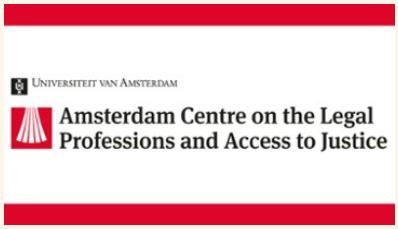
REPORT APRIL 2021

Judging The Robojudge

Ethical and Legal perspectives on using Artificial Intelligence in Legal Decision-Making Procedures

Nederlandse Onderzoeksschool Wijsbegeerte
Amsterdam Centre on the Legal Professions and Access to Justice
Vereniging voor Wijsbegeerte van het Recht



ORGANISERS

Marjolein Lanzing
University of Amsterdam

Iris van Domselaar
University of Amsterdam

KEY NOTE SPEAKERS

7th April

Frederik Zuiderveen Borgesius
Radboud University

Mireille Hildebrandt
Radboud University

8th April

Sven Nyholm
Utrecht University

Alicia Solow-Niederman
Harvard Law School

9th April

Tania Sourdin
Newcastle University, NSW

Caspar Chorus
Delft Technical University

Pim Haselager
Radboud University

Introduction and welcome by:

Iris van Domselaar
University of Amsterdam

Marjolein Lanzing
University of Amsterdam

- Introduction to the theme of Robojudge and focus on the interdisciplinary nature of this topic
- Highlight on AI and courts: how AI is used in the context of courts and laws for various purposes, such as for organising information, for communication, advise, predictions, deciding legal cases in courts
- Main challenges brought up by the speakers regarding this topic:
  - The issue of responsibility
  - The issue of discrimination
  - The issue of commercialisation vs public value
  - The challenge of independence
  - The challenge of Rule of Law/Justice
  - The issue of legitimacy/alienation (value of AI judges for citizens- are citizens okay with AI making decisions for them? This leads to the problem of alienation)
- For the future, some of the points are to be kept under consideration, such as governmental regulation, professional regulation, self-regulation, legal education and individual responsibility

Speaker

Mireille Hildebrandt
Radboud University

Presentation: The need to develop a philosophy of technology for computational law (and for modern positive law)

The speaker broadly covered the following topics during the session:

- Emphasised that for law to be articulated in code, it is important that lawyers understand the assumptions of computer science and that computer scientists understand the assumption of law and the rule of law
- Mentioned the need for a cross disciplinary research for example: Journal of Cross-Disciplinary Research in Computational Law (CRCL) https://journalcrcl.org/crcl
- Elaborated on text-driven laws: what is meant by text-driven laws, understanding text-driven laws in the context of legal texts and its effect within the legal domain
- Based on the above, the speaker opened up the floor to the audience addressing two main questions:
  - What should be preserved from text-driven laws?
  - What can be done better?
- Elaborated on code-driven laws: what is meant by code-driven laws and its effects within the legal domain
- Based on the above, the speaker raised the following two questions addressed to the audience:
  - When should code-driven 'laws' be resisted?
  - When and how should we integrate code-driven laws?

- Elaborated on data-driven laws: what is meant by data-driven laws and its effects within the legal domain
- The above topic of discussion raised the following two questions:
  - When should data-driven 'laws' be resisted?
  - When and how should we integrate data-driven laws?
- Some of the questions raised by the audience to the speaker concerned the following themes:
  - The issue of biases in legal domain and machine learning
  - The issue of interpretability of law when approaching computer science and law
- Concluding thoughts by the speaker highlighted the importance of interaction between lawyers and computer systems, as opposed to a rejection of computer system or the replacement of lawyers with computer systems

Speaker

Frederik Zuiderveen Borgesius
Radboud University

Presentation: Strengthening legal protection against discrimination by algorithms and artificial intelligence

The speaker broadly covered the following topics during the session:

- Touched upon the discrimination by Artificial Intelligence and Machine Bias, for example, in the case of software COMPAS
- Elaborated on the two most relevant legal protection laws within Europe that can safeguard against algorithmic discrimination (discrimination against people that can happen as a result of automated AI decisions), namely Data protection law and Non-discrimination law. Along with focusing on their limitations:
- Data protection law: EU's General Data Protection Regulation (GDPR)
  - Strengths of GDPR:
    1. Protecting the privacy of individuals
    2. Protecting fairness & human rights when personal data is used
  - Main roles of GDPR:
    1. For legal transparency it is important that organisations offer transparency about the use of individuals' personal data
    2. That it must conduct Data Protection Impact Assessment (DPIA) when dealing with individuals' personal data to minimise risk to the individual

