Hendrik Schopmans, M.Phil.   **Office Hours:** Wed., 10-11am
WZB Berlin Social Science Center   (please arrange via email)
Reichpietschufer 50, 10715 Berlin
Room B310
h.schopmans@fu-berlin.de

The Global Politics of Artificial Intelligence

Otto-Suhr-Institut, Freie Universität Berlin
Pro Seminar, SoSe 2020/2021
Tuesdays, 10.00-12.00

==**Please note: This is a preliminary syllabus, the individual readings for each week may still be subject to changes.**==

I. Course Description and Objectives

When reading the news these days, it seems that artificial intelligence (AI) is everywhere. AI translates our words into foreign languages on GoogleTranslate and recommends new songs on Spotify. It diagnoses cancer, steers autonomous cars, and decides on jail sentences. One day it may be employed in "killer robots" on the battlefield and, if you believe some people, be our most powerful weapon in the fight against climate change. Since AI technologies are poised to impact almost every facet of our lives, politicians across the world have taken notice. In fact, AI is increasingly becoming the concern of international politics. Some see AI as a source of power: Russian President Vladimir Putin, for instance, remarked that "whoever becomes the leader in this sphere will become the ruler of the world." Others are more worried about the possible harmful consequences of AI, from autonomous weapon systems to hyper-efficient surveillance states, and call for the global regulation of AI. Overall, the emergence of AI as a global issue raises a number of questions: First of all, what exactly is AI and why does it matter for global politics? Why do states care about it? What problems does it pose, and can AI be regulated? And which role do corporations, researchers, civil society and other non-state actors play in this process?

This course offers students an opportunity to explore this emerging area of international politics. Its objectives are twofold: First, it aims to equip students with foundational knowledge about artificial intelligence, its history, and the promises and risks associated with it. Second, the course invites students to think about the global political implications of AI with the help of International Relations (IR) theories. Each week, students will engage with a new theoretical perspective and use it to explore a specific case study related to AI. Upon completing the course, students should

be able to evaluate new developments in the field of AI and critically discuss their political implications.

II. Learning Outcomes

Upon completion of the course, students should be able to:

- ✓ Critically discuss developments in the field of AI, drawing on a basic understanding of its technical aspects, historical development, and contemporary application domains
- ✓ Understand key concepts and theories from the discipline of International Relations
- ✓ Apply theories and concepts to shed light on specific case studies
- ✓ Critically reflect on the merits and possible shortcomings of different theoretical approaches for understanding international politics
- ✓ Compose and structure a research paper consisting of a clear research question, a theoretical discussion, and an empirical case study

III. Teaching Methods and Class Structure

Due to COVID-19, all classes will take place online. In the course's **first block (Weeks 1-4)**, students will familiarize themselves with key concepts from the field of artificial intelligence. Students are expected to read and prepare the required materials in advance; the class sessions will consist of short group discussions and -exercises to clarify, and critically engage with, concepts and ideas from the literature. Depending on the session, I will make use of short presentations, featuring visual and audio materials, to supplement the discussions.

Classes during the **second block (Weeks 5-12)** will focus on exploring AI-related case studies with the help of IR theories. They all follow the same structure: The first half of the class will be reserved for group presentations and discussions. More specifically, each week one group will introduce the case study designated for that week to the rest of the class in a short presentation (10-15 minutes). Afterwards, the same group is asked to lead a discussion with the rest of the class, aimed at critically exploring the possible benefits and shortcomings of IR theories in illuminating the case. In the second half of the class, I will briefly summarize and contextualize the theoretical approach designated for that particular class (15-20 minutes). The remainder of the session will be used to discuss the theoretical approach more generally and address open questions.

The **final block (Weeks 13 and 14)** is reserved for feedback on ideas for the term paper. Students are asked to prepare a short outline of their paper idea prior to the class. Depending on the number of participants, we will either discuss these ideas with the whole group or split the class into small groups. Here, students will briefly present their ideas and provide feedback to each other based on the paper requirements (see below).

IV. Requirements and Assessment

All students are expected to come to class prepared and actively participate in discussions. For each session, **students are required do the core reading** indicated for that week. Reading

beyond the core reading and introducing new material in class discussions is strongly encouraged. Students aiming at a participation certificate ("Teilnahmeschein") are required to (1) continuously attend and actively participate in classes and (2) do a group presentation and structure the subsequent discussion.

Students who aim for a graded certificate ("Leistungsschein") are assessed along two dimensions:

➢ Continuous attendance and active participation in classes
➢ Group presentation and discussion (30%)
➢ Full-length research paper (70%)

More specific requirements for each assessment are outlined below.

1. Group Presentation and Discussion

Group presentations will take place between Weeks 5 and 12. Groups will be assigned during the first meeting of this course, with group size depending on the total number of people in the class. Each group has two main tasks: First, students are asked to prepare a short presentation (max. 15 minutes) to introduce the case study assigned for that week. The use of PowerPoint to support this presentation is encouraged. Importantly, in their presentation of the case, students are expected to go beyond the core reading and consult additional sources. While groups are free with regard to the structure and content of the presentation, the goal is to give the rest of the class a sound overview of the case, as this will form the basis of the subsequent discussion. Some questions to consider may be the following:

➢ **Background**. What is the case about? What is the current status and what are some of the most important historical milestones?
➢ **The role of AI**. What type of AI is involved in this case? Which elements of AI are contributing to the problem at the heart of this case?
➢ **Actors**. Who are the main actors involved and what do they want?
➢ **Institutions**. Where do politics take place?

