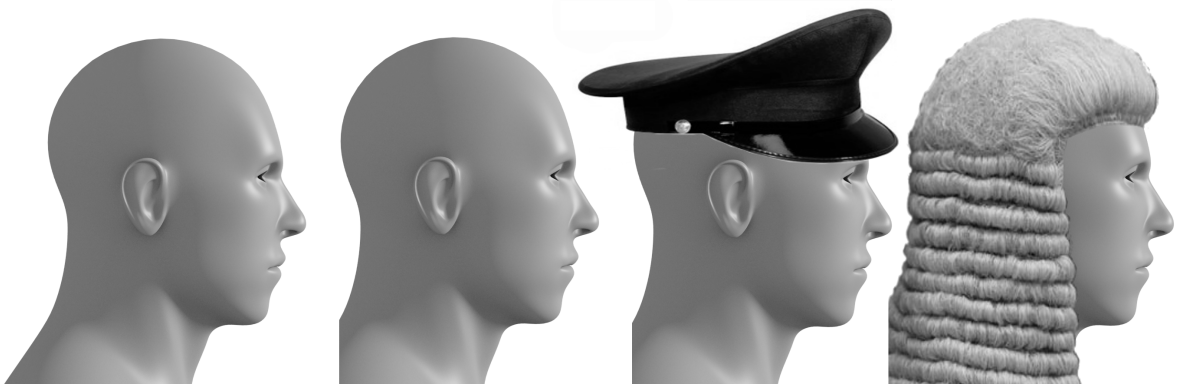
Roundtable on Uses of Artificial Intelligence in the Criminal Justice System

St Anne's College, Oxford
Seminar Room 7

November 23-24, 2017



Artificial Intelligence (AI) systems are increasingly widely used within the field of criminal justice. Police forces use AI systems to predict crime hotspots, to identify individuals at risk, and to assist on-the-ground policing and detective work. AI systems are also used in the courts, to gauge a defendant's risk of reoffending, and inform bail, sentencing and parole decisions. Similar systems are in use in prisons, to inform rehabilitation programmes.

While statistical systems have been in use in criminal justice for many years, the systems coming into use now are increasingly sophisticated, and are finding a variety of new roles. Their increasing prevalence raises many questions. How should the accuracy of these systems be measured? How can we ensure their operation is not biased towards particular social groups? How can we inspect the processes through which they reach a decision? How should human decision-makers interact with such systems? What ethical and legal frameworks do we need to ensure good practice in the use of such systems?

Many of these questions arise wherever AI systems are used to inform decision-making, whether in public policy or by the companies who shape our experiences on the internet. But the decisions made in the criminal justice system have a particularly large impact on people's lives: so the use of AI systems in these decisions demands particularly urgent attention.

This roundtable brings together experts working on AI and criminal justice from several perspectives. It includes lawyers, policy researchers, AI technologists, statisticians, ethicists and police officers. The event will be structured around five issues that we think are especially urgent to discuss.

- **Accuracy:** How reliable are the system's predictions or judgements? How can it be tested for accuracy? Should the results of such evaluations be made public?

- **Bias:** Is the system discriminatory towards particular social groups? Is bias ever acceptable, if it leads to higher accuracy? Are there ways of removing bias without compromising accuracy?

- **Control:** How do human decision-makers interact with the system? How can we most productively combine human decisions with the system's processes? What control should humans have over the system's outputs?

- **Transparency:** Should there be a requirement that the system's outputs be 'explainable'? If so, how can/should explanations be provided? Can these explanations be provided without infringing on individuals' privacy, or (for commercial systems) disclosing proprietary code?

- **Oversight and Regulation:** What ethical or legal frameworks should be established to ensure good practice on all of the above issues?

Thursday 23 November

ORIENTATION

12 – 1 pm

From Predpol to Parole: AI in policing and criminal justice

Colin Gavaghan
Associate Professor of Law, Otago

Alistair Knott
Associate Professor of Computer Science, Otago

James Maclaurin
Associate Professor of Philosophy, Otago

Lunch (provided)

ISSUES IN AI, POLICING AND CRIMINAL JUSTICE

1.45 – 3 pm

Accuracy

Nikolaos Aletras [1.45 – 2.15 pm]
Applied Scientist, Amazon

Discussion [2.15 – 3 pm]

Tea and coffee

3.30 – 5.15 pm

Bias

Janet Bastiman [3.30 – 4 pm]
Chief Scientist, StoryStream

Helena Webb [4 – 4.30 pm]
Senior Researcher in Social Science and Computing, Oxford

Discussion [4.30 – 5.15 pm]

Friday 24 November

ISSUES IN AI, POLICING AND CRIMINAL JUSTICE, cont'd.

