Volume **6** (3) 2024 https://doi.org/10.5204/lthj.3635

The Human in the Feedback Loop: Predictive Analytics in Refugee Status Determination

Niamh Kinchin
University of Wollongong, Australia

Abstract

Predictive analytics in law describes AI systems that can predict the outcome of legal cases via computational legal reasoning. Predictive analytics in refugee status determination (RSD) has been cautiously employed. However, active research into model development holds significant potential for scalable application. Predictive analytics, traditionally built upon inductive decision-making processes like supervised machine learning and decision trees, risks compromising the abductive reasoning processes that RSD relies upon. Even if models are built to effectively navigate this hurdle, problems with data remain. Insufficient data is an intractable element of forced displacement, which means inaccuracies or uncertainty. The prospective nature of a well-founded fear undermines how algorithms are traditionally trained from historical data. The inability of predictive analytics to measure subjective fear may mean that RSD credibility assessments are pushed to 'pseudo-scientific' tools such as lie detectors and emotion recognition technology. Without the ability to remove subjectivity, the training data will capture subjective fear from previous claimants that will inform other case outcomes. Case characteristics that describe an individual case within a dataset and upon which the algorithm is trained must be carefully chosen in consultation with legal experts. Without considering how case characteristics reflect legal standards, they risk subverting hard-fought and won international legal protections. Finally, commonly acknowledged issues of algorithmic bias and harmful feedback loops must not be forgotten. This conversation is crucial for ensuring that the likely implementation of predictive analytics in RSD upholds fairness and established legal standards and, crucially, does not forget the human at its heart.

Keywords*:* Predictive analytics; data; international refugee law; legal reasoning; refugee status determination.

1. Introduction

One of the most significant yet often overlooked aspects of the 'new wave' of artificial intelligence (AI) and its integration into all aspects of modern life is its rich potential for cross-disciplinary dialogue. These conversations, particularly between lawyers and computer scientists, are crucial because they may hold keys to increased efficiency and accuracy in legal administration and access to justice. However, lawyers and computer scientists do not always speak the same language. In their 2017 article 'Playing with Data', Lehr and Ohm present a compelling argument for lawyers to shift their perception of machine learning from a monolithic and abstract concept to a nuanced process that encompasses problem definition, data collection and model deployment, with several stages in between.[1] At the same time, in their enthusiasm for increasingly clever machines, computer scientists must not lose sight of the fact that the law and legal reasoning do not easily – and sometimes should not – adapt to computational data and modelling. Herein lies the crux of this article. The potential for predictive analytics in the field of refugee status determination raises pertinent questions about model appropriateness and how data can be collected and used in a way that may subvert legal standards to the detriment of the very real human at its heart.

[1] Lehr, "Playing with the Data," 653.

The increased use of technology for border control and asylum management has been a cause for concern for many academics, humanitarian actors, practitioners, international organisations and people with lived experience of forced displacement. From ‘robo-dogs’ at the US–Mexico border,[2] militarised surveillance drones across the Mediterranean[3] and biometrics for refugee registration,[4] the implications of often-experimental technologies[5] for human rights, privacy and international protection for vulnerable populations are troubling.[6] Central to this conversation is the role that automated decision-making (ADM) will play in refugee status determination (RSD). RSD is the decision-making process undertaken by states and, in some cases, the United Nations High Commissioner for Refugees (UNHCR) to declare whether an asylum seeker is a refugee according to the Refugee Convention.[7] As a form of administrative decision-making, RSD is subject to domestic legislation and jurisdiction-specific merits and judicial review processes. Its largely individualised case-based nature inevitably creates significant workloads, making RSD a fertile ground for integrating ADM to achieve greater administrative efficiency. A defining characteristic of ADM, which is where a computer system automates part or all of an administrative decision-making process, is its ability to execute tasks or make decisions based on predefined logical criteria without requiring human intervention.[8] ADM that incorporates predictive analytics, which uses historical data, statistical algorithms, and machine learning techniques to forecast future events, promises even greater benefits in increased consistency and better decision-making. However, the nature of the subjects of RSD – people who may be facing forced displacement – creates a pressing need to consider how the charge to technologically enhanced decision-making can impact the lives and prospects of vulnerable populations.

In section 2, the concept of predictive analytics in law is introduced. This section provides an overview of ‘computational legal reasoning’ and commonly used AI tools for law prediction. In section 3, this discussion is extended to the application of predictive analytics in RSD. Section 4 is a call for pause and reflection. The application of what are essentially inductive reasoning models to an abductive reasoning decision-making process needs to be unpacked. Further, the data used to feed these models must be questioned if established legal standards are to be maintained to avoid bias, increased standards of proof and harmful feedback loops.

2. Predicting Law: Machine Learning in the Court Room

The drive to harness data to identify trends and extract insights into behaviours and events is commonly facilitated by predictive analytics. Also referred to as ‘legal analytics’,[9] the purpose of predictive analytics in law is manifold. Most pragmatically, the ability to predict the outcome of legal cases via computational methods promises to drive more targeted case preparation and better outcomes. Predictive analytics does not just have the potential to support applicants in making better-informed decisions by predicting the likelihood of success and the strengths or weaknesses of their application,[10] but it may also assist lawyers in devising better case strategies and managing client expectations by understanding trends and precedents. With the ‘potential to increase the scope of legal services that can be automated’,[11] predictive analytics could make time-consuming manual tasks requiring legal expertise obsolete.[12] An example of a simple legal predictive analytics tool is LexisNexis’s *Lex Machina*, which uses machine learning to allow users to predict litigation outcomes by filtering cases by type (e.g. practice area), date range, courts, judges, case resolutions, damages, findings, parties and counsel.[13]

Conversations about the potential utility of predictive analytics in case preparation tend to overlook where its impact looms the largest – as a tool for government agencies and judicial systems to manage cases and assist in producing cohesive and consistent decision-making.[14] However, while automated decision-making of varying complexity is increasingly common in government decision-making,[15] the use of algorithmic prediction of case outcomes by courts remains largely in the theoretical realm. The

[2] Fussell, “Dystopian Robot Dogs.”
[3] Mazzeo, “Border Surveillance.”
[4] Ewert, “Displaced, Profiled, Protected?”
[5] Bergtora Sandvik, “Do No Harm.”
[6] Office of the High Commissioner for Human Rights, Digital Border Governance.
[7] Refugee Convention.
[8] Commonwealth Ombudsman, Automated Decision-making Better Practice Guide.
[9] Ashley, Artificial Intelligence and Legal Analytics.
[10] Branting, “Scalable and Explainable Legal Prediction.”
[11] Morris, “User-friendly Open-source Case-based Legal Reasoning,” 270–271.
[12] Undavia, “A Comparative Study of Classifying Legal Documents with Neural Networks,” 515–522.
[13] Surdeanu, “Risk Analysis for intellectual property litigation.” 116-120. Lex Machina.
[14] European Commission for the Efficiency of Justice, European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.
[15] NSW Parliament, Automated Decision-Making in NSW.

widely reported use of an 'AI judge' to adjudicate small claims in Estonia[16] and its subsequent denial by the Estonian Ministry of Justice[17] speaks to how fraught the possibility of human judges being replaced or supported by AI is.[18] Despite misgivings, research continues to enthusiastically develop means to apply AI for accurate case prediction. Before moving forward, it is important to explore, at least conceptually, how such technology works.

Predictive analytics in law has developed around several 'adapted legal reasoning' models, which McGregor Richmond identifies as rule-based, argument-based, case-based, inferential and hybrid legal reasoning.[19] It is important to pause here momentarily to acknowledge that the term 'legal reasoning' applied to AI models may be a misnomer. Whether AI systems, and therefore predictive analytics, actually reason in the legal sense should be interrogated. This point is taken up later as part of the discussion about reasoning models in RSD. For now, references will be made to the commonly used and accepted terminology related to 'reasoning'.

