curtail those risks? What guardrails might you put in place?
**Lecture:**Basic process for AI System development: PoCs and MVPs

6.     Weds, Oct 9: Prompt Engineering 2 •  AI System Risks and Opportunities

             Reading: AI Dilemma, Introduction and Chapter 2.

7.     Weds, Oct 16: Risks and Challenges

**Objective:**To explore the issues and challenges related to impact of AI systems on people.

**Class discussion:**To talk about case studies and how things might have happened differently.

8.     Weds, Oct 23:  Case Studies

**Objective:** To learn from one anothers’ case studies

**Presentations:** 5 minutes on each case study

            What happened?

            What were the consequences?

            What lesson do you draw from it?

9. Weds, Oct 30: Bias

**Objective:** To discuss the prevailing issues of ownership rights, bias, and human response to AI overreach.

**Readings:** *The AI Dilemma,* Chapter 5; *Unmasking AI,* Chapters 11, 13, 18

**Brightspace Discussion:** Describe one meaningful example of bias in AI: in technology, at NYU, in the world of government.

**Group exercise:**How do we effectively question and confront negative bias in or with AI systems?

**Opening:**Video-based exercise

Joy Buolamwini: [On Having to Wear a White Mask](https://www.youtube.com/watch?v=gkbNH39QE0Q) [Gender Shades Video](https://www.youtube.com/watch?v=gkbNH39QE0Q)

10.     Weds, Nov 6:  Reclaim Data Rights for People

**Objective:** To discuss the ownership of data and rights to data.

**Readings:** The AI Dilemma, Chapter 4

**Video:** “Mapping Data Flows,” John Battelle: <https://www.youtube.com/watch?v=b5H9tUSqnas>

**Brightspace Discussion:**How much is your personal data worth, and to whom?

**Small Groups:** Design a mechanism for controlling data. How would it work?

**Plenary discussion:** What future of data control is desirable? What future of data control is realistically possible? What would have to happen to generate a future we prefer?

11. Weds, Nov 13:  Public Accountability and Regulation

**Objective:** To explore the concept of accountability for risk and AI regulation.

**Readings:**The AI Dilemma, Chapter 6; Additional reading on the EU AI Act from ForHumanity

**Plenary discussion:** Government accountability for the US of AI: How to regard responsibility for the systems we create, use and regulate.

**Exercise:** The Moral Machine – Who should decide who should be saved?

**Assignment C due:**Proposal for standards or guidelines

Presentations: Each student presents their proposal; discussion

12.  Weds, November 20: Accountability, Transparency and Explainable AI

**Objective:** To look at up-to-date concepts of “opening the closed box” .

**Readings:**The AI Dilemma, Chapter 3

**Brightspace Discussion:** What level of openness and explainability is appropriate for a government agency AI system? What about a private sector AI system? What is the right balance between the public right to know about AI's impact and the need for organizational privacy, trade secrets, and managerial effectiveness?

**Small group exercise:**Each group proposes one rule – for managing transparency. (Use ChatGPT to help develop ideas.)

13.  Weds. December 4: Creative friction and organizational learning

**Objective:** To look at the management and organizational cultural issues relevant to AI.

**Reading:***The AI Dilemma,*Chapters 7 and 8.

**Brightspace Discussion:** Where have you seen creative friction in action? What made it happen? What was the community of interest involved in this?

**Video:** Tightly and Loosely Coupled Systems: Karl Weick, https://www.youtube.com/watch?v=MrH0kypHelA

**Exercise:** Simulated discussion using creative friction techniques

Students take turns playing the roles of each of four points of view:

            Engineers and data scientists: How responsible were we?

                                                            What could we do differently?

                                                            What support do we need?

            Government and regulation:    What harm was done?

                                                            How do we enforce accountability?

                                                            What pressures do we feel?

            Corporations and investors:   What is our responsibility?

                                                            What are we trying to achieve?

                                                            Can we regulate ourselves?

            Social activists:                        What is the issue at stake here?

                                                            What should those in power do?

                                                            How should this be different?

During the conversation, each student writes down two things that each of the other participants said. After the conversation, they check on whether these remarks were correctly interpreted.

14.  Weds, December 11: Final project presentation

**Due: Assignment D: Final project**

**Assignments**

**A.**    **Prototype for a use case of Generative AI in government**

Using a GenAI program (like ChatGPT), develop a proposal for an app or an actual prototype, showing how this tool could be used to address government agency issues, interact with citizens, etc.

**B.**    **A case study of a real-world AI dilemma in the public or not-for-profit sector.**

Prepare of a case study of a real-world use of AI in a public sector context, where there was an issue of responsibility or unintended consequences. Read or watch at least 3 credible sources on this case and share the case story in class.Be prepared to make a 7-minute verbal and visual presentation.

Turn in a 1-2 page writeup.

Sources may include: 1) media generated by the government or vendor; 2) independent reporting or observation; 3) scholarly or academic documentation; 4) interviews or primary source commentary.

At least one source should be independent of the government and vendor interests.

Answer these questions:

* What was the intent of this use of AI?
* What were the consequences: positive and negative?
* What did or would the four logics of power say about the case?
* Who - if anyone - was held accountable, and how?
* What has been the aftermath of the case?
* What did you learn that seems applicable to other situations?
*

**C.**    **Proposal for Standards or Guidelines Regarding Responsible AI**

Proposal to some government body about regulatory guidelines, based on what you’ve learned in this class and your own research. Cite sources. It can be an expansion or refinement of proposals that already exist, or it can be original. Examples of possible projects include: A proposal for giving people control over their data, standards for mitigating bias, requirements for involving people affected by AI in its design, etc. There should be a rationale for why this provision is important, an estimate of how people will be affected, and an analysis of who might oppose this proposal, and on what grounds.

**D.**    **Final project: Expansion of A, B, or C**

* A final project building on an earlier assignment.

Guest speakers (pending confirmation of particular dates):

* Juliette Powell, coauthor of The AI Dilemma
* Af Malhotra, founder of Diversity Equity AI (DEAI), UK
* Becky Duane, Data and AI for Social Impact, IBM
* Ryan Carrier, founder, ForHumanity