**Course Syllabus, Version 22.06.2020**


**GRAD-E1319: Governance and Politics of Artificial Intelligence**

*Concentration: Policy Analysis*

Joanna Bryson/Slava Jankin


1. **General information**

| Class time | Monday 10-12 |
|---|---|
| Course Format | This course is based on readings and in-class discussion. Each week there is a lecture component of no more than an hour during the two-hour discussion window. These lectures and discussions will be delivered live and simultaneously available online, as well as recorded. Under social distancing, students who desire to attend the live lectures will be allocated a chance to attend randomly each week based on the number of in-person seats available. Students are also responsible for preparing questions in advance for two lectures; each lecture will be followed by five students with prepared questions leading discussion. This too can be done either on line or in person. Students are encouraged to bring laptops to lectures they attend in person so that they can also participate in online discussion with absent students via chat.<br><br>This course is also offered in the spring semester. |
| Instructor | Joanna Bryson<br>Slava Jankin |
| Instructor's office | 3.21.1<br>3.15 |
| Instructor's e-mail | bryson@hertie-school.org<br>jankin@hertie-school.org |
| Instructor's phone number | 218<br>167 |
| Assistant | Name: Alex Karras |

| | Email: karras@hertie-school.org |
| | Phone: +49 30 259 219 156 |
| | Room: 3.45 |
| Instructor's Office Hours | TBA |

Link to Module Handbook **MIA** and **MPP**
Link to **Study, Examination and Admission Rules**

Instructor Information:
Joanna  Bryson is Professor of Ethics and Technology at the Hertie School. She is an academic recognised for broad expertise on intelligence, its nature and its consequences. She advises governments, transnational agencies, and NGOs globally, particularly in AI policy. She holds two degrees each in psychology and AI (BA Chicago, MSc & MPhil Edinburgh, PhD MIT). From 2002-19 she was Computer Science faculty at the University of Bath; she has also been affiliated with Harvard Psychology, Oxford Anthropology, Mannheim Social Science Research, The Konrad Lorenz Institute for Evolution and Cognition Research,  and the Princeton Center for Information Technology Policy. During her PhD she observed confusion generated by anthropomorphised AI, leading to her first AI ethics publication "Just Another Artifact" in 1998. In 2010 she coauthored the first national-level AI ethics policy, the UK's Principles of Robotics. She presently researches the impact of technology on human cooperation, and AI/ICT governance.

Slava Jankin is Professor of Data Science and Public Policy at the Hertie School. He is the Director of the Hertie School Data Science Lab. His research and teaching is primarily in the field of natural language processing and machine learning. Before joining the Hertie School faculty, he was a Professor of Public Policy and Data Science at University of Essex, holding a joint appointment in the Institute for Analytics and Data Science and Department of Government. At Essex, Slava served as a Chief Scientific Adviser to Essex County Council, focusing on artificial intelligence and data science in public services. He previously worked at University College London and London School of Economics. Slava holds a PhD in Political Science from Trinity College Dublin.

2. **Course Contents and Learning Objectives**

Course contents:

Innovations in Artificial intelligence (AI) are transforming economies and societies globally, leading to governance quandaries and political debates. AI is unusual for a transformative technology, in that many, perhaps all leaders and ordinary citizens alike come to it with preconceptions derived from fiction and popular understanding, which makes it unusually easy for policy to be deflected from a scientific basis. This course explores the transformations and corresponding policy challenges of AI,

with a focus on active areas of political debate and policy research. Integrating perspectives from both natural and social sciences, this course will provide learning experiences that examine the impact of AI on individual humans and their societies. We will explore the proliferation of algorithmic decision-making and autonomous systems; the issues of ethics, fairness, transparency and accountability raised by AI techniques such as machine learning; balances and interactions between regulation and innovation; the effects of AI on human rights and economic well being; increasing oppressive capabilities of state- and non-state actors; the global AI arms race, and the impacts of AI in the global arms race. We consider both public and private strategies of regulation, and local, national, and transnational aspects of governance.

Main learning objectives:

In this course, students will be introduced to a wide range of current debates concerning the role of artificial intelligence in society, and learn to write policy documents covering and evaluating the diverse perspectives on this technology. We will also explore the international policy, governance, and regulatory landscape around AI.

Target group:

Anyone interested in digital technology and the debates around the future of life and society and the role research and development in a policy school can play in these debates.