Rule-based legal reasoning applies a rule of law to a problem using a series of 'if–then' rule statements to deduce a solution. Rule-based systems, often referred to as 'expert systems', comprise three components: a set of rules (rule base), a fact base (knowledge base) and an interpreter for the rules (inference engine).[20] An example is the development of a rule-based reasoning program for the criminal domain in Lebanon.[21]

Argument-based legal reasoning involves the construction and choice of arguments and counter-arguments. For example, Colonette uses argument-based reasoning to deduce outcomes in the European Court of Human Rights.[22] Choosing cases based on 'violation or non-violation' of selected articles of the European Convention on Human Rights (ECHR),[23] the authors create a 'factor hierarchy' to represent the issues in the case. They then apply a decision tree. Decision trees use control statements to categorise data based on specific conditions (i.e. a 'classifier'). A decision tree begins at a single node, then splits into multiple branches. Each branch represents different possible outcomes, incorporating various decisions and chance events until a final result is reached.[24] The following passage explains how the tree is structured in this study:

> The root of the tree is the verdict, the overall question to be decided, such as whether there is a violation. The next layers represent issues, the various broad ways in which an answer to the question must be considered, such as the various ways in which an Article may be violated. Often, the issues can be found in the legislation. Below these are the abstract factors, the various considerations found relevant to the issues in previous cases. Below these are the base level factors, which are the legal facts as accepted by the court. These in turn unfold into the plain facts of the case, so that the base level factors can be resolved by posing questions to the user.[25]

Legal case-based reasoning (L-CBR)[26] predicts the outcome of future cases by retrieving past cases to draw conclusions about a legal issue or problem.[27] L-CBR relies on computational methods to approximate analogical reasoning. Analogical reasoning, which is central to legal argumentation and outcomes, is where one case's facts, policy or reasoning are compared or contrasted with the facts, policy and reasoning of another case (or cases).[28] This is the commonly known *stare decisis* model. In a disputed case, the decision-maker will use the reasoning of an analogous case, distinguished from other possible sources, to support a conclusion about the disputed case.[29] L-CBR requires a rich database of past cases with well-documented features and outcomes to represent knowledge about case facts to assess legally relevant similarities. [30] L-CBR stores instances of previous cases and uses them for general reasoning, continually learning by retaining new cases.

---

[16] Lo, "Can AI Replace a Judge in the Courtroom?"
[17] Republic of Estonia Ministry of Justice, "Estonia Does Not Develop AI Judge."
[18] See, for example, Sourdin, "Judge v Robot?" 1114–1133.
[19] McGregor Richmond, "Explainable AI and Law."
[20] McGregor Richmond, "Explainable AI and Law."
[21] El Ghosh, "Towards a Legal Rule-based System."
[22] Collenette, "An Explainable Approach," 21–32.
[23] European Convention on Human Rights.
[24] Hillier, "What is a Decision Tree?" For a more complex discussion, see De Ville, "Decision Trees."
[25] Collenette, "An Explainable Approach," 21–32.
[26] McGregor Richmond, "Explainable Legal Case-Based Reasoning (Xl-Cbr)."
[27] McGregor Richmond, "Explainable AI and Law."
[28] Meagher, "Law, Society and Civil Rights," 4.
[29] Guarini, "Resources for Research on Analogy," 84–197.
[30] Ashley, Artificial Intelligence and Legal Analytics, 77. Ashley discusses three examples of computational models of legal concepts and cases.

The process of L-CBR can be represented by four steps:

1. Retrieve: Find the most similar past cases to new problem.
2. Reuse: Apply the solutions from the past cases to the new problem.
3. Revise: Adapt and refine the solution if necessary.
4. Retain: Store the new case and solution for future reference.

An example of an early L-CBR model is CATO, which is an 'intelligent learning environment' that applied a simple algorithm to trade secrets law to find all cases that have satisfied a relevance criterion based on a given problem. If cases that met the relevance criterion were satisfied and all had the same outcome, then the winning side would be predicted. Otherwise, the model would abstain.[31]

Finally, inferential legal reasoning methods tend to combine narrative/scenario-based, probabilistic, and argumentative approaches to reason with the help of evidential data, such as witness statements and forensic expert reports.[32]

In practice, predictive analytics rely on various AI and computational methods, often used in combination, such as decision trees, machine learning (ML) and natural learning processes (NLP). ML, one of the most common methods, has been described as 'an automated process of discovering correlations (sometimes alternatively referred to as relationships or patterns) between variables in a dataset, often to make predictions or estimates of some outcome'.[33] *Supervised* ML uses data and statistical algorithms to learn from pre-labelled training or historical data to identify correlations, patterns or clusters so input data can be given as a probability to inform a decision. Alternatively, in *unsupervised* ML, the algorithm is trained on data with no labels or predefined outcomes. Its goal is to find hidden patterns or intrinsic structures in the input data. NLP focuses on the interaction between computers and humans through natural language. It encompasses various tasks, including text processing, sentiment analysis and translation.

In 2017, Katz conducted a study using ML to predict the outcome of cases in the Supreme Court of the United States (SCOTUS).[34] Using historical decisions from 1816 to 2015, a random forest method was applied to predict the voting behaviour of the court and its judges, classified as Reversed, Affirmed or 'Other'. For more complex problems, random forests combine multiple decision trees and aggregate their results to make a final prediction.[35] The model achieved 70.2 per cent accuracy at the case outcome level and 71.9 per cent at the judge level. A similar study used document categories predefined by human evaluators, organised into 15 legal categories and further divided into 279 subtopics, to train a model to classify legal opinions from SCOTUS.[36]

In a widely cited paper, Medvedeva[37] investigated how NLP, combined with supervised ML, could be used to analyse the texts of the court proceedings in the European Court of Human Rights for the purpose of prediction. NLP processes and analyses human language in a way that computers can understand. In the training phase, a tool called a Support Vector Machine (SVM) Linear Classifier[38] was provided with numerous cases' textual information and judgments to identify patterns associated with verdicts, defined as 'violation' or 'no violation' of the articles of the ECHR. Once patterns were identified, the tool determined which words or word sequences were most characteristic of a violation or a non-violation. Examples of word sequences included 'the killing of' and 'the public prosecution'. In the testing phase, the tool was provided with a case without judgment for which it provided the most likely judgment. For example, the Chechen Republic was an important feature in 'violation' cases, while Bosnia and Herzegovina had a higher weight for 'non-violation' cases. The tool had an average accuracy of 75% in predicting the violation of nine articles. However, the researchers also found that the results became less certain over time due to the lower amount of training data for more recent cases.[39]

[31] Aleven, "Using Background Knowledge in Case-based Legal Reasoning."
[32] McGregor Richmond, "Explainable AI and Law."
[33] Murphy "Machine Learning: A Probabilistic Perspective."
[34] Katz, "A General Approach for Predicting the Behavior of the Supreme Court of the United States."
[35] See Parmar, "A Review on Random Forest," 758–763.
[36] Undavia, "A comparative study of classifying legal documents with neural networks.", Morris, "User-friendly open-source case-based legal reasoning", Branting, "Scalable and explainable legal prediction."
[37] Medvedeva, "Using machine learning to predict decisions of the European Court of Human Rights."
[38] The SVM Linear classifier is a type of machine learning tool that separates data into different categories by drawing the best possible straight line (or hyperplane) between them.
[39] See also Aletras, "Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective." e93.

3. Predictive Analytics in Refugee Status Determination

Predictive analytics is limited and somewhat cautiously employed in the legal determination of refugee status and migration. Nonetheless, research into model development holds significant potential for scalable and practical application. This section examines the current deployment of predictive analytics in migration and asylum systems, as well as ongoing and emerging research in this domain.