After the presentation, the group will then be asked to structure and moderate a group discussion (max. 30 minutes) with the rest of the class. The goal of the discussion is to explore the case together by considering the theoretical approach introduced that week. Possible merits and potential shortcomings of the approach should be discussed. Keeping this goal in mind, students are again free in how they design the group discussion. Some guiding questions may be:

➢ What is the theory about? What are its assumptions about international politics, and what kind of propositions does it generate?
➢ Why is the theory relevant for the case?
➢ Does the theory explain the case well? Why or why not?
➢ What do you think will happen in the case in the future? Does the theory help give us answers to this?

2. Research Paper

The goal of the research paper is to apply a theory or concept from International Relations to a specific case related to either AI or the field of emerging technologies more broadly. Students are free to choose their topic, but I would encourage you to discuss it with me at least once.

- ➢ The paper should contain all of the following elements: (1) an introduction; (2) a theoretical section; (3) a case study section; and (4) a conclusion.
- ➢ In the **Introduction**, the student should briefly provide some background to the case; pose a clear research question; and outline the paper's main argument.
- ➢ In the **Theory** section, the student should discuss the theoretical approach that is used to answer the research question. How does the theory relate to other theories in the field of IR? What assumptions does the theory rely on and what propositions does it generally generate? Why is the theory well-suited to explain the case at hand?
- ➢ In the **Case Study** section, the student should apply the theoretical propositions to the case at hand. In other words, the student should answer the paper's original research question by relating empirical observations from the case to the theory's concepts, assumptions, and propositions.
- ➢ Finally, the **Conclusion** should summarize the paper's main findings. Rather than simply repeating the content of previous sections, however, the student should use this space to critically reflect on the theory. What aspects of the case has the theory helped to explain? Which other aspects did the theory fall short of explaining, and why? Based on your findings, what would be interesting questions or strands of research to pursue in the future?
- ➢ Students can use literature from this syllabus in their papers, but **it is expected that the theoretical section engages with literature beyond this course's core texts.** In the empirical section, students are free to engage with a variety of sources (government reports, policy papers, parliamentary discussions, newspaper articles, etc.). Please note that Wikipedia will not be accepted as a scientific source.

Technical requirements

- ➢ Length: 3000 words (+-10%) including in-text citations, excluding bibliography
- ➢ Citation style: The preferred citation style is Chicago Citation Style with in-text references in author-date format. Other citation styles are however also accepted—the most important criterion is consistency!
- ➢ Use 1.5 line-spacing and justify your text.
- ➢ A paper should have a title page with a student's name, paper's title, word count, instructor's name, date, and the following pledge: *I declare that this assignment is my own original work and that I have correctly acknowledged the work of others. I have not committed an act of plagiarism or any other act of academic dishonesty.*
- ➢ Document format: Word or PDF
- ➢ Submit at the **end of the semester (30 September 2021)** the latest
- ➢ Send to: h.schopmans@fu-berlin.de

Papers will be graded according to the following criteria. **Note:** In Week 8, we will have a workshop on writing a research paper, in which we will discuss these criteria in-depth and take a look at some best practice examples.

Criterion	Expectation
Set-up and overall structure	In the introduction, the paper poses a clear research question and presents an interesting argument. Overall, the paper follows the aforementioned structure (*introduction-theory-case study-conclusion*) and smoothly connects the different parts of the paper. There is a common thread running throughout the paper that the reader can easily follow.
Theoretical discussion	The paper puts the theory into the larger context of IR theories and relates it to other approaches in the field. It justifies why the particular theory is chosen to explore the research question outlined in the introduction. Finally, the paper discusses the main assumption and propositions of the chosen theory and generates ideas about observable implications.
Application of theory to case study	Rather than simply describing the chosen case, the paper analyzes the case by using concepts and mechanisms from the theoretical framework. More specifically, it puts case-specific evidence in relation to theoretical concepts. Finally, the paper critically discusses how the theory helps answering the research question, pointing out both areas that the theoretical approach explains well and possible shortcomings.
Engagement with literature, sources, and evidence	While it may build on the core readings in this syllabus, the paper engages with additional literature to provide depth to the theoretical discussion. In the case study, the paper draws on a variety of reliable sources to support its main argument with convincing evidence.
Language and readability	The paper uses clear, analytical, and jargon-free language. It avoids unnecessary repetitions and complicated sentence structures. The paper makes good use of signposts to illustrate how this section fits in the overall argument and thus ensure the reader does not get lost.
Format and citation	The paper uses a readable, standard font (e.g. Calibri, font size 11). Headings and sub-headings use a different font size and can be clearly discerned from the rest of the text. The text is justified and uses a line-spacing of 1.5. It includes a title page with all relevant information and contains page numbers at the top or bottom of each page. The paper employs a consistent citation style and lists all sources in a bibliography or reference list at the end of the paper.

Grading scale				
Excellent (1.0-1.3)	Very Good (1.7-2.3)	Satisfactory (2.7-3.3)	Sufficient (3.7-4.0)	Not sufficient (5.0)

V. Session Overview

Session	Topic	Remarks
Week 1—13 April 2021	Introduction and Overview	
Part I. Artificial Intelligence: Foundations		
Week 2—20 April 2021	What is thing called Artificial Intelligence?	Students prepare short weekly tasks to stimulate discussion. Classes consist of group discussions and short lecture components.
Week 3—27 April 2021	A Short History of Artificial Intelligence	
Week 4—4 May 2021	Magic Bullet or Existential Risk? The Promises and Perils of Artificial Intelligence	
Part II. Theories of International Relations		
Week 5—11 May 2021	International Politics as Competition: (Neo-)Realism	Classes are divided into two parts. In the first part, one group presents the case study and moderates a group discussion.