Teaching style:

Lectures covering a wide range of materials based on the professors' real-world experiences engaging in politics and governance of AI. Discussions based on questions from the students.

Prerequisites:

None, but those without an ICT background may need to invest more time on some of the readings.

Diversity Statement:

As you may know, the Hertie School is committed to implementing a new Diversity and Inclusion Strategy. We strive to have an inclusive classroom but ask your informal feedback on inclusivity throughout the course.

## 3. Grading and Assignments

Composition of Final Grade:

| Assignment 1: First policy brief | Deadline: TBA | Submit via Moodle | 30% |
|---|---|---|---|

| Assignment 2: Second policy brief | Deadline: TBA | Submit via Moodle | 30% |
|---|---|---|---|
| Assignment 3: Policy presentation | Deadline: TBA | Submit via Moodle | 30% |
| Participation grade | | | 10% |

Assignment Details

The assessment for the course consists of a series of **policy research projects** with policy briefs as deliverables for the first two assignments and a pre-recorded presentation (with an accompanying policy brief as a handout) for the final assignment. All projects must be carried out in teams of 2-4 students (individual submissions will not be accepted), and team composition will be **randomly assigned** for each assignment.

The aim of the assessments is five-fold:
- First, it will provide you with the opportunity to apply the concepts learned in this class creatively, which helps you with understanding material more deeply.
- Second, designing and working on a unique project in a team which is something that you will encounter, if you haven't already, in the workplace, and the projects help you prepare for that. There's largely no single-author work in the public sector, particularly at the intersection of technology and policy where you need a combination of relevant skills.
- Third, along with the opportunity to practice and the satisfaction of working creatively, students can use these projects to enhance their portfolio or resume.
- Fourth, random assignment to groups will allow you to get to know your classmates better and facilitate a more fair distribution of skills and interests.
- Fifth, the peer assessment component of the course will allow you to actively participate and engage with each other's work and thus have a more comprehensive overview of the topics covered in the course.

In the **policy brief**, you will concisely summarise the current state of scientific knowledge on one specific effect of artificial intelligence on society. Based on the current state of the art, you will formulate basic policy recommendations. These recommendations have to be realistic and practical, taking into account everything you learnt at the Hertie School (and beyond) about public policy and policy making. You are welcome to build on a specific example, as long as you draw from the literature more generally. In terms of style (though not necessarily content), the PRIO series is a good example, although you will add proper scientific references:
http://bit.ly/PRIO_policy_briefs

Conveying the results of the analysis to policy makers is often the hardest part of research. The third assignment is designed to simulate the presentation of scientific evidence to policy makers. You will carry out a policy relevant research project focusing on an issue covered in the second half of the course

(Sessions 7-11). You will undertake extensive research for the project, draft the policy brief based on your analysis (this is your handout circulated ahead of the presentation), and **present your results in the form of a pre-recorded presentation** (your deliverable).

Note that there is no "perfect project." While you are encouraged to be ambitious, the most important aspect of this research project is your learning experience. Hence, you don't want to pick something that is too easy for you, but similarly, you don't want to choose a project where you are not certain that is out of the scope of this class. Again, the objective of this project is to provide you with hands-on practice and an opportunity to learn.

### Assignment 1
The first policy research project should cover one or more topics covered in Sessions 1-3. The deliverable is a policy brief with maximum 1,000 words (including references, tables and figures) and maximum three references.

### Assignment 2
The second policy research project should cover one or more topics covered in Sessions 4-6. The deliverable is a policy brief with maximum 1,000 words (including references, tables and figures) and maximum three references.

The first two assignments will be graded by the professors, also providing you with the formative guidance for the third assignment.

### Assignment 3
This assignment consists of a policy research project that covers one or more topics in Sessions 7-11. The deliverable for the assignment is a pre-recorded presentation delivered in the final session of the course. This assignment is peer-reviewed where you will evaluate all projects (apart from your own) on three dimensions:
1. Creativity as a presentation
2. Quality of presentation of scientific evidence to policy makers
3. Feasibility of policy recommendations

You will evaluate each project on a scale 1-10 for each of the three components, where 10 is best, an average across three dimensions across all peer evaluations will be the group grade for Assignment 3. We will provide you with an evaluation form where along with the score for each component you will also add a brief note justifying the score. You will have 12 hours after the end of the class to complete the evaluation form.

