# PADM-GP 4152
# Ethics, Public Policy and Emerging Technology
# January 2022

## Instructor Information

- Amy Auton-Smith
- Email: autonsmith@nyu.edu
- Office Hours: Online, by prior appointment.

## Class Meeting Times/Locations

- Class Meeting Times: Fridays 9am–5pm, January 7th-21st
- Class Location: Silver, Room 507

## Course Prerequisites

- None

## Course Description

Commentators have noted that the use of AI technologies in government and nonprofits is lagging and part of the reason for this could be that many public service leaders do not feel equipped to make decisions on the use of these technologies. This course aims to equip current and future leaders with an understanding of how emerging (or AI type) technologies are created and how they work, and to provide a foundation for thinking strategically and ethically about their use in a variety of settings.

This is a non-technical class: no coding knowledge is required.

This course will look at practical applications of AI technologies, with a broad focus on public service and nonprofit use-cases. We will have a focus upon equipping public service leaders to take an informed and strategic stance when considering the use of these technologies in their organizations. Students will have the opportunity to undertake independent study on an area of interest to them.

We will start with a look at key concepts and the fundamentals of how 'AI' works. We will take a detailed look at some main types of AI, start to consider how they function and how they can be used to deliver organizational goals. Then we will move on to look at the wider societal implications of AI and dive into ethics, bias, and the future of public service work in an increasingly AI-driven world.

Finally, we will take a high-level look at some of the challenges facing policymakers and start to investigate how emerging technologies are managed and regulated across the globe.

This course complements PADM-GP.4700 Topics in Public Policy – Data, Evidence, Ethics and Bias in an AI World.

## Course and Learning Objectives

After completing this course, students will:
- Be equipped to think critically about questions of ethics in relation to AI/emerging technologies
- Understand the fundamentals of how the main types of AI are created
- Be able to identify strategic use-cases for these technologies in a variety of settings
- Have reflected upon the future of work in a technology-enabled public service world
- Have engaged critically with AI-related public policy challenges and issues nationally and globally

## Brightspace class site

Each class will feature a variety of readings from a wide range of sources. Some readings are posted on our class site, and some are weblinks as listed in the syllabus. If any of the weblink-only readings are non-functioning or the article is behind a paywall, please let me know as soon as possible.

All announcements, resources, and assignments will be delivered through the class site. I may modify assignments, due dates, and other aspects of the course as we go, with advance notice provided as soon as possible through the course website.

All assignments **must** be submitted through the class site. Exceptions will only be made where notification that a student has good reason for not being able to submit through Brightspace has been given in advance of the submission deadline.

## Academic Integrity

Academic integrity is a vital component of Wagner and NYU. All students enrolled in this class are required to read and abide by [Wagner's Academic Code](). All Wagner students have already read and signed the [Wagner Academic Oath]().

## Henry and Lucy Moses Center for Students with Disabilities at NYU

Academic accommodations are available for students with disabilities. Please visit the Moses Center for Students with Disabilities (CSD) website and click on the Reasonable Accommodations and How to Register tab or call or email CSD at (212-998-4980 or mosescsd@nyu.edu) for information. Students who are requesting academic accommodations are strongly advised to reach out to the Moses Center as early as possible in the semester for assistance.

## NYU's Calendar Policy on Religious Holidays

NYU's Calendar Policy on Religious Holidays states that members of any religious group may, without penalty, absent themselves from classes when required in compliance with their religious obligations. Please notify me in advance of religious holidays that might coincide with exams to schedule mutually acceptable alternatives.

## Assignments and Evaluation

- ○ Preliminary Memo:          10%
- ○ Participation:          30%
- ○ Light-touch Literature Review:    30%
- ○ Technology Exploration Paper:    30%

## Format – all written assignments

All written assignments must be in **pdf format**, 12-point Times New Roman font (or equivalent), with 1-inch margins and **double-spaced**. Assignments must be submitted via Brightspace (not by email or in hard copy).

### Preliminary Memo (10% of total grade)

You are asked to review the readings for Class 1: Sections 1.1 and 1.2 and to prepare a 2-page (max 500 words) preliminary memo setting out your reflections prompted by the readings. No prior knowledge is required and this memo aims to start you on a path of thinking critically about technology use cases. You may like to pose some questions in this preliminary memo, alongside your reflections.
**Grading**: This assignment will be assessed on the following basis (5 points each): (i) evidence of having read and engaged critically with the readings; (ii) your reflections on any issues or questions that you may identify in relation to the use of emerging technology.
**How to submit your work:** Please submit your memo via Brightspace no later than 5pm the day before Class 1. You do not need to bring a hard copy to class, but attendees should be ready to share their reflections with class colleagues.