In the second part, there will be a short input from the lecturer and a subsequent discussion. |
Week 6—18 May 2021	Cooperation under Anarchy? Neoliberal Institutionalism	
Week 7—25 May 2021	Opening the Black Box: The Domestic Sphere and International Politics	
Week 8—1 June 2021	**WORKSHOP: Writing an Academic Paper**	
Week 9—8 June 2021	The Power of Ideas: Social Constructivism and Normative Change in Global Politics	
Week 10—15 June 2021	Uncovering the Gendered Nature of International Politics: Feminist IR	
Week 11—22 June 2021	How "Global" are Global Politics? Post-Colonial IR	
Week 12—29 June 2021	Class, Hegemony, and Domination: Marxist IR & Neo-Gramscianism	
Part III. Discussion of Paper Ideas		
Week 13—6 July 2021	Discussion Session 1	Students should prepare short outlines of their intended research topics and send them around prior to class.
Week 14—13 July 2021	Discussion Session 2	

WEEK 1. Introduction and Overview

The first session will serve as an introduction to the course. We will use the session to get to know each other and discuss the course's objectives, structure, and requirements. Prior to this first meeting, students should reflect about their expectations towards the course, so we can explore together how each of you can get out most of it. I would be particularly interested in your responses to the following questions:

➢ What do you know about AI already—and what would you like to know?
➢ Are there specific facets of AI (for instance a specific application domain like facial recognition technology) that you are particularly interested in?
➢ Have you taken courses/do you have any prior knowledge on International Relations theories?

I would like to encourage you to read the syllabus carefully prior to this meeting, as this will give you a good idea which topics the course will cover. If you feel like there is something you would learn about, but it's not in the syllabus yet, this is your chance flag it—we can then see if and how it can be incorporated into the overall structure of the course.

Finally, with a look towards the group presentations, make sure to identify in advance a couple of the case studies you would be interested in. We will use this first meeting to divide you into your presentation groups, so that you will have enough time to prepare your session.

Part I
Artificial Intelligence: Foundations

WEEK 2. What is this Thing called "Artificial Intelligence"?

"As soon as it works, no one calls it AI anymore." – *John McCarthy*

"Innumerable tests are available for measuring intelligence, yet no one is quite certain of what intelligence is, or even just what it is that the available tests are measuring." - *R. L. Gregory*

"Narrow AI is just math. It's computational statistics on steroids." – *Meredith Broussard*
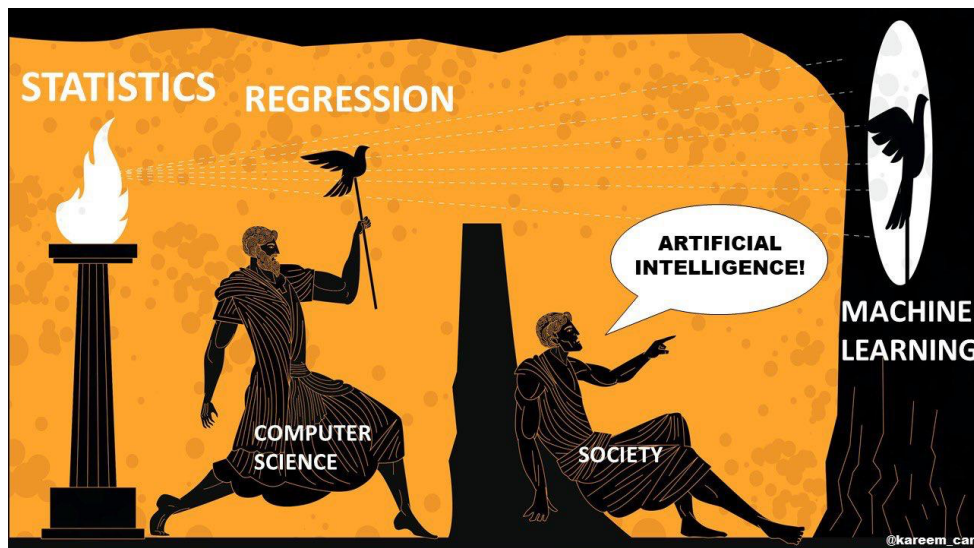
Task: For this week, your job will be to come up with your own short definition of *artificial intelligence*. If someone asked you what AI is, how would you respond? Be prepared to present your definition in class and justify why you think it is a good definition.

Core reading

Allen, Greg. 2020. *Understanding AI Technology*. Joint Artificial Intelligence Center, pp. 3-19.

Broussard, Meredith. 2018. *Artificial Unintelligence*. Cambridge, MA: The MIT Press, pp. 31-39.

Guiding questions

- ➢ What is artificial intelligence? Is AI "one thing"?
- ➢ Why do you think it may be difficult to define AI?
- ➢ Is machine learning the same as AI? Why or why not?
- ➢ How do artificial neural networks work? What is the function of algorithms more generally?
- ➢ What potential problems do you think could result from the way machine learning works?