A fertile ground for predictive analytics RSD may be, but it is in 'standard' migration decision-making where predictive analytics, largely in the form of unsupervised ML utilising rule-based legal reasoning, is primarily found. For example, Canada has utilised an automated decision-making system[40] called Advanced Data Analytics (ADA) since 2018. ADA is employed by Immigration, Refugees and Citizenship Canada (IRCC) officers to sort and process temporary visa applications submitted outside of Canada. The system, which is driven by the goals of efficiency in client service and addressing an increasing volume of visa applications, uses historical data to tier applications into 'routine', 'non-routine', or 'complex or non-complex' (Tier 1, Tier 2, and Tier 3). Tier 1 applications are deemed simple enough to be processed without further human interaction. An example of a Tier 1 application is where the applicant held a previous Canadian visa without incident. Applications triaged into Tiers 2 and 3 are sent to a human officer for eligibility and admissibility review. Only a human officer can refuse an application. The ADA, which is not currently used for asylum and humanitarian matters, relies on a simple identification of 'output variables' that trigger categorisation. Although these variables have not been disclosed, Tao suggests they would likely include 'the nationality of the applicants, previous travel history to Canada or elsewhere especially global south countries, marital status, number of dependents, the purpose of visit, previous refusal, age, or some other requirements of section 27 of the *Immigration and Refugee Citizenship Canada Act*.' [41]

Australia employs a similar system in which non-complex visa applications may be 'auto-granted' without human oversight. Once the system has assessed whether any further information is needed from the applicant, it will either proceed to auto grant (if the applicant satisfies all objective criteria and is not required to undertake a health examination), to delayed auto grant (if the only unresolved requirement after the application is made is the completion of health examinations) or refer the application to an officer for manual assessment. The last step may occur if there are outstanding requirements, such as where more information may be needed from the applicant.[42]

The integration of predictive analytics into RSD has been a cautious affair. RSD involves some of the world's most vulnerable people in refugees, and anxiety pervades about removing or supplementing human decision-makers in this fraught context.[43] Such anxiety has been implicitly recognised by a reluctance by countries to employ automatic decision-making in more complex cases such as asylum, citizenship and other humanitarian visas.[44] However, there is some evidence of the integration of simple decision-assistance technology into RSD, which may signal a willingness to embrace predictive analytics in the future.

The Netherlands Immigration and Naturalisation Service (IND) has implemented a system of unsupervised ML called 'case matcher'. The system is based on text analysis, or text mining, where the algorithm searches and filters data and then ranks cases and documents based on where a term is found in a text or whether multiple search terms are clustered within a text. Through filters, such as document type, opening date, nationality, and search terms, the system generates and scores similar documents and cases, enabling IND caseworkers to identify applications made on similar grounds.[45] 'Case matcher' does not predict the outcome of a case but groups or clusters subjects together based, roughly speaking, on how similar their input data values are.[46]

Despite the dearth of actively 'deployed' predictive analytics in RSD, research into its potential abounds. The purpose of this research and the chosen technological approach varies. Some projects have a values-based objective to design systems to help other researchers, applicants, and their advocates identify bias in RSD decision-making. For example, in a 2017 study, Chen

[40] Automated decision-making systems have been described as 'a fully or partially automated technical system, used by [an] organisation … in administrative decision-making, and that affects people. Automated Decision-Making in NSW.
[41] Tao, "A Closer Look at How IRCC's Officer and Model Rules Advanced Analytics Triage Works."
[42] Law Council of Australia, Positioning Australia as a leader in digital economy regulation.
[43] Molnar, "Robots and refugees," 134-151.
[44] See, e.g., for the Australian position – 'In the migration context, a computer program will only be making decisions on certain visa applications where the criteria for grant are simple and objective. There is no intention for complex decisions, requiring any assessment of discretionary criteria, to be made by computer programs. Those complex decisions will continue to be made by persons who are delegates of the Minister.' Explanatory Memorandum, *Migration Legislation Amendment (Electronic Transactions and Methods of Notification) Act 2001* (Cth) [8].
[45] The Algorithm Register of the Dutch government. Memon, "Artificial Intelligence (AI) in the asylum system."
[46] Lehr, "Playing with the data: what legal scholars should learn about machine learning,".

and Eagle used supervised ML to report biases in asylum systems that came from several factors unrelated to the legal merits of the applications.[47] The authors applied a random forest method to a dataset of over 500,000 decisions from asylum hearings in the US that were decided over a 32-year period. The random forest classified features of the cases as 'case information' (nationality, number of family members, date of hearing, whether the application was affirmative or defensive, refugee's reason for immigration); 'court and judge information and trend' (law school graduation year, gender, president whom they were appointed by, military, and experience in years); 'weather and news' (weather features, current events and media coverage) and 'missing data and dummy variables'. By weighing the classified features against case outcomes of denial or allowance, they achieved a prediction accuracy of 82%.

A related project by Dunn focused on the 'early prediction' of asylum cases. Early prediction was defined as the prediction of the case outcome when the case opened, based only on the identity of the judge and the applicant's nationality. The researchers found that 'highly predictable judges' tended to hold fewer hearing sessions before making their decision, 'which raised the possibility that early predictability is due to judges deciding based on snap or predetermined judgments rather than taking into account the specifics of each case.'[48]

Other biased-focused projects in refugee-related matters have had troubled paths. The SupraLegem project in France applied ML to migrant expulsion decisions to predict bias amongst certain judges. The project found, with a high level of accuracy, that some judges appeared to be more likely to reject appeals than others despite the caseload being distributed randomly between judges.[49] The project subsequently prompted a ban on using ML in France to make such predictions.[50] A project by academics at the University of Gothenburg considered whether an 'ml-driven *post hoc* intervention system' could reduce the overall risk of discrimination emerging from human discretion in RSD.[51] Focusing on gender discrimination, the idea was that a random forest would alert a decision-maker to the fact that a case shared features with earlier cases identified as prone to discrimination risks.

Some research that seeks to build predictive analytics into RSD is based upon the 'legal merits' of the case rather than features external to the legal process that can indicate bias. These projects tend to be focused on improving case management and decision-making. An early example is a 1997 Australian study that devised an L-CBR model to retrieve important components of refugee cases based on case structure for the purpose of formulating new decisions and arguments.[52] More recently, the University of Copenhagen's Nordic Asylum Law and Data Lab[53] has produced several pieces of research that explore the intersection of data with asylum law. One project posed the question: Are there asylum applications with such characteristics that make the Refugee Appeals Board more likely to overturn the initial decision taken at the first instance? The project considered a large data set of RSD cases that were initially rejected and re-tried by the Refugee Appeals Board in Denmark. Features of the cases were extracted based on 'year of decision, country of origin of applicant and type of asylum claim' and fed to several random forest classifiers in all possible combinations as predictors.[54]

Other projects have applied NLP to aid in the prediction of RSD outcomes. A project associated with the Nordic Data Lab used a deep neural network[55] tool on approximately 15,515 cases to determine whether 'the information provided in a particular case is credible enough to grant the refugee status to the applicant, based on the information provided during the interview, or if there is a violation in relation to specific Articles to reject the refugee status'.[56] Another project called AsyLex[57] applied NLP to a dataset of 59,112 Canadian decisions from 1996 to 2022 to predict the outcome of decisions using document text with a mix of rule-based and human (i.e. legal experts) annotations. The project aimed to assist practitioners with 'case review', defined as 'entity extraction' and outcome prediction tasks.

