Artificial Intelligence and Non-Proliferation

Elliott School of International Affairs
George Washington University

Instructor: Dr. Carol Kuntz
Semester: Fall 2020
Office Hours: Thursday, 1:30 – 2:30 and by appointment
Class Meeting Time: Tuesday and Thursday, 11:00 am – 12:15 pm
Syllabus Version: Final (1-22-21)

Course Description:

Artificial intelligence is a transformative technology whose application often creates new and acute public policy problems amidst the many benefits it provides. This class will study artificial intelligence and three of its applications – purposeful genomic manipulation, AI-enabled military capabilities, and facial recognition and the surveillance state – with a particular focus on seeking promising strategies to mitigate or manage in the international context the public policy problems that it creates or worsens. The class will examine at a general level how artificial intelligence works and how it can go wrong due to, for example, the poor quality or limited availability of data sets used to train the machine learning algorithms.

The class will seek to identify public policy problems created by artificial intelligence and consider proposals to manage those problems. These efforts will confront a classic problem: what are sources of evidence that a proposed policy would work as predicted? This class will examine one possible source of evidence: Classic non-proliferation tools – such as international inspections to ensure peaceful use of dual-use technologies – used in the post-World War II non-proliferation framework that sought to control chemical, biological and nuclear weapons. The class will identify the core mechanisms and enabling conditions that helped these classic non-proliferation tools “work” in the past and consider whether these enabling conditions would be present if these core mechanisms were applied to manage public policy problems created by Artificial Intelligence. The class will consider whether broader changes in the strategic environment – particularly the rise of China – invalidate the utility of these post-World War II tools.

Learning Outcomes:

• Students will learn at a general level how artificial intelligence works.
• Students will learn about policy issues inherent in artificial intelligence and three critical applications – purposeful genomic manipulation, AI-enabled military capabilities, and facial recognition and the surveillance state.
• Students will learn about classic post-World War II non-proliferation tools and their core mechanisms as well as the enabling conditions that contributed to their success.
• Students will consider whether there are fundamental changes underway in the strategic environment – particularly the rise of China – and whether these changes invalidate the utility of post-World War II tools or require, at least, their repurposing to serve different strategic objectives.

Organization of the Class:

The class (and its canvas site) is organized into seven “modules:"

1. Syllabus et al

2. Artificial Intelligence
   • Essential Questions: The dominant form of Artificial Intelligence used outside the laboratory is machine learning. What are the implications of the reliance of machine learning on data bases which may not be representative or may contain various sorts of bias? Or which may provide insurmountable monopoly powers or strategic advantage?

3. Strategic Environment
   • Essential Questions: Is the strategic environment changing in a fundamental way that invalidates the utility of post-World War II tools including classic non-proliferation tools? The principal source of change is the rise of China. Are China’s objectives reconcilable with the current rules-based international environment?

4. Non-Proliferation Tools
   • Essential Questions: Can a close examination of post-World War II non-proliferation tools provide useful evidence to inform the crafting of public policy proposals on emerging technology, including applications of Artificial Intelligence? What insights can be derived from identifying the core mechanism and enabling conditions for these tools? Can these tools be repurposed to serve other objectives in a new strategic era?
5. **Purposeful Manipulation of Genomes**

- **Essential Questions:** The convergence and low entry costs of advanced biotechnologies and artificial intelligence enables humans to circumvent evolution to alter the fundamental nature of plants and animals, including humans. Can humanity shape this future or merely endure its consequences, however disquieting they may be?

6. **AI-Enabled Military Capabilities**

- **Essential Questions:** Does the literature on AI-enabled military capabilities overemphasize the risk of spiral model failures? Are deterrence failures or spiral model failures the greater risk with AI-enabled military capabilities in this strategic environment? What strategies could minimize each of these risks?

7. **Facial Recognition and the Surveillance State**

- **Essential Questions:** Facial recognition technology, powered by artificial intelligence, is eliminating privacy. Unlike other applications of artificial intelligence, this issue has received very different policy responses in the United States, the European Union, and China. What are the risks and benefits of this new capability? Which regulatory regime holds the greatest promise for maximizing benefits and minimizing risks?