WEEK 3: A Short History of Intelligent Machines

Although AI is experiencing an unprecedented hype these days, it did not simply appear overnight. In fact, the field of AI is already several decades old! Its birth is commonly attributed to the *Dartmouth Summer Research Project on Artificial Intelligence*, a meeting of U.S. researchers that took place in 1956. At the meeting, the researchers expressed their ambition to create "genuinely intelligent machines"—a goal that has driven research on AI ever since. The decades that followed, however, were characterized by controversies about how to exactly achieve this goal. To this day, the concept of AI remains very much contested.

Task: For this week, think about how you would summarize the history of AI to this day. If you had to draw a timeline, what would you consider major "epochs", what do you think are major turning points?

Core reading

Cardon, Dominique, Jean-Philipp Cointent & Antoine Mazières. 2011. "Neurons spike back: The invention of inductive machines and the artificial intelligence controversy." *Réseaux 5*(211): 173-220.

Guiding questions
  ➢ What approaches toward AI have historically existed? What have they agreed and disagreed about?
  ➢ What are the so-called "AI winters"? When and why did they come about?
  ➢ What role did the U.S. government play in the research on AI?
  ➢ Which factors have driven the recent revival of, and hype around, AI?
  ➢ Have we reached an end point in the development of AI?
  ➢ Do you think there is a geographical bias in the development of AI?

Further reading and resources

Lewis-Kraus, Gideon. 2020. "The Great A.I. Awakening." New York Times Magazine. Available at https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html

Metz, Cade. 2021. *Genius Makers: The Mavericks Who Brought AI to Google, Facebook, and the World*. New York: Dutton.

McCorduck, Pamela. 2004. *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*. Natick, MA: A.K. Peters.

Nilsson, Nils J. 2010. *The Quest for Artificial Intelligence: A history of ideas and achievements.* New York: Cambridge University Press.

Olazaran, Mikel. 1996. "A sociological study of the official history of the perceptron controversy." *Social Studies of Science 26*(3): 611-659.

Wooldridge, Michael. 2020. *The Road to Conscious Machines: The Story of AI*. Pelican Books.

WEEK 4. Magic Bullet or Existential Risk? The Promises and Perils of Artificial Intelligence

In the same manner that scientists, experts, and politicians disagree about what AI is, they are divided on the question what its consequences are. Some believe that AI is as significant as electricity or the steam engine were in the past: It will transform our lives, creating better health, more efficient transportation, and taking over mundane tasks so we have more freedom to do what we like. Others are seeing our future with AI more critical. They believe that AI will destroy jobs and be misused for cyberattacks and mass surveillance. Some experts even consider the creation of a future superintelligence an existential risk to humanity. Not everyone concerned with AI is looking at the future, however. AI, they say, is already here, and so are its many possible dangers.

Task: For this week, your task will be to critically engage with the possible benefits and risks that recent advancements in AI are believed to bring. To this end, do some research on your own and identify one existing or future application or development in AI that you think may be useful, and one that you think may be harmful. Be prepared to justify why you chose these cases.

Core reading

Russel, Stuart. 2019. "Overly intelligent AI." In: *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking.

Crawford, Kate, & Calo, Ryan. 2016. "There is a blind spot in AI research." *Nature.* Available at
https://www.nature.com/news/there-is-a-blind-spot-in-ai-research-1.20805

Guiding questions
➢ What are the major risks associated with artificial intelligence?
➢ What is superintelligence and why are some researchers concerned about it? Why can it be considered an existential risk?
➢ Do you think concerns about superintelligence are warranted, or are they merely "science fiction"?
➢ What do you think we can and should do right now to mitigate the risks of AI?

Further reading and resources

Bostrom, Nick. 2014. *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.

Eubanks, Virginia. 2018. *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: Picador.

Future of Life Institute. 2021. "Benefits and risks of artificial intelligence." *Future of Life Institute*. Available at https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/

O'Neil, Cathy. 2016*. Weapons of mass destruction: How big data increases inequality and threatens democracy*. Penguin Random House.

Roff, Heather. 2019. "Artificial Intelligence: Power to the People." *Ethics & International Affairs 33*(2): 127 – 140.

Tegmark, Max. 2017. *Life 3.0: Being human in the age of artificial intelligence*. Penguin Books.

Wolchover, Natalie. 2015. "Concerns of an artificial intelligence pioneer." *QuantaMagazine*. Available at https://www.quantamagazine.org/artificial-intelligence-aligned-with-human-values-qa-with-stuart-russell-20150421/

Wooldridge, Michael. 2020. "How Things Might Actually Go Wrong." In: *The Road to Conscious Machines: The Story of AI*. Pelican Books.

Zwetsloot, Remco and Dafoe, Allan. 2019. "Thinking about risks from AI: Accidents, Misuse, and Structure." *Lawfare Blog*. Available at: https://www.lawfareblog.com/thinking-about-risks-ai-accidents-misuse-and-structure

Part II
Theories of International Relations

WEEK 5. International Politics as Competition: (Neo-)Realism

Case study: The United States, China, and the "AI Arms Race"
Since artificial intelligence is widely believed to bring transformative benefits, governments across the world have begun investing heavily in AI research and development (R&D). According to many observers, the fact that AI offers "first-mover advantages" has led to a competitive dynamic between powerful states: Each of them wants to develop the most advanced technologies and attract the most talented researchers. Some even speak of an international "AI arms race", headed by the United States and China, and draw parallels to the superpower competition taking place during the Cold War. This gloomy view of international politics raises a number of questions: First, why should states care about each other's AI activities? Can't they just pursue technological development in isolation? Second, if we accept the assumption that they cannot, what are the consequences of international competition? Aren't there any ways to overcome it?