Predictive analytics for RSD should be approached with a balance of caution and optimism. There is little doubt that it could be a useful tool to help applicants and their representatives better prepare their cases. It may become an important accountability

[47] Chen, "Can machine learning help predict the outcome of asylum adjudications?"
[48] Dunn, "Early predictability of asylum court decisions."
[49] Benesty, "NLP Applied to French Legal Decisions."
[50] Benesty, "The Judge Statistical Data Ban – My Story."
[51] Arvidsson, "Decision Making in Asylum Law and Machine Learning."
[52] Yearwood, "Case-based Retrieval of Refugee Review Tribunal Text Cases."
[53] Nordic Asylum Law and Data Lab.
[54] Katsikouli, "Machine Learning and Asylum Adjudications." See also Piccolo, "On Predicting and Explaining Asylum Adjudication."
[55] A deep neural network (DNN) is a type of artificial neural network (ANN) that has multiple layers between its input and output layers. These networks are capable of learning complex representations from data through a process called deep learning.
[56] Muddamsetty, "Danish Asylum Adjudication."
[57] Barale, "AsyLex."

mechanism that scrutinises bias in decision-making. Predictive analytics also holds significant potential for helping decision-makers streamline workload backlogs in a way that will actively contribute to faster RSD outcomes. The latter possibility – well, probability – prompts a deeper dive into how model design and data selection could have far-reaching consequences for legal reasoning, standards, case outcomes and, crucially, the humans in the RSD system.

4. Data and Model Selection: Implications for Legal Standards in RSD

Among the enthusiasm for the 'AI Spring' of the last decade are many voices of caution. The complexities of algorithmic bias have provided particularly fertile ground for rumination. Scholars have not only identified where and how bias occurs in the life-cycle of 'AI systems';[58] they have suggested and designed a plethora of mitigation techniques.[59] The conversation continues to be multidisciplinary. As just two examples among many, algorithmic bias has been considered through the lens of critical race theory[60] and as the manifestation of socially entrenched gender bias.[61]

In unpacking automated systems, legal scholars have paused to consider the ways AI can impact legal guarantees and frameworks. As early as 2014, Keats Citron and Pascale expressed concerns that individuals are increasingly being ranked and rated by algorithms in ways that undermine due process.[62] The following year, Kroll pointed out that oversight, 'in which real decisions are reviewed for their correctness, fairness, or faithfulness to a rule', rarely happens – if at all.[63] Since then, legal scholars have been avidly pursuing the ethical dimensions of the intersection of technology and law across the legal spectrum.[64]

The way data are collected, and what data are collected, have been fruitful topics for legal scholars pondering the ramifications of the rise of algorithms. An early klaxon for the harms of algorithmic bias from data collection and its impact on anti-discrimination law, Barocas and Selbst warned that 'unthinking reliance on data mining' can ignore data issues, such as algorithms inheriting the prejudices of prior decision-makers, data reflecting the widespread biases that persist in society and the misidentification of 'useful regularities' that are just pre-existing patterns of exclusion and inequality.[65] Lehr and Ohm's article builds on the new focus by lawyers on 'data problems' to point out that the harms of machine learning are not restricted to data collection or variable specification, but occur at every step along the way to system deployment – a point often missed by legal scholars. Indeed, the authors argue that two of the most notable harms of automated systems – opacity and lack of explainability – largely occur after algorithms are chosen and developed.[66]

While predictive analytics holds the potential for greater transparency of judicial proceedings,[67] integrating digital prediction tools into the courtroom raises its own concerns. Predictive analytics tools may create a greater opportunity to critique the justice system and for judges to use information produced by 'judicial analytics tools to reflect on and improve upon their practices'.[68] However, they also carry the potential to undermine the requirements of the right to a fair trial, such as adversarial proceedings, equality of arms and the impartiality and independence of a judge.[69] There is also a risk of the increased surveillance of judges, leading to strategic behaviour that could impact their health and well-being.[70] The remainder of this article borrows the critical lens from this ongoing conversation to examine how choices about models and data in predictive analytics for RSD may have consequences for viability, legal standards and the standard of proof.

4.1 Predictive Analytics and Legal Reasoning in RSD

A principal function and sometimes overlooked challenge of predictive analytics is to model legal reasoning to predict the outcome of future cases. Depending on the model, predictive analytics can reflect deductive, inductive or analogical reasoning.

---

[58] See, for example, Mehrabi, "A Survey on Bias and Fairness in Machine Learning"; Suresh, "A Framework for Understanding Sources of Harm."
[59] See Kusner, "The Long Road to Fairer Algorithms."
[60] Hanna, "Towards a Critical Race Methodology."
[61] Hall, "A Systematic Review of Socio-technical Gender Bias in AI Algorithms."
[62] Citron, "The Scored Society," 1.
[63] Kroll, "Accountable Algorithms."
[64] See, for example, McGregor, "International Human Rights Law as a Framework"; Harlow, "Proceduralism and Automation"; Završnik, "Algorithmic Justice".
[65] Barocas, "Big Data's Disparate Impact," 671.
[66] Lehr, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning."
[67] Stewart, "Judicial Analytics and Australian Courts," 84.
[68] McGill, "Judging by Numbers," 263–264.
[69] Steponenaite, "Judicial Analytics on Trial," 765.
[70] McGill, "Judging by Numbers," 273.

In deductive reasoning, a general theory is applied to facts to form specific conclusions. In law, deductive reasoning proceeds from a general rule, typically a statute, to determine the appropriate outcome in a specific case.[71] In predictive analytics, deductive reasoning underpins rule-based legal reasoning models. Conversely, inductive reasoning observes patterns in established facts through chains of inference to propose a general rule. In law, general rules are extrapolated from different cases where specific facts vary.[72] In predictive analytics, argument-based legal reasoning models are often based on chains of inference, reflecting inductive reasoning. One argument-based legal reasoning model is L-CBR, which relies upon analogical reasoning. Analogical reasoning assumes that similar problems have similar solutions. Analogical reasoning, which is a form or 'sub-set' of inductive reasoning, implicitly relies on recognising patterns across cases, comparing a current case to a previous case, and arguing that because the cases are similar, the same legal principles or outcomes should apply.

Evans Cameron describes the three basic premises of inductive reasoning as follows:[73]

1. A conclusion must be probably true.
2. A conclusion can only follow from premises that have themselves been established as a fact from evidence.
3. The decision-maker must build towards an ultimate conclusion by linking preliminary findings together in chains of inference.

Somewhere among the zeal for machines to predict law, a central and somewhat inconvenient truth seems to have been forgotten: artificial intelligence does not *reason* in the legal sense. Instead, it creates correlations. In 2012, James Franklin questioned the capacity of 'common sense' and legal reasoning to be 'formalisable'. Among his concerns was the application of open-textured or 'fuzzy' language to legal reasoning, which he exemplified through the concept of 'tall'. What is considered 'tall' differs in numerical degree according to context, such as for a woman, man or child. In law, rules involving open-textured language can be applied in context because there is typically a reason or intuitive *ratio decidendi* that is mentally present in the decision-maker to support the application of a rule.[74] That 'reason', often based upon factors such as memory, experience, policy and discretion existing beyond the immediate material, will not be present in the algorithm.

In 2018, the European Commission for the Efficiency of Justice (CEPEJ) rejected the notion that modern AI reproduces or models legal reasoning. Instead, AI systems 'infer a probable estimate of the best match between groups of lexical structures and translations already done'.[75] Referring to the work of Xavier Linant de Bellefonds, it was noted that 'the complexity of the law lies in its teleological and contentious nature: two coherent arguments can lead to different judgments according to two different priorities'.[76] The evolution of more advanced systems, such as generative AI, does not change this fact.