### Participation grade
As pre-work before the first lecture of the course, students will be asked (through an online poll disseminated over moodle) to each take two lecture topics for which they will prepare and ask questions. Initial preparation should be based on the readings, but questions may be adjusted in response to the

lecture. For each lecture, up to five students will be assigned to ask questions and lead discussion on it for 5-10 minutes each. Participation will be marked on a simple scale:

- 0: failure to attend
- 50: question asked not on topic, or discussion points by others dropped
- 85: good question asked, fair discussion led (expected mark)
- 100: best question / discussion of that particular lecture (there may rarely be ties)

All students will be encouraged to participate in discussion during the lecture and in response to their peer's questions, but this participation will not be graded.

**Late submission of assignments:** For each day the assignment is turned in late, the grade will be reduced by 10% (e.g. submission two days after the deadline would result in 20% grade deduction).

**Attendance:** Students are expected to be present and prepared for every class session. Active participation during lectures and seminar discussions is essential. If unavoidable circumstances arise which prevent attendance or preparation, the instructor should be advised by email with as much advance notice as possible. Please note that students cannot miss more than two out of 12 course sessions. For further information please consult the Examination Rules §10.

**Academic Integrity:** The Hertie School is committed to the standards of good academic and ethical conduct. Any violation of these standards shall be subject to disciplinary action. Plagiarism, deceitful actions as well as free-riding in group work are not tolerated. Free riding must be reported. See Examination Rules §16.

**Compensation for Disadvantages**: If a student furnishes evidence that he or she is not able to take an examination or complete a coursework as required in whole or in part due to disability or permanent illness, the Examination Committee may upon written request approve learning accommodation(s). In this respect, the submission of adequate certificates may be required. See Examination Rules §14.

**Extenuating circumstances:** An extension can be granted due to extenuating circumstances (i.e., for reasons like illness, personal loss or hardship, or caring duties). In such cases, please contact the course instructors *and* the Examination Office *in advance* of the deadline.

## 4. General Readings

At the moment there is no single textbook we can use for this course. We will draw from a large number of recent articles across multiple disciplines.

## 5. Session Overview

| Session | Session Date | Session Title |
|---|---|---|
| 1 | 07.09.2020 | **What is Artificial Intelligence?** |
| 2 | 14.09.2020 | **Overview of machine learning** |

| 3 | 21.09.2020 | Bias in algorithms |
|---|---|---|
| 4 | 28.09.2020 | Malicious AI and Democracy |
| 5 | 05.10.2020 | Militarisation of AI |
| 6 | 12.10.2020 | Politics of national AI policies and approaches |
| Mid-term Exam Week: 19.10 - 23.10.2020 – no class | | |
| 7 | 26.10.2020 | AI and Grand Challenges |
| 8 | 02.11.2020 | Social Impacts and Future of Work |
| 9 | 09.11.2020 | Economics of AI |
| 10 | 16.11.2020 | Guest lecture: German National AI Policy (provisional) |
| 11 | 23.11.2020 | {Transnational policy} & Regulation of AI |
| 12 | 30.11.2020 | Student presentations |
| Final Exam Week: 14.12 - 18.12.2020 – no class | | |

### 6. Course Communication, Course Sessions, and Readings

All readings will be accessible on the Moodle course site before semester start. In the case that there is a change in readings, students will be notified by moodle. All students are required to keep up with moodle announcements, and encouraged to participate on moodle forums.

Required readings are to be read and analysed thoroughly. Optional readings are intended to broaden your knowledge in the respective area and it is highly recommended to at least skim them.

Some readings and even the ordering of lectures may be subject to change, but all changes will be made at least two weeks in advance so students preparing questions will have adequate warning.

| Session 1: What is Artificial Intelligence? [JB] | |
|---|---|
| Learning Objective | We establish the baseline for what systems are and are not considered to contain AI, and from a technological perspective, how such systems can be governed. Students should understand by the end the costs of computation, combinatorial complexity, and what accountability and transparency can mean in the context of intelligent digital systems. We also establish the difference between governance and politics at least for this course, and introduce the question of who should govern what in the context of AI. |