3. Requirement of specific rules on automated decision- making. In some cases, fully automated decision making is present, however, people subjected to it can ask for intervention. The r right to explanation: the organisations must provide meaningful explanations as to the logic involved in producing a certain result through automated decision-making

- Weakness of GDPR:
  1. It only applies to personal data (which in some cases can be a strength but a drawback in other cases)
  2. It is nuanced but still vague
  3. The issue of compliance and enforcement deficit
  4. Difficult in explaining AI decisions
- Thus, GDPR is necessary but it is not sufficient

- Non-discrimination law:
  - Non-discrimination law can help protect individuals against algorithmic discrimination
  - It is important to note that discrimination may also be due to deliberate human decision making, however algorithmic decisions can also be discriminatory as it can learn from the discriminatory decisions made by humans
  - Two key points within this include: 1. Direct discrimination and 2. Indirect discrimination
    1. Direct discrimination: Non-discrimination law can protect people from direct discrimination due to algorithmic decisions
    2. Indirect discrimination: Non-discrimination law can protect people from discrimination by considering the concept of indirect discrimination as well

- Weakness of non-discrimination law: (1) Indirect discrimination ban is nuanced but vague (2) Automated discrimination can be hidden
- Thus, non-discrimination law is necessary but it is not sufficient

- In response to the above, the speaker mentioned a new field of computer science that focuses on fairness and non-discrimination in AI, namely, ACM FAccT (ACM Conference on Fairness, Accountability, and Transparency)
- The speaker presented concluding remarks on the topic of discussion which highlighted the following:
  - The need for cooperation between different disciplines
  - Enforce GDPR and non-discrimination law
  - Examining the additional rules needed besides the aforementioned ones
- Some of the questions raised by the audience to the speaker concerned the following themes:
  - The black box problem of AI: to what extent can AI decisions be explained?
  - Privacy concerns
  - Direct and indirect discrimination

Speaker

Sven Nyholm
Utrecht University

Presentation: Responsibility Gaps, value alignment and meaningful human control over AI

The speaker broadly covered the following topics during the session:

- The speaker focused on how 'responsibility gaps' associate to 'value alignment problems' regarding AI
- What is meant by responsibility gaps? A responsibility gap in the context of technology and human engagement, points to the gap in attributing responsibility to either technology or humans when something (bad) happens
- What is meant by value alignment? It means for AI to be able to match up to the values of the domain in which they are used. If AI matches/aligns to the said values, then it would be considered value aligned. However, if an AI does not match/does not align with the said value then it will give rise to a value alignment problem
- The speaker emphasised on two main approaches to responsibility:
  1. First approach considers (i) negative responsibility (a case where something bad happens which requires for blame to be assigned to an actor and punish them) and (ii) positive responsibility (a case where something good happens which calls for a reward or praise to an actor)

2. Second approach considers (i) backward looking responsibility (after something good or bad has happened and we want to assess who is responsible for it) and (ii) forward looking responsibility (it is about mitigating the risk when something bad happens or assessing what/who contributed to the instance where something good happened)

- The speaker further described the four combinations that can be made possible when considering the aforementioned approaches to responsibility: (1) Negative-Backward looking responsibility (2) Negative-Forward looking responsibility (3) Positive-Backward looking responsibility (4) Positive-Forward looking responsibility
- Mentioned some of the reasons why responsibility gap might arise: (1) Novel form of AI (2) 'the problem of many hands' (3) No strong basis for positive-forward looking responsibility
- Some problems with making AI value aligned:
  1. A weak common sense duty to encourage good outcomes than avoiding bad ones
  2. Different notions of good or which values matter the most
  3. Making AI more value aligned may make it difficult to control
- Elaborated on the solutions to solving responsibility gap problem: 1. Collaborative agency between AI and humans
  2. Santoni de Sio & Mecacci's "track & trace" theory, AI system should track human values and it should be possible to trace the system back to humans who understand how it works (however, some of the key worries remain)
  3. Re-conceptualising positive-forward looking responsibility gaps
- Some of the questions raised by the audience to the speaker concerned the following themes:
  - Agency of robots; 'agency laundering' and information technology and how are agency solutions relevant for law?