In RSD, the façade of AI 'legal reasoning' is arguably at its thinnest. RSD is not based on inductive reasoning; instead, it is a form of the lesser-known *abductive* reasoning, which relies upon an inference to the best explanation,[77] meaning that it starts with an observation or set of observations and seeks to find the simplest and most likely explanation. It focuses on finding the best explanation for the available evidence by combining new information with background knowledge. A simple way to explain abductive reasoning is through the following three-step process:

> Observation -> Generation of hypotheses->Selection of the best hypothesis considering context and background knowledge

Abductive reasoning is supposition based in that it requires an outcome to be 'plausible enough' based on 'an intelligent guess'[78] without a requirement that the premises for that outcome be established as fact. 'Intelligent guess' combines the perceiver's imagination, association and intuition.[79] Abductive reasoning asks the decision-maker to consider competing hypotheses, allowing her 'to accept a hypothesis if she feels sufficiently confident that it is the best explanation'.[80] The part played by

[71] Meagher, "Law, Society and Civil Rights."
[72] Meagher, "Law, Society and Civil Rights."
[73] Cameron, Refugee Law's Fact-Finding Crisis.
[74] Franklin, "How Much of Commonsense and Legal Reasoning is Formalizable?" 227–228.
[75] CEPEJ, European Ethical Charter on the use of Artificial Intelligence, 33.
[76] Linant de Bellefonds, "L'utilisation d'un système expert en droit comparé."
[77] Harman, "The Inference to the Best Explanation."
[78] Walton, Abductive Reasoning.
[79] Askeland, "The Potential of Abductive Legal Reasoning."
[80] Cameron, Refugee Law's Fact-Finding Crisis, 204.

abductive reasoning in legal argumentation and the law of evidence is increasingly acknowledged,[81] and inference to the best explanation is referred to as 'the best theory of the case'.[82]

In her important book on fact-finding in RSD, Evans Cameron argues that RSD is an abductive reasoning process.[83] Refugee adjudication has been described as conditioned by 'radical uncertainty',[84] the decision-maker more akin to a 'probability estimator' than a fact-finder, knowing 'that their state of knowledge can only ever be imperfect and who weighs various possibilities and decides to give or withhold the benefit of the doubt'.[85] This 'radical uncertainty' comes from several factors, including the need for the decision-maker to appreciate the complexity of socio-political conditions in a country of origin, the characteristic lack of documentary and other evidence and the need to test the asylum seeker's credibility, applying notoriously imprecise indicators such as 'plausibility' 'coherence' and 'consistency'.[86]

In inductive reasoning, while a conclusion may be 'probably true', it must follow from facts or evidence that have been proven to be fact. In most asylum cases, personal narratives of circumstances leading to a well-founded fear cannot be proven to be fact. This is complicated by the fact that the 'counter theory', the adjudicator's own theory of a case, is obscured. Evans Cameron points out that in traditional adjudicative settings, a clear theory and counter-theory are presented by two parties to an impartial adjudicator. In RSD, there is only one side and one clear theory of the case: the one presented by the claimant. However, a counter-theory still exists – that of the decision-maker, which is commonly premised on the theory that the claimant is actually an economic migrant and lying. This counter-theory 'lurks in the shadows, unexamined', and although it may be implausible, it is not scrutinised because it is not presented by the parties.[87]

RSD's fact-finding process requires a holistic approach to evidence that embraces uncertainty.[88] Evans Cameron points to the UK case of *Karanakaran*,[89] which has been adopted in UK and Australian jurisprudence; it rejected the traditional inductive legal reasoning approach for RSD in favour of a risk-based approach. The nature of risk means the probability that something has occurred cannot be proven as a fact.[90] Facts are not 'eliminated on the way to the final conclusion';[91] nor are they 'taken off the table'[92] as occurs in traditional chains of inference unless the decision-maker has no 'real doubt' that they are false. Instead, lingering doubts are carried forward and considered together in the final conclusion. This holistic approach to evidence means the decision-maker considers evidence 'in the round'[93] and, if doubts remain at the conclusion, they are to be resolved in the claimant's favour.[94]

However, Evan Cameron argues that the *Karankaran* approach does not go far enough in framing reasoning in RSD. The counter-theory (that of the decision-maker) remains obscured, meaning the claimant does not see competing hypotheses essential for identifying the 'best explanation'. She suggests that John R. Josephson's 'leading hypothesis' work on evaluating competing explanations in abductive reasoning[95] can be applied to the 'gap' between considering the evidence and applying the correct standard of proof in RSD. Josephson suggests that abductive reasoning involves five steps:

1. how decisively the leading hypothesis surpasses the alternatives
2. how well the hypothesis stands by itself, independently of the alternatives
3. how thorough the search was for alternative explanations
4. how strong the need is to come to a conclusion at all, especially considering the possibility of gathering further evidence, and
5. the cost of being wrong and the rewards of being right.[96]

[81] Cameron, Refugee Law's Fact-Finding Crisis, 195–196.
[82] Cameron, Refugee Law's Fact-Finding Crisis.
[83] Cameron, Refugee Law's Fact-Finding Crisis, 195–211.
[84] Macklin, "Coming Between Law and the State," 51.
[85] Macklin, "Coming Between Law and the State," 25.
[86] See Luker, "Decision Making Conditioned by Radical Uncertainty."
[87] Cameron, Refugee Law's Fact-Finding Crisis, 199–202.
[88] Cameron, Refugee Law's Fact-Finding Crisis, 188.
[89] *Karanakaran v Secretary of State for the Home Department* [2000] EWCA Civ 11, [2000] 3 ER 449.
[90] Cameron, Refugee Law's Fact-Finding Crisis, 187.
[91] *Minister for Immigration and Ethnic Affairs v Wu Shan Liang* [1996] HCA. at para 26, Sheppard, Lee, Carr JJ.
[92] *Karanakaran v Secretary of State for the Home Department* [2000] EWCA Civ 11, [2000] 3 ER 449 at 22, per Brooke LJ.
[93] *Ravichandran.*
[94] UNHCR, Benefit of the doubt.
[95] Josephson, "On the Proof Dynamics."
[96] Josephson, "On the Proof Dynamics,"1626–1627.

According to Evans Cameron, Josephson's first three steps reflect the *Karanakaran* approach, while steps 4 and 5 are 'pragmatic considerations' that should be integrated into RSD to ensure the benefit of the doubt is resolved in the claimant's favour. Step 1 reflects the role of credibility assessments in identifying the leading hypothesis. Step 2 creates a minimum threshold for accepting a hypothesis as the best explanation as long as it is not impossible or implausible. Step 3 requires the decision-maker to consider all relevant possibilities, recognising that sometimes an explanation is not needed or the correct explanation is unprecedented. Step 4 requires the decision-maker to consider whether they should try to gather further evidence, being mindful of the need to share the claimant's burden by helping frame the case and playing an active role in evidence gathering and identifying what is an 'important matter'. In step 5, error costs are weighed and considered when accepting an abductive conclusion. A legal requirement to afford the claimant the benefit of the doubt should guide this consideration. Therefore, unless an alternative hypothesis is decisively stronger, the leading hypothesis should be the one put forward by the claimant.

The duality of complexity and ambiguity in abductive reasoning in the RSD context challenges the notion that predictive analytics can mimic legal reasoning to predict case outcomes. Evolving predictive analytics for RSD has tried and tested various tools and reasoning approaches on the journey to model optimisation. Katsikouli applies decision trees, random forests, a support vector machine, logistic regression, neural networks, and naive bayes.[97] Chen and Eagel, and Dunn, apply a random forest model.[98] Barale proposes a methodology that combines supervised ML with a human component.[99] Arvidsson and Noll attempt to create an ML-based decision support system that builds on the work of Chen.[100] Piccolo relies upon decision trees, linear SVM, logistic regression, naive bayes and random forests.[101] Predictive analytics in other legal fields rely on decision trees[102] and support vector machines.[103] Despite differing approaches, the models generally utilise supervised ML, which is primarily an inductive reasoning process. In supervised ML, the algorithm learns from specific training instances to create generalised rules that can be applied to future data to make predictions.