| Required Readings | <ul><li>Joanna J. Bryson, The Artificial Intelligence of the Ethics of Artificial Intelligence: An Introductory Overview for Law and Regulation. solicited and reviewed for M. Dubber, F. Pasquale, & S. Das (Eds.) 2020.</li><li>One Hundred Year Study on Artificial Intelligence (Stanford, 2016): Read Preface (pp. 1-3), Sections I, II, and III (pp. 12-49); Optional: Appendix I (pp. 50-52), for those who would like a short history of AI</li></ul> |
|---|---|
| Optional Readings | <ul><li>Kroll, J. A., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., _ Yu, H. (2016). Accountable algorithms. University of Pennsylvania Law Review, 165 (3), 633.</li><li>Bryson, J. (2018) Patiency is not a virtue the design of intelligent systems and systems of ethics.</li></ul> |

| Session 2: Overview of machine learning [SJ] | |
|---|---|
| Learning Objective | In this session we provide a very broad overview of machine learning algorithms. This will provide the basis for our subsequent discussions trying to separate the hype from reality. |
| Required Readings | <ul><li>Kakia Chatsiou and Slava Jankin Mikhaylov, "Deep Learning for Political Science" in *The SAGE Handbook of Research Methods in Political Science and International Relations* (eds. Luigi Curini and Robert Franzese), SAGE, 2020. Pre-print version http://arxiv.org/abs/2005.06540</li></ul> |
| Optional Readings | <ul><li>Domingos, P. (2015). The master algorithm: How the quest for the ultimate learning machine will remake our world. Basic Books.</li></ul> |

| Session 3: Bias in algorithms [JB] | |
|---|---|
| Learning Objective | We cover the origins and consequences of bias, stereotypes, and prejudice exhibited by digital systems, how these relate to bias in society more generally, and how the challenges of behaving equitably vary when actions are expressed by AI systems rather than directly by people. |
| Required Readings | <ul><li>Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan, Semantics derived automatically from language corpora contain human biases. Science 356 (6334)183-186</li></ul> |

| | |
|---|---|
| | ○ Be sure to also look at the supplement, which gives the stimuli and shows similar results for a different corpus and word-embedding model.<br>● "How we Analyzed the COMPAS Recidivism Algorithm" by Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin (ProPublica, 2016)URL |
| **Optional Readings** | ● Kusner, Matt J., Joshua Loftus, Chris Russell, and Ricardo Silva. "Counterfactual fairness." In Advances in neural information processing systems, pp. 4066-4076. 2017.<br>● "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights," pp. 1-18, 22-24 (White House, May 2016)<br>● Benjamin, Ruha. "Race after technology: Abolitionist tools for the new jim code." *Social Forces* (2019). (warn library!) |


| **Session 4: Malicious AI and Democracy [SJ]** | |
|---|---|
| **Learning Objective** | In this session we discuss malicious applications of AI that undermine the institutions of democracy. From election interference to fundamental rights to deepfakes, we will cover recent use-cases and technological tools of malicious AI. We will also consider suggested technological and governance solutions like differential privacy. |
| **Required Readings** | ● "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation" Brundage et al. (ArXiv, 2018)<br>● "Malevolent soft power, AI, and the threat to democracy" Elaine Kamarck, 29 November 2018. Brookings.<br>● "OpenAI: Please Open Source Your Language Model." Hugh Zhang, *The Gradient*, 19 February 2019. |
| **Optional Readings** | ● "Privacy and human behavior in the age of information" by Alessandro Acquisti, Laura Brandimarte, and George Loewenstein (Science, 2015)<br>● Burgess, M. (2018) What is GDPR? The need to know guide. Wired, 21 January 2019.<br>● "What is Differential Privacy?" by Matthew Green (A Few Thoughts on Cryptographic Engineering, 2016)<br>● Ross Anderson, *Security Engineering*, 3rd edition, Chapter 11 "Inference Control", available online here: https://www.cl.cam.ac.uk/~rja14/book.html |

| Session 5: Militarisation of AI [SJ] | |
|---|---|
| **Learning Objective** | In this session we discuss military applications of AI. We go beyond the usual "killer robots" and discuss how AI starts to permeate different aspects of both combat and command & control systems -- complicating international governance and regulation efforts. |
| **Required Readings** | <ul><li>"It's already too late to stop the AI arms race - We must manage it instead" by Edward Moore Geist (Bulletin of the Atomic Scientists, 2016)</li><li>The Militarization of Artificial Intelligence, Sisson et. al on behalf the United Nations, June 2020.</li><li>Asaro, P. (2012). "On banning autonomous weapon systems: human rights, automation, and the dehumanization of lethal decision-making." International Review of the Red Cross, 94, 687-709.</li></ul> |
| **Optional Readings** | <ul><li>"The role of AI in future warfare." Michael E. O'Hanlon, 29 November 2018. Brookings.</li><li>"Weapons of the weak: Russia and AI-driven asymmetric warfare." Alina Polyakova, 15 November 2018. Brookings.</li><li>Allen, G. and Chan, T. (2017) "Artificial Intelligence and National Security." Harvard Kennedy School Belfer Center study.</li></ul> |