## Participation (30% of total grade)

Participation is awarded based on preparation, attendance and engagement in class discussions (and optionally also via the forum on the class site). This intensive class requires that all students participate actively and you are expected to do readings prior to class, attend class, and contribute to the discussions. **Grading:** Participating by being present, volunteering your thoughts, actively engaging with colleagues and sharing insights and reflections is expected. Missing or being more than a few minutes late for class will negatively impact your participation grade. Extraordinary circumstances include religious observance and illness, but you must give notice via email in advance (religious observance) or as soon as possible (illness). If you foresee missing a class due to personal obligations, please notify me via email.

## 'Light-touch' Literature Review (30% of total grade)

The literature review will be undertaken in the period between Class 1 and Class 2 and will be due the day before Class 2 at 5pm.

This **brief** literature review must be no more than 750 words and should provide your reader with a high-level overview of some of the work and findings that have been done in relation to a topic. Your aim is to give your reader a clear and concise picture of a representative sample of the research, knowledge and opinion in that area.

Please note the 750-word limit: you are **not** being asked to create a full and exhaustive review of all relevant research. Instead, this assignment aims to start you on a path of independent inquiry and evaluation in relation to the use of an emerging technology in a given sphere and to ensure a degree of familiarity with the state of knowledge in your chosen area.

You are strongly recommended to choose a topic relevant to a use-case that you could use for your technology exploration paper (see below).

The literature review assignment requires you to dive into not only the set readings for the class, but also to undertake an element of your own research on a chosen topic. You may choose a topic from the list below, or you may ask for approval to choose an independent topic, however, your literature review must focus upon (i) ethical and/or public policy considerations; (ii) that do or could arise in relation to the use of a chosen emerging technology in the topic area; (iii) considered from a public service, NGO or non-profit perspective. If you wish to select your own topic, please make sure this is approved by **me no later than the end of Class 1**.

Topics: (a) law enforcement, (b) healthcare, (c) education, (d) workforce management

**Grading**: This assignment will be assessed on the following basis:
    (i)        evidence of having engaged critically with the chosen topic (5/30);

(ii)     inclusion of a references appendix (not included in word limit) that sets out at least 8 full citations, of which at least 5 must be relevant new or outside resources that are not in the reading lists for the course (5/30);

(iii)    a clear and concise paper that includes as a minimum (1) an introduction, (2) a main body with a clear and comparative flow of ideas deriving from the research considered and (3) a conclusion (10/30);

(iv)    reflections on some of the ethical and public policy considerations that could arise for public service organizations, NGOs or non-profits considering using AI in the topic area (5/30);

(v)     no more than 750 words, submitted on time, in pdf format, via Brightspace (5/30).

Extra credit is available (up to 5 additional points) for presenting your reflections on future research that is needed to complete our understanding of a topic: a gap analysis of the research in your chosen area. Please note that this must be included in the 750-word limit.

**How to submit your work:** Please submit your Literature Review in pdf format via Brightspace no later than 5pm the day before Class 2. You may choose to use the content template provided, but this is not required.

## Technology Exploration Paper (30% of total grade)

For your technology exploration paper, you will create a high-level overview of the technological, ethical and public policy considerations of a selected use-case for an emerging technology in an organization or setting of your choice and create a compelling analysis for or against the use-case you explore.

We will spend some time in class going over the expectations and requirements for the assignment and you will have the opportunity to work with class colleagues in refining your ideas in Class 2. Please submit one single pdf file via Brightspace and please do not exceed 750 words.

You should include references (with citations) to support your arguments. Please include a references appendix that will not count towards the word limit.

**Grading**: This assignment will be assessed on the following basis:
(i)     clear and concise description of the emerging technology use-case you have identified (5/30);

(ii)    analysis of the benefits and risks of using the technology in the way you propose (5/30);

(iii)   analysis of feasibility for the organization or setting you identify (dive into whether the technology is suitable for use in that setting and consider the requirements for building or implementing the chosen technology) (5/30);

(iv)    clear expression of the ethical and public policy considerations to be taken into account, along with your opinion on whether the use-case is appropriate in the setting you explore (10/30);

(v)     no more than 750 words, submitted on time, in pdf format, via Brightspace (5/30).