From the perspective of this law academic, carrying with her all the limitations to technological knowledge that that title implies, supervised ML seems ill-fitted for abductive reasoning in RSD. Abductive reasoning requires 'commonsense reasoning over a broad range of activities and problems', which causes difficulties for 'developers who are trying to keep the problem domain as narrow as possible and as free from the mysterious "commonsense reasoning" and "world knowledge" as possible'.[104] 'World knowledge' in relation to forced displacement is multifarious, constantly evolving and shaped by perspective – whether the claimant's, the state's or from other 'objective' sources such as country of origin information. The CEPEJ report makes a similar observation when analysing the work of Aletras, referenced earlier,[105] which attempted to predict the outcomes of the European Court of Human Rights. The report argues that the ability of the AI to establish a 'high probability of correspondence between groups of words and a decision that had already been formalised' did not reproduce the reasoning of the judges, nor predict an outcome, on the basis, for example, of a future applicant's raw account before the Strasbourg court'.[106]

It is not, of course, that researchers cannot see the challenges that reasoning, evidence and credibility in RSD pose to predictive analytics. McGregor Richmond describes credibility assessments as a situation where 'legal rules and the evidential matrix are less explicitly rendered'.[107] McGregor Richmond suggests that RSD is a 'mode of adjudication less consistent with rationalist approaches than with abductive heuristics'.[108] In trying and testing two 'rationalist' approaches based on logical reasoning (Wigmorean and Bayesian), McGregor Richmond proposes an L-CBR methodology for RSD. Her model adapts 'episodic features and solutions of previous cases' to 'derive a robust co-variance "heatmap" of material adjudicative factors' that will demonstrate the comparative value of their interrelations. The aggregated sets of covariant maps will allow for multivariate outlier detection to determine the presence and nature of bias.[109]

[97] Katsikouli, "Machine Learning and Asylum Adjudications."
[98] Chen, "Can machine learning help predict the outcome of asylum adjudications?" Dunn, "Early predictability of asylum court decisions."
[99] Barale, "AsyLex: A Dataset for Legal Language Processing of Refugee Claims."
[100] Arvidsson, "Decision Making in Asylum Law and Machine Learning," 72–73.
[101] Piccolo, "On Predicting and Explaining Asylum Adjudication," 217–226.
[102] Collenette, "An Explainable Approach to Deducing Outcomes"; Katz, "A General Approach for Predicting the Behavior."
[103] Medvedeva, "Using Machine Learning to Predict Decisions."
[104] Waterman, "Expert Systems for Legal Decision Making."
[105] Aletras, "Predicting Judicial Decisions."
[106] CEPEJ, European Ethical Charter, 39.
[107] McGregor Richmond, "Explainable AI and Law," 18.
[108] Heuristic reasoning involves efficient cognitive processes, both conscious and unconscious, that ignore part of the information. Gigerenzer, "Heuristic Decision Making."
[109] McGregor Richmond, "Explainable Legal Case-Based Reasoning (Xl-Cbr)."

McGregor Richmond's model, which measures the interrelation of 'material adjudicative factors', could address the issue other predictive analytics models have in inductively 'discarding data' through inferential reasoning on the way to a conclusion. Using the heatmap to see how the evidence 'fits together'[110] may reveal uncertainty and doubts about the evidence and, as Evans Cameron, Goldfarb and Morris suggest, guide on intermediate material adjudicative factors, leaving the final decision to the human.[111] This is a promising step towards addressing how predictive analytics can tackle the complex, abductive nature of fact-finding and reasoning in RSD. However, as discussed in the remainder of this article, many obstacles remain regarding data characteristics and choices.

4.2 Prospective, Subjective and Simply Not Enough: Data in RSD

For an asylum seeker to be declared a refugee, they must prove they have a well-founded fear of being persecuted on at least one of the Refugee Convention grounds.[112] Credibility assessments, which involve consideration of the sufficiency of detail and specificity, consistency and plausibility of the claimant's narrative and evidence,[113] are central to this process. Credibility assessments are complicated by the fact that a 'well-founded' fear is a bipartite standard. A claimant must be able to demonstrate *subjective* fearfulness and an *objective* validation of that fear.[114] The bipartite standard has been confirmed by national courts[115] and the UNHCR, which has said:

> To the element of fear – a state of mind and a subjective condition – is added the qualification 'well-founded'. This implies that it is not only the frame of mind of the person concerned that determines his refugee status, but that this frame of mind must be supported by an objective situation. The term 'well-founded fear' therefore contains a subjective and an objective element, and in determining whether well-founded fear exists, both elements must be taken into consideration.[116]

Some scholars have suggested that there is no subjective element in the well-founded fear standard.[117] It imposes an additional burden because a claimant may be found not to be a refugee if they are not perceived to be 'subjectively fearful', irrespective of evidence of objective risk.[118] Despite these concerns, the bipartite standard remains, posing a unique 'data challenge' to predictive analytics in RSD.

Predictive analytics rely on historical data to train the ML algorithm to make predictions for the future. Training data in RSD can only be labelled according to evidence of past persecution based on the outcomes of other cases. While positive findings of persecution in similar situations may assist the decision-maker to identify whether the claimant has an objectively provable fear of being persecuted, past persecution will not provide insight into a claimant's subjective fear. Predictive analytics are thus not well designed to judge subjectivity. There may be a place for subjective assessment in AI, such as emotion-recognition technology, but this controversial tech is beyond the scope of this article. Predictive analytics will not separate subjective fear from its objective validation because case outcomes combine them into one monolithic 'fear'. Through the back door, other people's subjective fear could implicitly influence the outcome of a person with an entirely separate experience of being persecuted.

A related data challenge is the prospective nature of a well-founded fear. Fear is forward-looking, meaning a claimant need not show that they experienced persecution but have a well-founded fear of *being persecuted*. Historical evidentiary facts 'do not provide a sound basis for a determination that any asylum seeker is entitled to protection *now*'.[119] The fact that threats have not

---

[110] SM (Section 8: Judge's process) Iran.
[111] Cameron, "Artificial Intelligence for a Reduction of False denials," 493–510.
[112] Refugee Convention, Art 1A(2).
[113] UNHCR, Beyond Proof, 27. See also UK Home Office, Asylum Policy Instruction Assessing Credibility, para 5.4.
[114] Hathaway, The Law of Refugee Status, 91–92.
[115] *Re Minister for Immigration and Multicultural Affairs; Ex parte Miah* [2001] HCA 22, at 76 per Gaudron J; *Ward v Canada (Attorney General)* [1993]2 SCR 689 (Can. SC, Jun. 30, 1993), at 723; *Zgnat'ev v Minister for Justice, Equality and Law Reform* [2001] IEHC 70 (Ir. HC, Mar. 29, 2001), at para 6.
[116] UNHCR, UNHCR Handbook, para 38.
[117] See, for example, Hathaway, "Is There a Subjective Element?," 505. Bossin, "A Canadian Perspective," 108.
[118] *Gzim Bela v Canada* [2013] FC 784 (Can. FC, July 12, 2013). The Canadian Federal Court has expressed concerns about subjectivity in the well-founded fear standard: '[But] I find it hard to see in what circumstances it could be said that a person could be right in fearing persecution and still be rejected because it is said that fear does not actually exist in his conscience. The definition of a refugee is certainly not designed to exclude brave or simply stupid persons in favour of those who are more timid or more intelligent. Moreover, I am loath to believe that a refugee status claim could be dismissed solely on the ground that as the claimant was a young child or a person suffering from a mental disability, he or she was incapable of experiencing fear, the reasons for which clearly exist in objective terms.' *Yusuf v Canada* (1992) 1 FC 629 (Can. FCA, Oct. 24, 1991).
[119] *R. (Saber) v Secretary of State for the Home Department* [2007] UKHL 57 (UKHL, Dec. 12, 2007) at para 2 (emphasis added).