| Session 6: Politics and AI policies: National AI policies and approaches [JB] | |
|---|---|
| **Learning Objective** | In this session we will discuss issues localised governments have to consider with respect to regulating AI, and will consider some present variation and proposed strategies. Issues examined will include: giving AI legal personality, taxation and revenue considerations for technology and data, data privacy and national infrastructure (including cybersecurity), and the obligations of corporations including their capacity for self-governance. |
| **Required Readings** | <ul><li>Joanna J. Bryson, Mihailis E. Diamantis, and Thomas D. Grant, Of, For, and By the People The Legal Lacuna of Synthetic Persons. Artificial Intelligence and Law 25(3)273–291 (2017).</li><li>Smith, Brad. "Facial recognition technology: The need for public regulation and corporate responsibility." Microsoft on the Issues (2018).</li><li>Something about taxes</li></ul> |

| | |
|---|---|
| | ● Sambuli, Nanjira. "Challenges and opportunities for advancing Internet access in developing countries while upholding net neutrality." Journal of Cyber Policy 1, no. 1 (2016): 61-74. |
| **Optional Readings** | ● Pasquale, Frank. "A rule of persons, not machines: the limits of legal automation." Geo. Wash. L. Rev. 87 (2019): 1.<br>● Solaiman SM (2017) Legal personality of robots, corporations, idols and chimpanzees a quest for legitimacy. Artificial Intelligence and Law 25(2)155–179<br>● FEPS 2020 AI cybersecurity report (not out yet) |

**Mid-term Exam Week: 19 – 23.10.2020 – no class**

| Session 7: AI and Grand Challenges [SJ] | |
|---|---|
| **Learning Objective** | In this session we will cover the role of AI in international development, climate change, and health. There is a lot of promise but we will also consider the limitations of current technology solutions, and the interplay between technology and politics of the grand challenges. |
| **Required Readings** | ● "Tackling Climate Change with Machine Learning" (ArXiv, November 2019); ICML 2019 Workshop; NeurIPS 2019 workshop.<br>● "Combining satellite imagery and machine learning to predict poverty." Neal Jean et al. (Science, 2016). Two technical papers behind the work: one, two [skim]; and later extension "Predicting Economic Development using Geolocated Wikipedia Articles." Sheehan et al., ArXiv, May 2019.<br>● "High-performance medicine: the convergence of human and artificial intelligence." Eric J. Topol, Nature. |
| **Optional Readings** | ● "How AI-powered maps help improve vaccination campaigns and rural electrification" Facebook AI.<br>● "From Satellite to Village, Turning Data into Action." USAID, Medium.<br>● "AI and Compute." OpenAI<br>● "Energy and Policy Considerations for Deep Learning in NLP." Strubell et al. ArXiv, 5 June 2019.<br>● "Deep Learning, Natural Language Processing and Climate Change." Julian Harris, Medium, 6 October 2019.<br>● "Social media interventions for precision public health: promises and risks." Dunn et al., npj Digital Medicine, 2018.<br>● "Technology to advance infectious disease forecasting for outbreak management." George et al., Nature Communications, 2019. |

| | |
|---|---|
| | ● "[How artificial intelligence will affect the future of energy and climate](#)." David G. Victor, 10 January 2019. Brookings. |

| Session 8: Social Impacts and Future of Work [JB] | |
|---|---|
| **Learning Objective** | We discuss what "competition" with AI tells us about the nature of labour and human identity, and begin in what ways AI may contribute to increasing economic and power inequality. We also discuss positive visions: what kind of society would we like to achieve if digital technology is pervasive. Is that plausible? |
| **Required Readings** | ● Raji, I.D. and Buolamwini, J., 2019, January. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 429-435).<br>● Why Are There Still So Many Jobs? The History and Future of Workplace Automation, David H. Autor, Journal of Economic Perspectives, 29(3), Summer 2015, 3-30.File |
| **Optional Readings** | ● Weeden, Kim A. "Why do some occupations pay more than others? Social closure and earnings inequality in the United States." American Journal of Sociology 108, no. 1 (2002): 55-101.<br>● Korinek, A., _ Stiglitz, J. (2017). Artificial Intelligence and its implications for income distribution and unemployment, National Bureau of Economic Research, IncFile<br>● "What Would Happen If We Just Gave People Money?" by Andrew Flowers (FiveThirtyEight, 2016)URL |