Required Readings

Wolforth, Wiliam C. 2008. "Realism." In Christian Reus-Smit & Duncan Snidal (eds.) *The Oxford Handbook of International Relations*. Oxford: Oxford University Press.

U.S. National Security Commission on Artificial Intelligence. 2021. Final Report: Introduction (pp. 19-28). https://www.nscai.gov/wp-content/uploads/2021/03/Full-Report-Digital-1.pdf

Guiding Questions
➢ What assumptions do (neo-)realists have about international politics?
➢ How do classical realists and neorealists differ from each other?
➢ Do you think realist theories are correct? What contrary evidence can you think of?
➢ What is the role of technology in realist thought? How would realists think about AI? Why should states care about it, and what should they do?
➢ Some critics call realism a "self-fulfilling prophecy." Would you agree?

Further reading and resources

On Realism and Technology in Global Politics:

Conversations in International Relations. 2006. "Interview with John J. Mearsheimer (Part I)." *International Relations 20*(1): 105-123.

Fritsch, Stefan. 2011. Technology and Global Affairs. *International Studies Perspectives 12*(1): 27-45.

Koubi, Vally. 1999. "Military technology races." *International Organization 53*(3): 537-565.

Dafoe, Allan. 2015. "Technological determinism: A typology, scope conditions, and a mechanism." *Science, Technology, and Human Values 40*(6): 1047-1076.

On the "AI Arms Race":

Armstrong, Stuart, Bostrom, Nick, & Shulman, Carl. 2013. "Racing to the precipice: A model of artificial intelligence development." *Technical Report. Future of Humanity Institute*.

Asaro, Peter. 2019. "What is an 'artificial intelligence arms race' anyways? *I/S: A Journal of Law and Policy for the Information Society 15*(1): 45-64.

Jensen, Benjamin M., Whyte, Christopher, and Cuomo, Scott. 2019. "Algorithms at war: The promise, perils, and limits of artificial intelligence." *International Studies Review 22*(3): 526-550.

Lee, Kai-Fu. 2018. *AI superpowers: China, Silicon Valley, and the new world order*. New York: Houghton Mifflin Harcourt.

Scharre, Paul. 2019. "Killer Apps: The real dangers of an AI arms race." *Foreign Affairs* 98(3): 135-144.

Zwetsloot, Remco, Toner, Helen, and Ding, Jeffrey. 2018. "Beyond the AI arms race: America, China, and the dangers of zero sum thinking." *Foreign Affairs*. Available at: https://www.foreignaffairs.com/reviews/review-essay/2018-11-16/beyond-ai-arms-race

WEEK 6. Cooperation under Anarchy? Neoliberal Institutionalism

Case study: The Global Partnership on Artificial Intelligence (GPAI)

If we believe (neo-)realists, international politics is characterized by war and competition. However, a look at the contemporary international landscape reveals that states are not always in conflict with each other. Instead, we see that states cooperate through international institutions, such as the OECD or United Nations, to address collective challenges such as AI. Not only that: Several states are even coming together to form new institutions to specifically tackle AI—the most prominent example being the Global Partnership on Artificial Intelligence (GPAI), which was officially launched in June 2020 by 15 governments. As its objective, the GPAI states among other things to "facilitate international collaboration" and "ultimately promote trust in and the adoption of trustworthy AI." How, then, is such cooperation possible in an anarchical international system? And what types of institutions should states set up to address the challenges posed by AI?

Core reading

Dai, Xinyuan, Snidal, Duncan, & Sampson, Michael. 2010. "International cooperation theory and international institutions." *Oxford Research Encyclopedia of International Studies.*

GPAI [Video]. 2020. "Address by President of France Emmanuel Macron." *The Global Partnership on Artificial Intelligence*. Available at https://gpai.ai/events/montreal-2020/

Cihon, Peter, Maas, Matthijs, & Kemp, Luke. 2020. "Fragmentation and the future: Investigating architectures for international AI governance*." Global Policy 11*(5): 545-556.

Guiding Questions
  ➢ What assumption do neoliberal institutionalists have about state behavior at the international level?
  ➢ How do they differ from (neo-)realists? What do they have in common?
  ➢ Do states always cooperate with each other, or only under certain conditions?
  ➢ In your opinion, what purpose does an institution like the GPAI fulfill?
  ➢ Can you imagine how neorealists would respond to institutionalists? How would they interpret an institution like the GPAI?

Further reading and resources

On neoliberal institutionalism, regimes, and the "neo-neo-debate":

Alter, Karen, and Meunier, Sophie. 2009. "The politics of international regime complexity." *Perspectives on Politics 7*(1): 13-24.

Axelrod, Robert, and Keohane, Robert. 1985."Achieving cooperation under anarchy: Strategies and institutions." *World Politics 38*(1): 226-254.

Keohane, Robert O., and Martin, Lisa L. 1995. "The Promise of Institutionalist Theory."
         International Security 20(1): 39-51.

Mearsheimer, John J. 1994. The False Promise of International Institutions. *International
         Security 19*(3): 5-49.

On global AI governance:

Butcher, James & Beridze, Irakli. 2019. "What is the State of Artificial Intelligence Governance
         Globally?" *The RUSI Journal 164*(5-6): 88-96.

Dafoe, Allan, & Journal of International Affairs. 2019. "Global politics and the governance of
         artificial intelligence: An interview with Allan Dafoe." *Journal of International Affairs
         72*(1): 121-126.