Speaker

Alicia Solow-Niederman
Harvard Law School

Presentation: Contextualizing Artificially Intelligent Justice

The speaker broadly covered the following topics during the session:

- Talked about what a robot judge means in legal system along with how the different forms of AI will change more than the mechanics of adjudication as they will also impact the adjudicatory values which are held by both legal actors and the public
- Touched upon how technology interacting with new systems creates newer challenges
- Talked about two contrasting models of legal change i.e. Rule Updating and Value Updating
- Elaborated on two adjudicatory codified justice and equitable justice, and specifically on the predicted effect of AI adjudication in the context of equitable justice and codified justice
- What is new about robojudge: Codified justice
- Touched upon the changes that will be brought about due to AI adjudication. Concerns regarding it include:
  1. Incomprehensibility (compromises understanding)
  2. Datafication (compromises adaptation)
  3. Effect on legal actors: disillusionment (compromises trust) and (
  4. Alienation (compromises participation)

- Potential solutions to the above:
  1. There is no perfect fix
  2. Four responsive approaches: (i) Experimentation (ii) Division of Labour (iii) 'Coding for Equity' and (iv) Market intervention
- New ways of AI can be good and helpful
- Some of the questions raised by the audience to the speaker concerned the following themes:
  - The influence of these solutions on legal representation
  - Is AI better suited for law or facts?

Speaker

Tania Sourdin
Newcastle University, NSW

Presentation: Judge v Robot?

The speaker broadly covered the following topics during the session:

- Judged AI and technology: a dystopia or an opportunity
- Touched upon facial recognition technology used in criminal conviction
- Elaborated on three levels of technological change that are shaping the justice system:
  1. Supportive technology: technology assisting in informing, supporting and advising the people involved in justice activities, for example in conducting remote court hearings, e-forms, justice café and so on
  2. Replacement technology: technology replacing the functions and activities previously done by humans, for example in case management, letters and so on
  3. Disruptive technology: technology changing the way in which determinate, advisory and facilitation processes work and AI replaces some decision making through means of AI and analytics
- Emphasis on the Digital Age: people are more connected than ever before

- Further elaboration on the three levels of technological changes in the context of justice reform (these can have both advantages and drawbacks):
  1. Supportive technology: people now access justice services online and can obtain information about justice processes. With the help of web-based information systems people can access different legal alternatives. Supportive technology enables people to access legal support and legal services while being online. However, this also results in shift to managing disputes online which can be beneficial, but also presents its own challenges
  2. Replacement technology: There has been significant growth in this. Services such as email, video calling, web-based information portals can assist, support or even replace in-person court procedures. This can also result in altering the environment in which court hearings tend to take place. Besides this, it can be effective in reducing cost and time among other certain advantages
  3. Disruptive technology: There are possible benefits of this type of technology such as more access to legal information and legal options, but it also raises issues of hostility, stress and disconnection. Job loss is another point within this. Besides this there are threats to privacy and the challenge of loss of social interaction among people
- The evolution of AI can have big impacts on adjudicative processes. How will the role of the Judge change? Who drives these reforms? And what data can be used to train Judge AI? The following points are to be considered: displacement, control, legality of decisions made by AI, translating law into code, and discretionary judgement
- Developments in some jurisdictions may not be appropriate in others

- Concluding remarks focused on the possible changes in the years to come: AI can replace some people as it can make use of automated responses, Big Data analytics and services such as visual recognition will replace many jobs and can transform courts. Replacing some decision-making processes with simple technology. And lastly, with some jobs parts, of it will be replaces by automation
- Some of the questions raised by the audience to the speaker concerned the following themes:
  - How to detect people who are who are trying to gain something out of the system?
  - The question of equity in the context of adversarial system
  - How to access people who are digitally illiterate?

Speakers

Caspar Chorus
Delft Technical University

Tania Sourdin
Newcastle University, NSW

Pim Haselager
Radboud University

Round Table on "Judge v Robot"

The round table broadly focused on the following topics:

- Bias in judicial decision making. Can disputes ever be settled without bias in the context of AI decision making?
- The responses to the above topic were varied and they raised many important concerns:
  - As there are different connotations of bias, it is useful to have a working definition of the term bias. Biases can also often evolve over time. It may be impossible to have a bias free environment as even human decision making can also contain bias and thus this can apply in the case of AI decision making as well
  - The role that culture plays in generating/keeping up the biases. AI can help recognise certain biases, but it can also impose an over uniformity
  - Bias may also open up the scope for creativity in some cases
  - Bias is value laden

- Within the definition of the term bias, people may be concerned with different aspects of it, the question then becomes whether one is transparent in what argument is presented. AI can present an opportunity to potentially get rid of some human biases but at the same time it can also introduce new biases. A multi person panel of people can also introduce more and new biases

- With AI decision making, humans may fear losing their ability to make decisions due to a reliance on AI decision making. Thus AI can be used as a helping tool that can help humans reflect and think on their own decision and the legal processes, in place of replacing humans completely with AI and relying on AI completely. AI can be used as critical 'friend'
- Questions concerning human dignity: how to make sense of decisions made by AI vs the decisions made by humans, when a person is at the receiving end of a decision
- The role educating lawyers about the computing systems used in legal practices, which can also help address certain ethical questions within the field of law
- The issue of transparency with AI decision making and its involvement in legal domain