9 – 10.45 am

Transparency

Alistair Knott [9 – 9.30 am]
Associate Professor of Computer Science, Otago

Michael Veale [9.30 – 10 am]
PhD Researcher in Responsible Public Sector Machine Learning, UCL

Discussion [10 – 10.45 am]

Tea and coffee

11.15 am – 1 pm

Control

Nigel Harvey [11.15 – 11.45 am]
Professor of Judgment and Decision Research, UCL/LSE

James Maclaurin and Colin Gavaghan [11.45 – 12.15 pm]
Associate Professors of Philosophy and Law, Otago

Discussion [12.15 – 1 pm]

Lunch (provided)

1.45 – 3.30 pm

Solutions: Oversight and Regulation

Jamie Grace [1.45 – 2.30 pm]
Senior Lecturer in Law, Sheffield Hallam

Discussion [2.30 – 3.30 pm]

Tea and coffee

4 – 5 pm

Future Directions: Where to Now?

Julian Savulescu [4 – 4.30 pm]
Uehiro Chair in Practical Ethics, Oxford

Discussion [4.30 – 5 pm]

Close

SPEAKERS

Nikolaos Aletras is an Applied Scientist at Amazon Research Cambridge. Prior to that, he worked as a Research Associate at the Department of Computer Science at UCL and completed a PhD in Natural Language Processing at the Department of Computer Science at the University of Sheffield. His main research interests are in Natural Language Processing and Machine Learning. He is particularly interested in applying statistical methods for detecting and interpreting the underlying topics in large volumes of text data. He also develops methods to analyse text and uncover patterns in data to solve problems in other scientific areas such as social and legal science.

Janet Bastiman is Chief Science Officer at StoryStream where she leads a team of talented Machine Learning researchers to solve a multitude of problems across all fields of AI. She is treasurer of the UK IEEE STEM committee, co-founder of Tech Women London meetup and a regular speaker and writer on all aspects of AI. She provided written evidence to the recent UK parliamentary inquiry on algorithmic transparency and is particularly concerned with ensuring that the algorithms that will play a part in our future are fair and unbiased. Janet holds a PhD in Computational Neuroscience and an undergraduate degree in molecular and cellular biochemistry.

Colin Gavaghan is an Associate Professor in the Faculty of Law at the University of Otago. He is the first director of the New Zealand Law Foundation sponsored Centre for Law and Policy in Emerging Technologies. The Centre examines the legal, ethical and policy issues around new technologies. To date, the Centre has carried out work on biotechnology, nanotechnology, information and communication technologies and artificial intelligence. In addition to emerging technologies, Colin lectures and writes on medical and criminal law. He is a member of the Advisory Committee on Assisted Reproductive Technology and the Advisory Board of the International Neuroethics Network. He was an expert witness in the High Court case of *Seales v Attorney General*, and has advised members of parliament on draft legislation. He is co-convenor of the AI and Society discussion group at Otago, and co-investigator in the AI and Law in New Zealand project.

Jamie Grace is a Senior Lecturer in Law in the Department of Law and Criminology at Sheffield Hallam University, holding this post since January 2014. He is the current course leader of the MA and LLM in Applied Human Rights courses taught in his Department. Jamie is an active researcher in the Helena Kennedy Centre for International Justice, and is also a Fellow of the Sheffield Institute of Policy Studies, both part of Sheffield Hallam University. Jamie was previously Senior Lecturer in Law in the School of Law & Criminology at the University of Derby (where he undertook various roles from September 2007 until January 2014). Jamie was an active researcher in the Law in Society Research Group at the University of Derby from January 2010 to January 2014, submitting an impact case study in relation to his research on information law and privacy in healthcare to the 2014 Research Excellence Framework.