WEEK 7. Opening the Black Box: The Domestic Sphere and International Politics

Case Study: China and the Rise of Digital Authoritarianism

So far, the theoretical schools discussed share one assumption: that states are unitary actors who have similar preferences, and should thus behave similarly on the international stage. But can this really be true? States across the world are quite different—for instance, some have political systems that are much more democratic than others. Should this not be somehow reflected in how they conduct politics? Consider China, an authoritarian state that uses AI technologies in the domestic sphere to surveil and repress their people. There are concerns that China may export these technologies to other authoritarian states. Other authoritarian states use AI technologies to spread misinformation to other countries, and undermine their political systems. Against this background, can we really say that all states always behave the same?

Core reading

Moravcsik, Andrew. 2008. "The New Liberalism." In Christian Reus-Smit and Duncan Snidal (eds.), *The Oxford Handbook of International Relations*. Oxford: Oxford University Press, pp.

Andersen, Ross. 2020. "The panopticon is already here." *The Atlantic*. Available at https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/

Guiding questions

  ➢ How do liberal theories challenge mainstream accounts of International Relations?
  ➢ What determines the preferences of states?
  ➢ Are democracies more peaceful than authoritarian states?
  ➢ Would you say democracies deal with AI differently than authoritarian states?

Further reading and resources

Dragu, Tiberiu & Lupu, Yonatan. 2021. "Digital Authoritarianism and the Future of Human Rights." *International Organization*: 1-27.

Feldstein, Steven. 2019. The Road to Unfreedom: How Artificial Intelligence is Reshaping Repression. *Journal of Democracy 30*(1): 40-52.

Freedom House. 2018. "The Rise of Digital Authoritarianism: Fake news, data collection, and the challenge to democracy." *Freedom House*. Available at https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism

Milner, Helen. 1997. *Interests, Institutions, and Information: Domestic Politics and International Relations*. Princeton: Princeton University Press.

Polyakova, Alina, & Meserole, Chris. 2019. "Exporting digital authoritarianism: The Russian and Chinese models." Brookings. Available at https://www.brookings.edu/research/exporting-digital-authoritarianism/

WEEK 8. WORKSHOP: Writing an Academic Paper

This week, we will take a break from the world of states, power, and intelligent machines. Instead, we will have a small workshop that is meant to help you with the preparation of your own research paper, which constitutes the main requirement for this course.

The goal of the workshop is to provide you not only with an idea of this course's specific requirements, but rather to reflect on what makes a good research paper in general. Among other things, we will look at some best practice examples to shed light on the following aspects:

- ➢ Research question
  - o What is a good research question? How do I decide on a research question?
  - o Should I always go for a "Why?" question? Or can I also ask "how" or "what"?
- ➢ Literature review and theoretical framework
  - o What is the purpose of a literature review? How detailed do I engage with the existing literature?
  - o How do I justify the theoretical approach I chose? Do I need to describe the entire theory?
- ➢ Empirics and case studies
  - o How do I select a case study for my paper?
  - o What are good sources for my case study? How do I know that a source is reliable?
  - o How do I make a connection between abstract theoretical propositions and concrete empirical observations?
  - o At what point do I know that I have collected "enough" evidence?
- ➢ Structure
  - o What structure should my paper follow?
  - o Which elements belong into a good introduction? Should I already give away my main argument in the beginning?
  - o Is the conclusion of my paper merely a summary?
- ➢ Format and style
  - o What should I pay attention to with regard to language and the format of my paper?
  - o How do I properly cite a source? What should my reference list look like?

WEEK 9. The Power of Ideas: Social Constructivism and Normative Change in Global Politics

Case study: Towards a Global Ban on Lethal Autonomous Weapon Systems

Since 2014, states have convened at the United Nations to discuss what to do about the use of AI technologies in weapon systems. A particular bone of contention has been the prospect of lethal autonomous weapon systems (LAWS), which could one day be used to select and kill targets without human intervention in these decisions. A number of powerful states seem keen on developing LAWS. They argue that warfare could become less deadly if fewer humans are involved. At the same time, global resistance against LAWS is building. A growing coalition of actors is demanding an international ban on "killer robots", claiming that removing human control over weapon systems would violate international humanitarian law and human rights.

Core reading

Finnemore, Martha and Kathryn Sikkink. 1998. "International norm dynamics and political change." *International Organization 52*(4): 887-917.

Campaign to Stop Killer Robots. 2020. "Campaigner's kit." Available at
https://www.stopkillerrobots.org/wp-content/uploads/2020/02/2020_Campaigners-Kit_FINAL.pdf

Guiding questions
- ➢ Do ideas have a role to play in international politics?
- ➢ What are norms and how do they affect the behavior of states?
- ➢ Where do norms come from? Under which conditions do they become internationally recognized?
- ➢ Do you think it is possible to pre-emptively ban killer robots? Why or why not?
- ➢ Should killer robots be banned? What arguments could speak for or against such a ban?

Further reading and resources

On norms and policy entrepreneurs:

Carpenter, Charli. 2011. "Vetting the advocacy agenda: Network centrality and the paradox of weapons norms." *International Organization 65*(1): 69-102.

Cross, Mai'a K.D. 2013. "Rethinking epistemic communities twenty years later." *Review of International Studies 39*: 137-160.

Finnemore, Martha, & Hollis, Duncan. 2016. "Constructing norms for global cybersecurity." *American Journal of International Law 110*(3): 425-479.