**How to submit your work:** Please submit your paper in pdf format via Brightspace no later than 5pm on the due date.

## Overview of the Semester

| Item | Topic | Preparation |
|---|---|---|
| Preliminary Memo | See assignment instructions in syllabus and class site.<br>Readings in Sections 1.1 and 1.2 | Due the day before Class 1 at 5pm. |
| Class 1 | **Introduction**. What is emerging tech? Overview of examples: ML, Robotics, NLP.<br>**Use-cases in public service** - opportunities and the balancing-act of benefit and cost<br>**Ethics** – intro to ethics and emerging tech<br>**Bias** – why it's important to understand tech bias and some mitigations<br>**Guest speaker**<br>**Case study and discussion** | Readings for all sections of Class 1<br><br>Please select a topic for your Literature Review (and likely also your Technology Exploration Paper) before the end of Class 1. |
| Literature Review | See assignment instructions and guidance template. | Due the day before Class 2 at 5pm. |
| Class 2 | **Future of work**<br>**Public policymaking** and the governance and regulation of emerging tech<br>**Global perspectives** – a review of public policy on emerging tech<br>**Guest speaker**<br>**Discussion of Technology Exploration Paper proposals**<br>**Synthesis and close** | Readings for all sections of Class 2<br><br>Please come prepared with a topic for your Technology Exploration Paper assignment. |
| Technology Exploration Paper | See assignment instructions and refer to instructions given in class. | Due date as per Assignments in class site. |

# Readings

There is no set textbook for this class, however, by way of further reading you may wish to consider purchasing or reading online the following:

1. Crawford, K., "Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence", Yale University Press, New Haven/London, 2021. Available online here;
2. Coeckelbergh, M., "AI Ethics", MIT Press, 2020. Available online here;
3. O'Neill, C., "Weapons of Math Destruction", Broadway Books, New York, 2017. Available online here.
4. Foster, I. et al, "Data Science Methods and Tools for Research and Practice", Coleridge Initiative, 2nd Ed. Available online here.

## CLASS 1

## Section 1.1: Introduction. What is emerging tech?

### Readings

1. Moltzau, A., "Artificial Intelligence and Nonprofits", Towards Data Science, Aug 31, 2019.
2. Agrawal, A. et al "What to Expect from Artificial Intelligence" MIT Sloan Management Review, Vol 58, Iss. 3, Spring 2017, 23-27.
3. **Machine Learning**: McNulty, K., "How does Machine Learning Work?", Towards Data Science, 2018. And watch this brief video.
4. **Robotic process automation**: watch this brief video.
5. **Natural Language Processing**: Review these examples summarized by Tableau.

6. [Optional] Lagi, M., "Natural Language Processing – business applications", Emerj, 2018.
7. [**Optional**] Browse the articles on 'Monday' on the Oxford Insights website here: https://www.oxfordinsights.com/aiinaweek/monday.
   Suggested areas of focus are: **Introduction**: "What is Artificial Intelligence Exactly?", **Why it Matters**: "What AI is, and isn't" and **The Turing Test**: the 5-minute Ted-Ed video on "The Turing Test" linked on the page.

   **Guest Speaker - TBC**

## Section 1.2: Use-cases - opportunities and the balancing-act of benefit and cost

### Readings

We will use predictive policing as a use-case example for in-class discussion.

1. Apolline, R., "Ethics, AI and Predictive Policing", Blog: The Security Distillery, July 23, 2021.
2. Advertorial for Appen (Ai service provider): "AI in Police Work", May 17, 2021.
3. O'Neill, C., "Weapons of Math Destruction", Broadway Books, 2017. Ch 5 "Civilian Casualties.

The following are optional readings on some use-case examples and you are encouraged to find other examples independently. Assignment pre-approved topics include: (a) law enforcement, (b) healthcare, (c) education, (d) workforce management.

4. [Optional] Gita (jee-ta) Robot Cart and Astro. Consider ageing, disability and family use-cases.
5. [Optional] Tesla's 'Autopilot' function. Consider mass transit, public safety, climate change. Consider also automated decision-making risks.
6. [Optional ] Chatbots in use in customer service. **Warning, the following link contains offensive material:** And 'Tay' as a cautionary example as part of an article looking at issues with chatbots.
7. [Optional] Da Vinci and robotic surgery. Consider value for money, cyber security, impact on physician training and skills, accessibility.
8. [Optional] AI Now Testimony on Predictive Policing: February 20, 2020.
9. [Optional] Agarwal, S., "Surveying how the public sector is approaching an AI-enabled future", Deloitte Insights, November 8, 2019.
10. [Optional] Truswell, E., "How AI could help the public sector", HBP, Jan 25, 2018.