Nigel Harvey is Professor of Judgment and Decision Research in the Department of Experimental Psychology at University College London and Visiting Research Fellow in the Department of Statistics at the London School of Economics and Political Science. He is a past President of the European Association for Decision Making. He is an Associate Editor of the International Journal of Forecasting and on the Editorial Board of the Journal of Behavioral Decision Making. His current research deals with how people integrate their judgment with statistical forecasts and with identifying when judgment improves and when it impairs the final forecasts.

Alistair Knott is an Associate Professor at the Department of Computer Science in the University of Otago, New Zealand. He studied Psychology and Philosophy at Oxford University, then took an MSc and PhD in Artificial Intelligence at Edinburgh University. Ali has worked in AI for 25 years, focussing on models of natural language processing, human-computer dialogue and neural models of language and memory; he has published over 100 papers on these topics. He also works for the Auckland-based AI startup Soul Machines, where he is implementing the embodied model of language developed in his book *Sensorimotor Cognition and Natural Language Syntax* (MIT Press, 2012). He is co-convenor of the AI and Society discussion group at Otago, and co-investigator in the AI and Law in New Zealand project.

James Maclaurin is an Associate Professor in the Department of Philosophy and Associate Dean for Research in Humanities at the University of Otago. His MA in biological applications of mathematical information theory is from Victoria University of Wellington and his PhD in the philosophy of science is from the Australian National University. His research focuses on the relationship between science, public policy and ethics. His books include *What is Biodiversity?* (with Kim Sterelny, University of Chicago Press) and *A New Science of Religion* (with Greg Dawes, Springer Science). He has also published on philosophical methodology and on the application of evolutionary science in economics and computer science. He has a long-standing interest in the use of information technology in higher education, in which context he convenes the University of Otago's Information Technology Advisory Committee which advises the Vice Chancellor on the university's use of new information technology. He is co-convenor of the AI and Society discussion group at Otago, and co-investigator in the AI and Law in New Zealand project.

Julian Savulescu has held the Uehiro Chair in Practical Ethics at the University of Oxford since 2002. He has degrees in medicine, neuroscience and bioethics. He directs the Oxford Uehiro Centre for Practical Ethics within the Faculty of Philosophy, and leads a Wellcome Trust Senior Investigator award on Responsibility and Health Care. He directs the Oxford Martin Programme for Collective Responsibility for Infectious Disease at the Oxford Martin School at the University of Oxford. He co-directs the interdisciplinary Wellcome Centre for Ethics and Humanities in collaboration with Public Health, Psychiatry and History. In 2017, he joined the Murdoch Childrens Research Institute, spending four months per year as Visiting Professorial Fellow in Biomedical Ethics where he is working to establish a programme in biomedical ethics, and Melbourne University as Distinguished International Visiting Professor in Law. He is a leader in medical and practical ethics, with more than 300 publications, an h index of 55 and over 11, 000 citations in total. He is Editor of the *Journal of Medical Ethics*, the highest impact journal in the field, and founding editor of *Journal of Practical Ethics,* an open access journal in Practical Ethics. He received an honorary doctorate from the University of Bucharest in 2014.

Michael Veale is a PhD researcher at the Department of Science, Technology, Engineering and Public Policy at University College London. He studies how issues of fairness, accountability, privacy and resilience that arise around machine learning technologies intersect with the law, as well as in the day-to-day practices of public sector modellers designing and deploying these systems today. His research has been drawn upon by international governments, regulators, and in seminal policy reports on the future of AI governance. He is a machine learning advisor at the Red Cross Red Crescent Climate Centre, was drafting author and consultant on the Royal Society and British Academy report 'Data Management and Use: Governance in the 21st Century', and previously worked at the European Commission on ageing and the IoT. He tweets at @mikarv.

Helena Webb is a Senior Researcher at the Department of Computer Science, University of Oxford. She is a social scientist by training and works in the Human Centred Computing research group, which studies the inter-relationships between computing and social practices. She is currently working on the UnBias project. This collaboration between the Universities of Nottingham, Oxford and Edinburgh analyses the user experience of algorithm-driven internet services and the processes of algorithm design. It investigates controversies around 'bias' in the use of algorithms and also solicits stakeholder concerns about these issues. The project will produce a series of resources to inform policy, user awareness and developer understanding in relation to algorithms and 'fairness'.