been carried out does not mean a claimant's fear is not well-founded. Instead, what is important is that the 'group making the threat has the will and ability to carry it out'.[120]

It should be remembered that RSD is a risk assessment. Evidence of past persecution acts as a 'guide' to what may happen if a person returns to their country of origin,[121] but an assessment of risk is speculative.[122] While past persecution may be a sound indicator that future persecution may occur, there may be a good reason why it will not be repeated.[123] Indeed, claimants may 'prevail on a theory of future persecution despite an … adverse credibility ruling as to past persecution so long as the factual predicate of [her] claim … is independent of the testimony … found not to be credible'.[124] Therefore, past persecution alone cannot determine well-founded fear. If other people's experiences exclusively form the basis of training data, past persecution is given a higher weight in the decision-making process than it should receive according to refugee law.

The challenges of capturing a claimant's subjective fear in predictive analytics, accounting for embedding previous claimants' subjective fear in the training data and adapting historical data to a prospective legal standard, risk elevating the standard of proof to the detriment of the refugee. Although standards of proof in refugee law vary across jurisdictions, including a 'real chance' in Australia,[125] a 'reasonable possibility' in the United States,[126] a 'reasonable degree of likelihood' in the United Kingdom[127] and a 'reasonable chance,' reframed as a 'serious possibility' in Canada,[128] they are low compared with civil (balance of probabilities) and criminal (beyond a reasonable doubt) standards. Considering that the refugee bears the burden of proof, these 'data challenges' introduce factors that claimants must implicitly address above and beyond making out their own case.

Exacerbating the issues of subjectivity and prospectivity is the fact that RSD will either have insufficient training data or be unreliable.[129] It is the nature of a fact-finding process such as RSD, which relies on personal narrative and unreliable or non-existent documentary evidence for facts to vary widely according to the case. Given 'the sparse data and uncertain environment in refugee claims', 'machine predictions are likely to generate wide distributions'.[130] Medvedeva made a similar observation in the context of predicting ECHR decisions. The researchers found that it was easier to predict the judgement of similar cases than where very diverse issues were grouped under a single article of the ECHR, making the prediction performance lower.[131] In addition, training on one period and predicting for another becomes difficult when the gap between the training and testing data increases. Further, because case law evolves over time, changes may not be picked up by a system trained on historical data.[132] Consequently, continuous integration of published judgments in the system is necessary to keep up with the changing legal world and maintain adequate performance.[133]

Connected to the issue of insufficient data is a 'class imbalance problem'. Class imbalance occurs when the distribution of classes in the training data is heavily skewed, meaning some classes have significantly more instances than others. As Muddamsetty points out, class imbalance can pose challenges during the data training stage, 'leading to biased predictions and reduced performance on minority classes'.[134] Imbalanced classes are likely in RSD because some case characteristics, such as particular nationalities, will be highly represented in the data, particularly in times of mass influx, such as the 2016 Syrian crisis. Consequently, more unusual situations, such as persecution related to a specific family group in a region without civil unrest, will result in fewer cases and fewer data. 'In such cases, standard classifiers tend to be overwhelmed by the large classes and ignore the small ones',[135] which could have consequences for meaningful outcome prediction.

[120] *Marcos v Attorney General* 2005 US App., Lexis 10698 (USCA9, Jun. 9, 2005).
[121] *S152/2003 v. MIMA*, [2004] HCA 18 (Aus. HC, Apr. 21, 2004), per McHugh J. See also *Katrinak v. SSHD*, [2001] EWCA Civ 832 (Eng. CA, May 10, 2001).
[122] Anderson, "A well-founded fear of being persecuted," 156.
[123] Directive 2011/95/EU**,** Art 4(4).
[124] *Tatsiana Boika v Attorney General* 727 F.3d 735 (USCA7, Aug. 16, 2013).
[125] *Chan v MIEA* (1989) 63 ALR 561 (Aus. HC).
[126] *INS v Cardoza-Fonseca* (1987) 467 U.S. 407 (USSC).
[127] *R. v. SSHD ex parte Sivakumaran* [1988] 1 All E.R. 193 (UKHL).
[128] *Adjei v MEI*, [1989] 57 DLR 4th 153 (Can. FC A.): SCC (1995).
[129] Cameron, "Artificial intelligence for a reduction of false denials," 12.
[130] Cameron, "Artificial Intelligence for a Reduction of False Denials," 494.
[131] Medvedeva, "Using Machine Learning to Predict Decisions."
[132] Collenette, "An Explainable Approach to Deducing Outcomes".
[133] Medvedeva, "Using Machine Learning to Predict Decisions."
[134] Muddamsetty, "Danish Asylum Adjudication Using Deep Neural Networks."
[135] Muddamsetty, "Danish Asylum Adjudication Using Deep Neural Networks."

4.3 The Tail Wagging the Dog? Case Characteristics

For 'non-tech scholars' (i.e. not computer or data scientists), unpacking technology in migration and refugee decision-making systems, border control and other migration-related areas – for example, biometric registration,[136] algorithms for resettlement,[137] satellite-based predictions of migratory movements[138] – has meant focusing on the manifestation and consequences of that tech over the data that feed it.[139] Consequently, the role played by data in these systems remains relatively under-explored.

In their fascinating autoethnographic article chronicling a research project that attempted, and failed, to create a machine that identified bias in refugee decision-making, Arvidsson and Noll encountered the problem of 'data wrangling'. In designing a 'hybrid human–machine collaborative system', the researchers found that decisions about the selection and management of data depended on the explanations by the Migration Agency – the very author of the decisions the system was supposed to be impartially scrutinising. In the context of this project, 'data wrangling', which broadly refers to processes that transform raw data into more readily used formats, means notoriously opaque refugee decisions are shifted elsewhere, compromising the ability of the researchers to identify bias in decision outcomes.[140] The study is a timely warning that the way data is 'wrangled' in RSD predictive analytics systems must be careful not to undermine the legal process of determining who is a refugee according to the Refugee Convention.

Predictive analytics driven by supervised ML rely on the selection of case characteristics. A case characteristic is a specific attribute, feature or variable that describes an individual case or data point within a dataset. Case characteristics make up the data set upon which the algorithm is trained. McGregor Richmond discusses 'case characteristics' in terms of 'classification', which she describes as 'key' and central to the development of a legal ontology that recognises that the 'application of legal rules is often contingent on the satisfaction of a particular definition'.[141] In some RSD ML systems, the case characteristics are associated directly with the merits of the case, or 'the legal factors that are most influential in determining the [case] outcomes'.[142] For example, Barale's study included case characteristics such as the constitution of the decision-making panel, the country of origin and characteristics of the claimant, the year the claim was made in relation to a particular geopolitical situation, the legal procedures involved, the grounds for the decision and relevant legislation.[143] Other models, such as Chen and Eagel's[144] and Dunn's,[145] measured legal characteristics against those not associated with the merits of the case. For example, Chen and Eagel combined nationality, number of family members, date of hearing, whether the application was affirmative or defensive and the refugee's reason for 'immigration' with the law school graduation year and gender of the judge, who was president when they were appointed and whether they served in the military.