On LAWS:

Bahcecik, Serif Onur. 2019. "Civil society responds to the AWS: Growing activist networks and shifting frames." *Global Policy 10*(3): 365-369.

Bode, Ingvild, and Huelss, Hendrik. 2018. "Autonomous weapons systems and changing norms in international relations*." Review of International Studies 44*(3): 393-413.

Garcia, Denise. 2016. "Future arms, technologies, and international law: Preventive security governance." *European Journal of International Security 1*(1): 94-111.

Future of Life Institute [Podcast]. 2019. "Why ban lethal autonomous weapons?" Future of Life Institute. Available at https://futureoflife.org/2019/04/02/fli-podcast-why-ban-lethal-autonomous-weapons/

Haner, Justin, and Garcia, Denise. 2019. "The artificial intelligence arms race: Trends and world leaders in autonomous weapons development." *Global Policy 10*(3): 331-337.

Human Rights Watch. 2020. "Stopping Killer Robots: Country Positions on Banning Fully Autonomous Weapons and Retaining Human Control." Human Rights Watch. Available at: https://www.hrw.org/report/2020/08/10/stopping-killer-robots/country-positions-banning-fully-autonomous-weapons-and

Rosert, Elvira and Sauer, Frank. 2019. "Prohibiting autonomous weapons: Put human dignity first." *Global Policy 10*(3): 370-375.

WEEK 10. Uncovering the Gendered Nature of International Politics: Feminist IR

Case Study: AI as a Tool of the Global Patriarchy

In political discussions on AI, one question that often receives less attention is *who* is making decisions—on how AI systems are designed, how they are used, and how they should be regulated. At closer inspection, we quickly see that both the research field of AI and that of international policymaking are characterized by the same problem: a lack of diversity. In short, the people who design AI systems, and who decide over their uses, are predominantly male and white. For the global politics of AI, this raises a number of questions: First, how does the lack of diversity affect the nature of AI systems? Does the dominance of patriarchal structures have an effect on how AI is governed internationally? And would international politics unfold differently if there was more diversity in terms of gender and race?

Core reading

Youngs, Gillian. 2004. "Feminist International Relations: A contradiction in terms? Or: Why women and gender are essential to understanding the world 'we' live in." *International Affairs 80*(1): 75-87.

West, Sarah, Whittaker, Meredith, and Crawford, Kate. 2019. "Discriminating Systems: Gender, Race, and Power in AI." *AI Now Institute*, pp. 5-32.

Stolton, Samuel. 2020. "Poland rejects Presidency Conclusions on artificial intelligence, rights." *Euractiv*, 26 October 2020. Available at https://www.euractiv.com/section/digital/news/poland-rejects-presidency-conclusions-on-artificial-intelligence-rights/

Guiding questions

➢ What is the feminist critique of mainstream IR approaches?
➢ Which issues and problems in international politics do feminist scholars shift our attention to?
➢ What does a gender perspective reveal about AI technology? Why are AI systems "gendered" and which problems does this raise? Is biased data the only problem?
➢ Why are discriminatory AI systems a problem for *international* politics?

Further reading and resources

Coded Bias[Movie]. 2020. Available on Netflix.

Buolamwini, Joy, & Gebru, Timnit. 2018. "Gender shades: Intersectional accuracy disparities in commercial gender classification." *Proceedings of Machine Learning Research* 81: 1-15.

Cupac, Jelena, & Ebetürk, Irem. 2020. "The personal is global political: The antifeminist backlash in the United Nations." *The British Journal of Politics and International Relations 22*(4): 702-714.

Crawford, Kate. 2016. "Artificial intelligence's white guy problem." *New York Times.* Available at https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html

Leavy, Susan. 2018. "Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning." *Proceedings of the 1st International Workshop on Gender Quality in Software Engineering*: 14-16.

Perez, Caroline Criado. 2019. *Invisible women: Exposing data bias in a world designed for men*. London: Chatto & Windus.

Prügl, Elisabeth. 2012. "Feminist International Relations: The state of the field." In: Bayes, Jane (ed.) *Gender and politics: The state of the discipline*. Barbara Budrich.

Tickner, J. Ann. 1997. "You just don't understand: Troubled engagements between feminists and IR theorists." *International Studies Quarterly 41*(4): 611-632.

Waylen, Georgina. 2006. "You Still Don't Understand: Why Troubled Engagements Continue between Feminists and (Critical) IPE." *Review of International Studies 32*(1):145–64.

Weber, Jutta. 2006. "From Science and Technology to Feminist Technoscience." In: Lorber, Judith, Davis, Kathy, and Evans, Mary (eds.) *Handbook of Gender and Women's Studies*. Sage.

WEEK 11. How "Global" are Global Politics? Post-Colonial IR

Case study: *A "Global" Movement for AI Ethics*
In response to the many challenges posed by AI, recent years have witnessed a new global movement towards establishing a universal AI *ethics*. In other words, researchers, civil society organizations, and even technology corporations have come together develop principles such as transparency, privacy, accountability, and beneficence that should be observed by all those developing and deploying AI systems. Abiding by these principles, they argue, will ensure that AI will be used for good of all of humanity. International organizations such as UNESCO or the OECD are working towards formalizing these principles and ensuring that the benefits of AI will be distributed equally among everyone in the world. However, how universal are these principles really? And is the development of AI technology really a "global" endeavor that everyone is included in?

Core reading

Seth, Sanjay. 2011. "Postcolonial theory and the critique of International Relations." *Millennium: Journal of International Studies 40*(1): 167-183.