| Session 9: Economics of AI [SJ] | |
|---|---|
| **Learning Objective** | In this session we discuss the broader impact of AI on the economy. We look at economic inequality, growth, trade, innovation, and industrial organisation. |
| **Required Readings** | ● "[Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics](#)." Erik Brynjolfsson, Daniel Rock, Chad Syverson, NBER Working Paper No. 24001<br>● "[The Impact of Artificial Intelligence on Innovation](#)" Iain M. Cockburn, Rebecca Henderson, Scott Stern, NBER Working Paper No. 24449 |

| | |
|---|---|
| | - "[Artificial Intelligence and Economic Growth](#)" Philippe Aghion, Benjamin F. Jones, Charles I. Jones, NBER Working Paper No. 23928.<br>- "[AI and International Trade](#)." Avi Goldfarb, Daniel Trefler, NBER Working Paper No. 24254.<br>- "[Artificial Intelligence, Economics, and Industrial Organization](#)." Hal Varian, NBER Working Paper No. 24839<br>- "[Artificial Intelligence and Its Implications for Income Distribution and Unemployment](#)." Anton Korinek, Joseph E. Stiglitz, NBER Working Paper No. 24174<br>- "[Political Consequences of Technological Change](#)." Special Issue in *Research & Politics*, 6(1), 2019. |
| **Optional Readings** | - "[The impact of artificial intelligence on international trade](#)." Joshua P. Meltzer, 13 December 2018. Brookings.<br>- Susan Athey (2018). "[The Impact of Machine Learning on Economics](#)." In *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press, 2018.<br>- Ajay Agrawal, Joshua Gans, and Avi Goldfarb. "[Prediction, Judgment, and Complexity: A Theory of Decision-Making and Artificial Intelligence](#)." In *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press, 2018. |


| Session 10: Guest lecture: German National AI Policy (provisional) | |
|---|---|
| **Learning Objective** | The guest speaker will cover policy and politics of AI in Germany and in comparative perspective. |
| **Required Readings** | Readings will be made available prior to class once the guest speaker is confirmed. |
| **Optional Readings** | |


| Session 11: Transnational Efforts at  Regulating AI [JB] | |
|---|---|
| **Learning Objective** | We discuss the international efforts being made to regulate AI, looking at a comparison between the world's three leading economies: the USA, China, and the EU; looking also at the G7, OECD, UN, OSCE, IGF, and corporate attempts like the efforts of Google, Microsoft and Facebook. Again we will consider a vision: can we imagine an ideal? What are nations for in the face of transnational technology? |

| Required Readings | ● The EU White paper (new version expected) |
| --- | --- |
| | ● "Perspectives on Issues in AI Governance", Google AI policy document, 2018. |
| | ● "Recommendation of the Council on Artificial Intelligence" (OECD Principles of AI), 2019. |
| | ● Ó hÉigeartaigh, Seán S., Jess Whittlestone, Yang Liu, Yi Zeng, and Zhe Liu. "Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance." Philosophy & Technology (2020): 1-23. |
| |     ○ Optionally , see the Chinese and Japanese versions on https://www.baai.ac.cn/research/translation-series-on-ai-ethics-en |
| Optional Readings | ● Kai-Fu Lee, AI Superpowers, 2018.URL |
| | ● White House (February 2019) "Executive Order on Maintaining American Leadership in Artificial Intelligence" |
| | ● Joanna J. Bryson, The Meaning of the EPSRC Principles of Robotics, Connection Science 29(2):130-136. |
| | ● Joanna Bryson. The Past Decade and Future of AI's Impact on Society chapter in Towards a New Enlightenment? A Transcendent Decade, published by BBVA OpenMind (2019). |


| Session 12: Student presentations | |
| --- | --- |
| Learning Objective | Transferable skills in communication of science and policy of AI. |
| Required Readings | ● Tips on communicating scientific results to policy makers: https://www.forbes.com/sites/marshallshepherd/2016/11/22/9-tips-for-communicating-science-to-people-who-are-not-scientists/ |
| Optional Readings | |