## Section 1.3: Ethics – intro to ethics and emerging tech

**Readings**

1. The Trolley Dilemma and Self-Driving Cars:
    a. D'Olimpio, L., "The trolley dilemma: would you kill one person to save five?", The Conversation, June 2, 2016.
    b. Bradshaw-Martin, H., "Could your self-driving car choose to kill you?", Science Focus, November 17, 2020.
    c. 'Play' the Moral Machine scenarios and think about whether the data collected by this exercise should play a part in real-world design.
2. Dhasarathy, A. et al, "When Governments Turn to AI: Algorithms, Trade-Offs and Trust", McKinsey Special Report, October 2020.

3. [Optional] Walsh, M., "Algorithms Are Making Economic Inequality Worse", HBR, October 22, 2020.
4. [Optional] Stapleton, A., "Machines that Speak and Write will make Misinformation Worse", HBR, May 14, 2019.
5. [Optional] Stone, A., "Ethics in the Balance: AI's implications for Government", GovTech Article, June 26, 2020.
6. [Optional] Watch the two short videos (total 5 mins) on this Accenture site: "Human and Machine: AI in Public Service" and "The Ethics of AI".

7. [Optional] Hofmann, B., "Progress bias versus status quo bias in the ethics of emerging science and technology", Bioethics, March 2020, Vol 34, Issue 3, pp252-263.

## Section 1.4: Bias – why it's important to understand tech bias and some mitigations

**Readings**

1. Coeckelbergh, M., "AI Ethics", MIT Press, 2020. Ch 9 "Bias and the Meaning of Life".
2. Buranyi, S., "Rise of the Racist Robots", The Guardian, August 8, 2017.
3. Metz, C., "Who is making sure the AI Machines aren't racist?", NYT, March 15, 2021.

4. [Optional] Briefly review the Berkeley HAAS compendium of examples of bias arising in AI.
5. [Optional] Feast, J., "4 Ways to Address Gender Bias in AI", HBR, November 20, 2019
6. [Optional] Chang, F., "To Build More-Inclusive Technology, Change Your Design Process", HBR, October 19, 2020.
7. [Optional] Feast, J., "Root Out Bias at Every Stage of Your AI-Development Process", HBR, October 30, 2020.

## Section 1.5: Case study exercise

**Readings**

1. Case Study 1 materials – TBC

[IoT Case Study - Please read the Smart Cities case study information sheet and the article entitled 'Thinking Ethically'. Start to think about your response to the questions posed.]

## CLASS 2

## Section 2.1: Future of work

**Readings**

1. Crawford, K., "The Atlas of AI: Power, Politics and the Planetary Costs of AI", Yale University Press, 2021. Ch 2 "Labor".
2. McKinsey Survey Report, "The State of AI in 2020", November 17, 2020.

## Section 2.2: Public policymaking and the regulation of emerging tech

**Readings**

1. Crawford, K., "The Atlas of AI: Power, Politics and the Planetary Costs of AI", Yale University Press, 2021. Ch 6 "State".
2. Coeckelbergh, M., "AI Ethics", MIT Press, 2020. Ch 10 "Policy Proposals" and Ch 11 "Challenges for Policymakers".

3. [Optional] Chowdhury, R. and Sloane, M., "The risks of using AI for Government work", MarshMcLennan Brink News, November 30, 2020.
4. [Optional] Downes, L., "A Measured Approach to Regulating Fast-Changing Tech", HBR, October 23, 2020.
5. [Optional] Fox Cahn, A., "The first effort to regulate AI was a spectacular failure", Fast Company, November 26, 2019.
6. [Optional] Ghosh, D., "What Will Tech Regulation Look Like in the Biden Era?", HBR, December 17, 2020.

## Section 2.3: Global perspectives – public policy globally on emerging tech

**Readings**

1. Crawford, K., "The Atlas of AI", ibid. Ch 7 "Conclusion".
2. Please browse the OECD AI Policy Observatory as a reservoir of information on comparative global activity on AI. Spend a few moments comparing activity across various countries, for example China, Russia and the US.

Please choose **one** of the following examples, or a non-US country/region of your choice, and take a deeper look at AI policy and regulation in that region.

3. Canada: "Responsible use of AI" website and news postings.
4. Mexico: "Mexican National Agenda of AI" website and linked resources. Please make use of tools such as Google Translate to read the Spanish text if needed.
5. China: New Generation AI Development Plan, July 20, 2017. Please make use of tools such as Google Translate to read the Chinese text if needed.
6. [TBC - Non-state actors and their influence in AI]

## Section 2.4: Case Study Exercise

**Readings**

1. Case Study 2 materials – TBC

[Hiring By Machine Case Study - Please read the Hiring By Machine case study materials and start to consider the questions posed.]

## Section 2.5: Technology Exploration Paper, Synthesis and Close.

Guest Speaker - TBC

No set readings for this section.