**STATES, CORPORATIONS, ROBOTS**

**MPhil Course Outline 2020-21**

Time: 1-3pm Monday, starting on 25th January 2021.

Overview

The course is designed to examine the historical parallels and potential future relationship between the rise of states and corporations and the rise of artificial intelligence (AI). It starts from the premise that states and corporations have more in common with AIs than we might think. They are artificial agents (AAs), often described as having minds or wills or lives of their own. States and corporations are constructs designed to pool information, intelligence and other resources to engineer wider efficiencies and to generate collective outcomes. These outcomes transcend the input that generated them. States and corporations are meant to be different from, and more than, the sum of their parts.

There are also clear and obvious differences between AAs and AIs. States and corporations long pre-date even basic computing, never mind AI. They do not necessarily have more intelligence than the humans who constitute them. The algorithms that underpin them are often extremely crude by the standards of twenty-first century technology: constitutions, voting systems, laws, business regulations. Yet at the same time, states and corporations have enormous power: as agents they are much more powerful – for now at least – than AIs. These differences raise interesting questions. Can artificial agency be distinguished from artificial intelligence? Do collective human agents have power without responsibility? What happens when states and corporations integrate their agency with the capabilities of machine-learning technology? What will be the future relationship of the AAs and the AIs?

A reading list and class itinerary is below. There is not much of a specific literature on these questions – they are rarely asked in this form. This is an experimental course being taught for the first time to explore these important issues. It aims to synthesise different literatures – from history of political thought, IR, philosophy, political economy and technology studies – where different aspects of these problems are discussed. It starts with one of the founding texts of political philosophy, IR and the history of AI: Thomas Hobbes’s *Leviathan*. The book, published in 1651, begins with the startling claim that the state is an automaton i.e. a kind of robot. What does that mean? Should we take Hobbes seriously? If so, what follows?

The course will be taught through weekly 2 hours seminars in Lent term. The final week is left free for a general discussion and/or student presentations.

Assessment will be via a 3000-word essay at the end of term (19th March 2021).

Week 1 – The Leviathan as Robot

What are the implications of Hobbes’s description of the state as a kind of robot? Does this have any connection with artificial intelligence? What does it imply for the state’s role both domestically and in the international order?

***Reading***

Thomas Hobbes, *Leviathan* (ed. Richard Tuck) (Cambridge University Press, 1996)

Noel Malcolm, *Aspects of Hobbes* (Oxford University Press, 2002)

* Chapter 7: ‘The Title Page of Leviathan, Seen in a Curious Perspective’
* Chapter 13: ‘Hobbes Theory of International Relations’

Quentin Skinner, ‘Hobbes and the Purely Artificial Person of the State’, *Journal of Political Philosophy* (1999, 7: 1-29)

David Runciman, ‘What Kind of Person is Hobbes’s State? A Reply to Skinner’, *Journal of Political Philosophy* (2000, 8: 268-278)

Samantha Frost, *Lessons of a Materialist Thinker: Hobbesian Reflections on Ethics and Politics* (Stanford, 2008)

Minsoo Kang, *Sublime Dreams of Living Machines. The Automaton in the European Imagination* (Harvard University Press, 2011)

George Dyson, *Darwin Among the Machines* (Penguin, 2012)

* Chapter 1: ‘Leviathan’

Ziyaad Bhorat, ‘Automata in the Political Thought of Thomas Hobbes’ (working paper 2019)

<http://www.wpsanet.org/papers/docs/WPSA_Hobbes%20Automata_Ziyaad%20Bhorat.pdf>

Week 2 – Groups as Agents

Do groups have minds of their own? Could they even have consciousness? What might group agency mean in the absence of group intelligence?

***Reading***

Christian List and Philip Pettit, *Group Agency: The Possibility, Design and Status of Corporate Agents* (Oxford University Press, 2011)

Philip Pettit, ‘Groups with Minds of their Own’, in *Socializing Metaphysics* (ed. F. Schmitt) (Rowman & Littlefield, 2003)

Christian List, ‘What is it Like to be a Group Agent?’, *Nous* (2018, 52: 295-313)

D.P. Tollefsen, *Groups as Agents* (Polity Press, 2015)

R. Briggs, ‘The Normative Standing of Group Agents’, *Episteme* (2012, 9: 283-9)1

Toni Erskine (ed.), *Can Institutions Have Responsibilities? Collective Moral Agency and International Relations* (Palgrave MacMillan, 2003)

S. Logan, T. Erskine, 'New Technologies: Questions of Agency, Responsibility, and Luck', in *Technologies of International Relations: Continuity and Change* (Springer International Publishing, 2018)

David Chalmers, *The Conscious Mind* (Oxford University Press, 1996)

Kathryn Hayles, ‘Cognitive Assemblages: Technical Agency and Human Interactions’, *Critical Inquiry* (2016, 43: 32-55)

James Suroweicki, *The Wisdom of Crowds: Why the Many are Smarter than the Few* (Doubleday, 2004)

Week 3 – The Agency of the State

Is the state a person in its own right? How does its agency get expressed? What can states do that other forms of collective human agents can’t?