Jobin, Anna, Ienca, Marcello, Vayena, Effy. 2019. "The global landscape of AI ethics guidelines." *Nature Machine Intelligence* 1(9): 389-399.

Hao, Karen. 2020. "The problems AI has today go back centuries." *MIT Technology Review*. Available at https://www.technologyreview.com/2020/07/31/1005824/decolonial-ai-for-everyone/

Guiding questions

  ➢ What do post-colonial theories suggest about the nature of international politics, and the discipline of IR more broadly?
  ➢ How do post-colonial theories differ from mainstream IR with regard to their conception of power?
  ➢ What are the potential consequences of a "Eurocentric" AI, for instance with regard to the AI arms race, AI governance, or the global movement for AI ethics?

Further reading and resources

Capan, Zeynep. 2017. "Decolonizing International Relations?" *Third World Quarterly 38*(1): 1-15.

Fjeld, Jessica, Achten, Nele, Hilligoss, Hannah, Nagy, Adam, & Srikumar, Madhulika. 2020. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI." *Berkman Klein Center Research Publication*.

Hao, Karen. 2020. ""I started crying": Inside Timnit Gebru's last days at Google—and what happens next." *MIT Technology Review*. Available at https://www.technologyreview.com/2020/12/16/1014634/google-ai-ethics-lead-timnit-gebru-tells-story/

IEEE. 2019. *Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems*. Available at: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf?utm_medium=undefined&utm_source=undefined&utm_campaign=undefined&utm_content=undefined&utm_term=undefined

Mhlambi, Sabelo. 2020. "From rationality to relationality: Ubuntu as an ethical and human rights framework for artificial intelligence governance." Discussion Paper, *Harvard Kennedy School Carr Center for Human Rights Policy*.

Mohamed, Shakir, Png, Marie-Therese, & Isaac, William. 2020. "Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence." *Philosophy & Technology 33*: 659-684.

WEEK 12. Class, Hegemony, and Domination: Marxist IR & Neo-Gramscianism

Case study: The Power of Big Tech and the Rise of "Surveillance Capitalism"
When examining the global politics of AI, it is impossible to look past the role played by technology corporations such as Google, Facebook, Amazon, and Baidu. In the past, research on frontier technologies had often taken place in state-run research labs; the data-intensive nature of today's AI, however, means that much of the research and application of AI is taking place at the headquarters of "Big Tech". It seems only logical, then, that these actors should be able to wield considerable power. In fact, the scholar Shoshanna Zuboff claims we have reached a new stage of capitalism—surveillance capitalism—in which corporations use our personal experiences as commodities. At the same time, corporations try to downplay their power, pointing to ethical codes of conduct as proof of their responsible behavior. Understanding this new form of power seems to pose a problem to International Relations theories that have largely looked at global politics through a state-centric lens. Can International Relations tell us anything about the role that capital and technology corporations play in international politics?

Core reading

Robinson, William. 2005. "Gramsci and Globalization: From nation-state to transnational hegemony." *Critical Review of International Social and Political Philosophy 8*(4): 559-574.

Zuboff, Shoshanna. 2019. *The age of surveillance capitalism: The fight for a human future at the new frontier of power.* London: Profile Books Ltd., Chapter 18.

Guiding questions

- In the Gramscian view, who are the main groups or actors dominating world politics?
- How does the Neo-Gramscian conception of hegemony differ from a Realist conception?
- Do you think we are entering a new "historical block"? Who are the actors that are in power and how did they get there?
- Can international political developments such as the "AI arms race" be explained through the influence of transnational classes?
- Can you think of actors who could be driving possible "counter-hegemonies"? Under which conditions could they succeed?

Further reading and resources

Bietti, Elettra. 2020. "From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy." *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*: 210-19.

Cox, Robert. 1981 . "Social forces, states, and world orders: Beyond International Relations theory." *Millennium: Journal of International Studies 10*(2): 126-155.

Frey, Carl Benedikt (2019). *The Technology Trap: Capital, Labor, and Power in the Age of Automation*. Princeton & Oxford: Princeton University press, Chapter 12.

Greene, Daniel, Hoffmann, Anna Lauren, and Stark, Luke. 2019. "Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning." *Hawaii International Conference on System Sciences*.

Hagendorff, Thilo. 2020. "The ethics of AI ethics: An evaluation of guidelines." *Minds and Machines 30*: 99-120.

Worth, Owen. 2011. "Recasting Gramsci in international politics." *Review of International Studies 37*(1): 373-392.

WEEKS 13 & 14. Discussion of Ideas for Research Papers

The final two sessions of this class offer you the opportunity to present your ideas for a research paper to your peers and receive feedback. What these sessions will look like in detail will depend on the class size. For now, the plan is the following:

1. If you are taking this class for credit, prepare a ½-page outline for the paper you would like to write. The outline should contain your research question, the theory you would like to work with, and some preliminary ideas about the case study you want to look at. Please send this outline to h.schopmans@fu-berlin.de **one week prior** to the session you are assigned to.

2. During class, you have the opportunity to **briefly** present your main idea and articulate open questions you may have. There will be time allocated for a short discussion, in which your peers comment on your outline and make constructive suggestions.

3. When commenting on your peers' proposals, make sure to keep in mind the evaluative criteria for this course's research papers (see page 5). Most importantly, make sure to not only criticize your fellow students' ideas, but to offer some constructive advice on how they could address the issue or problem you have identified.