***Reading***

Alexander Wendt, ‘The state as person in international theory’, *Review of International Studies* (2004, 30: 289-316)

Sean Fleming, ‘Leviathan on Trial: Should States be Held Criminally Responsible?’, *International Theory* (2020)

David Runciman, ‘Is the State a Corporation?’, *Government and Opposition* (2000, 35: 90-104)

Jens Bartleson, *The Critique of the State* (Cambridge University Press, 2001)

Margaret Canovan, *The People* (Polity Press, 2005)

Hannah Arendt, *The Human Condition* (Chicago University Press, 1958)

Michael Oakeshott, *On Human Conduct* (Clarendon Press, 1975)

E. Schitzgebel, ‘If Materialism is True, the United States is Probably Conscious’, *Philosophical Studies* (2015, 172: 1697-1721)

Quentin Skinner, ‘What is the State?’ (2010) <https://vimeo.com/14979551>

Peter Lomas, *Unnatural States: The International System and the Power to Change* (Routledge, 2017)

James Scott, *Seeing Like A State* (Yale University Press, 1999)

Mark Neocleous, *Imagining the State* (Open University Press, 2003)

Week 4 – The Agency of the Corporation

Do corporations have independent agency? How does this relate to the agency of the state? Should corporations have rights comparable to the rights of individuals?

***Reading***

F. W Maitland, St*ate, Trust and Corporation* (ed. D. Runciman & M. Ryan) (Cambridge University Press, 2003)

* Chapter 4: ‘Moral Personality and Legal Personality’

Peter Drucker, *The Concept of the Corporation* (Transaction Publishers (2006 [1946])

Adam Winkler, *We the Corporations: How American Businesses Won Their Civil Rights* (Liveright, 2018)

Kent Greenfield, ‘If Corporations are People They Should Act Like it’, *The Atlantic* (Feb 2015) <https://www.theatlantic.com/politics/archive/2015/02/if-corporations-are-people-they-should-act-like-it/385034/>

Leonardo Davoudi, Christopher McKenna, Rowena Olegario, ‘The Historical Role of the Corporation in Society’, *Journal of the British Academy* (2018, 6: 17-47)

Bruce Kogut, Ugo Zander, ‘Knowledge of the Firm and the Evolutionary Theory of the Multinational Corporation’, *Journal of International Business Studies* (1993: 625-45)

Jessica Green, *Rethinking Private Authority* (Princeton University Press, 2015)

Maha Rafi Atal, ‘The Nature of (Corporate) Power: Making Room for Ideology in the Study of Private Governance’ (draft paper, from author)

M. M. Blair, ‘Of Corporations, Courts, Personhood and Morality’, *Business Ethics Quarterly*, (2015, 25: 415-31)

A. Singer, *The Form of the Firm: A Normative Political Theory of the Corporation* (Oxford University Press, 2019)

Week Five – The Rise of Artificial Intelligence

Are machines capable of intelligent thought? How does machine agency compare to other forms of artificial agency? Should robots have rights?

***Reading***

Samuel Butler, ‘Darwin Among the Machines’ (June 1863) <http://nzetc.victoria.ac.nz/tm/scholarly/tei-ButFir-t1-g1-t1-g1-t4-body.html>

Stuart Russell, *Human Compatible: AI and the Problem of Control* (Allen Lane, 2019)

Steven Pinker, Stuart Russell, ‘The Foundations, Benefits and Possible Existential Threat of AI’, *Future of Life Institute* podcast transcript

<https://futureoflife.org/2020/06/15/steven-pinker-and-stuart-russell-on-the-foundations-benefits-and-possible-existential-risk-of-ai/>

Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford University Press, 2014)

Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence (*Allen Lane, 2017)

Margaret Boden, *AI: Its Nature and Future* (Oxford University Press, 2016)

Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (Harper Collins, 2017)

Anne Phillips, *The Politics of the Human* (Cambridge University Press, 2016)

B. Russell (ed.), *Robots: the 500-year Quest to Make Machines* Human (Scala, 2017)

David Gunkel, *Robot Rights* (MIT Press, 2018)

Week Six – The Singularity?

Is AI going to transform the human condition? How does this compare to the transformation of human societies by modern states and corporations? Will AI transform AA?

***Reading***

Peer Vries, *Escaping Poverty: The Origins of Modern Economic Growth* (Vienna University Press, 2013)

Gregory Clark, *A Farewell to Alms: A Brief Economic History of the World* (Princeton University Press, 2007)

Daron Acemoglu & James Robinson, *The Narrow Corridor: States, Societies and the Fate of Liberty* (Penguin, 2020)

Jonathan Daly, *Historians Debate the Rise of the West* (Routledge, 2015)

Toby Ord, *The Precipice: Existential Risk and the Future of Humanity* (Bloomsbury, 2020)

Ray Kurzweil, *The Singularity is Near* (Viking, 2005)

David Chalmers, ‘The Singularity: A Philosophical Analysis’, *Journal of Consciousness Studies*, (2010, 17: 7-65)

U. Awret (ed.), *The Singularity: Could Artificial Intelligence Really Outthink Us (and Would We Want it to)?* (Imprint Academic, 2016)

Katherine Hayles, *How We Became Posthuman* (University of Chicago Press, 1999)

Week Seven – States, Corporations, Robots

Where does the balance of power lie between states, corporations and intelligent machines? Does the international order suggest different models of this relationship? Are states or corporations better equipped to handle the impact of AI?

***Reading***

Shoshanna Zuboff, *The Age of Surveillance Capitalism* (Profile Books, 2019)

Kai-Fu Lee, *AI Superpowers: China, Silicon Valley and the New World* *Order* (Houghton Mifflin, 2019)

Rana Foroohar, *Don’t Be Evil: The Case Against Big Tech* (Allen Lane, 2019)

Langdon Winner, ‘Do artifacts have politics?’, *The Whale and the Reactor* (Chicago University Press, 1986)

John Naughton, ‘The evolution of the Internet: from military experiment to General Purpose Technology’, *Journal of Cyber Policy* (1, 2016)

M. Babic et al, ‘State vs Corporation: Rethinking the Power of Business in International Politics’, *The International Spectator* (2017:4)

Judy Wajcman, ‘Automation: Is it Really Different this Time?’, *British Journal of Sociology* (2017, 68: 119-27)