Each study considered in this article outlines how 'legally relevant' case characteristics were selected. Chen and Eagel report that their case information was based on their 'intuition about the relevance of case-centric factors'.[146] Dunn models their 'data dictionary' on information from a previous study and conversations with practising migration attorneys.[147] Barale uses a mixture of NLP and annotations by legal attorneys to select appropriate characteristics.[148] Katskouli states that two members of their group, one a legal expert, independently studied and manually extracted features from 50 randomly sampled cases.[149] Piccolo relies on a literature review of similar previous work.[150] Arvidsson's choice of outcome-relevant variables depended on Chen and Eagel's choices.[151] Notably, none of the studies explicitly bases case characteristics upon the legal criteria for a refugee in the Refugee Convention nor relevant domestic legislation. Considering that judges apply these criteria to the facts in the cases, why might this be so? And, importantly, what might the consequences be for the law?

---

136 Jacobsen, "Experimentation in Humanitarian Locations."
137 Ahani, "Placement Optimization in Refugee Resettlement."
138 Ahmed, "Artificial Neural Network and Machine Learning Based Methods."
139 Kinchin, "Technology, Displaced?" See, however, Bringas Colmenarejo, "Artificial Intelligence and Big Data Analytics"; McCarroll, "Weapons of Mass Deportation."
140 Arvidsson, "Decision Making in Asylum Law and Machine Learning."
141 McGregor Richmond, "Explainable Legal Case-Based Reasoning (Xl-Cbr)."
142 Ashley, Artificial Intelligence and Legal Analytics, 123.
143 Barale, "AsyLex."
144 Chen, "Can Machine Learning Help Predict?"
145 Dunn, "Early Predictability of Asylum Court Decisions."
146 Chen, "Can Machine Learning Help Predict?" 237.
147 Dunn, "Early Predictability of Asylum Court Decisions," 234.
148 Barale, "AsyLex."
149 Katsikouli, "Machine Learning and Asylum Adjudications," 130958.
150 Piccolo, "On Predicting and Explaining Asylum Adjudication."
151 Arvidsson, "Decision Making in Asylum Law and Machine Learning," 77.

The obvious answer to the first question is that researchers seek common case characteristics that can be used to aid prediction within the same jurisdiction. An insight of sorts may be found in the observation by Piccolo that 'truly predictive models that try to anticipate future decisions cannot include all the legally relevant factors as some of them can only be established by the judges as part of their decisions (such as the applicant credibility)'.[152] However, this conflates the *law* with *applying the law to the facts*. When the predictive analytics model is used to alert parties and observers to individual judges' potential bias or decision-making patterns, this raises no issues. However, when predictive analytics become 'recommender systems' to guide future decisions – which they inevitably will – the case characteristics become proxies for the legal criteria for RSD. The tail wags the dog, if you will. The problem with the 'tail' is that because case characteristics can only be represented by simple concepts (e.g. religion, family size), they do not individually, nor in a 'correlation' with each other, represent the legal definition of a refugee. There is also a risk that case characteristics will be represented as 'legally relevant' when they are not. When they are carried forward as predictors for future decisions, the law risks becoming skewed.

Under international law, an asylum seeker must meet the criteria outlined in Article 1A(2) of the Refugee Convention to be declared a refugee. 'Exclusions' to protection under the Convention and the *non-refoulement* obligations of states must also be taken into consideration.

Table 1: Legal criteria for refugee status in international law

Legal criteria or exclusion for refugee status	Refugee Convention
Outside their country of citizenship or habitual residence (i.e. alienage).	Article 1A(2)
Well-founded fear of being persecuted	Article 1A(2)
For reasons of race, religion, nationality, membership of a particular social group or political opinion.	Article 1A(2)
State of origin failing to protect	Article 1A(2)
Non-refoulement exception: Reasonable grounds for regarding a person as a danger to the security of the country or who, having been convicted by a final judgement of a particularly serious crime, constitutes a danger to the community of that country.	Article 33(2)
Exclusion: 'He has committed a crime against peace, a war crime, or a crime against humanity, as defined in the international instruments drawn up to make provision in respect of such crimes'	Article 1(F)(a)
Exclusion: 'He has committed a serious non-political crime outside the country of refuge prior to his admission to that country as a refugee.'	Article 1(F)(b)
Exclusion: 'He has been guilty of acts contrary to the purposes and principles of the United Nations.'	Article 1(F)©

Take Katsikouli's project as an example. In the following passage, the authors reveal their chosen case characteristics or 'feature set':

> the applicant's country of origin (or nationality), the applicant's identified gender, the applicant's identified religion, the applicant's identified ethnicity, the year the applicant entered Denmark, their marital status, their involvement in political parties and organizations, military involvement or experience, whether the applicant has applied for asylum in another country before coming to Denmark, whether discrepancies were identified in the applicant's case, in cases of torture we check whether relevant investigation was carried out, the type(s) of asylum claim and finally, the Refugee Appeal's Board decision on the case.[153]

Some of these case characteristics are not legal criteria for a refugee in the Refugee Convention or Danish domestic law, which mirrors the Convention.[154] Two case characteristics stand out in this regard. The first is 'whether the applicant has applied for asylum in another country before coming to Denmark'. Failure of an asylum seeker to claim protection in their region of origin or in the first safe country she reaches is not grounds for refusal under the Refugee Convention.[155] While policy and third-party

[152] Piccolo, "On Predicting and Explaining asylum Adjudication."
[153] Katsikouli, "Machine Learning and Asylum Adjudications," 130958.
[154] Article 7(1) of the Danish *Aliens (Consolidation) Act* states: '7(1) Upon application, a residence permit will be issued to an alien if the alien falls within the provisions of the Convention Relating to the Status of Refugees (28 July 1951).'
[155] 'There is no requirement in the Convention that a refugee seek protection in the country nearest her home, or even in the first state to which she flees. Nor is it requisite that a claimant travel directly from her country of first asylum to the state in which she intends to seek

regional agreements may have much to say on this issue,[156] the Refugee Convention demands nothing of the asylum seeker in this regard. Second, 'the Refugee Appeal's Board decision on the case' is an indicator of a previous finding in a similar case that risks carrying forward legal error.

Other case characteristics reflect legal criteria for refugee status but are framed both broadly and simplistically, distorting their impact. The applicant's country of origin (or nationality), identified gender, religion, ethnicity, marital status, involvement in political parties and organisations, and military involvement or experience could all constitute grounds for persecution. However, because Article 1A(2) defines a refugee as a composite phrase, case characteristics cannot be considered in isolation. As the UNHCR has said, 'The Article 1 definition can, and for purposes of analysis should, be broken down into its constituent elements. Nevertheless, it comprises only one holistic test.'[157] To properly reflect the Refugee Convention, case characteristics would need to be combined with an assessment of 'well-founded fear' supported by credibility assessments. Further, 'being persecuted' involves assessing serious harm through applying a human rights approach that considers the appropriate exemptions in international law.[158] Alienage, state protection and protection exemption must also be taken into account. In predictive analytics, case characteristics are utilised in various, often random, combinations and inputted as predictors into classifiers. This approach not only fails to control for all the relevant variables – increasing, as Piccolo argues, the risk of facing the 'Simpson's paradox' or finding spurious correlations,[159] but it also risks distorting the law and introducing irrelevant variables into the model.

The final point to be made regarding the selection of case characteristics is a reminder that RSD is an abductive reasoning process. In abductive reasoning, the identification of issues does not happen in isolation. It is a process informed by background knowledge. The background knowledge is the law. As Aleven says in his important work on abductive reasoning in the law:

> the challenge in modelling arguments about the significance of differences between cases is to apply the background knowledge in a context-sensitive manner, so that the significance of a given distinguishing factor may vary depending on the specifics of the cases being compared and the purpose for which the argument is made.[160]

Background knowledge is used to construct arguments, identify issues that a problem raises, judge the relevance of available past cases 'and marshal the cases in a variety of argument moves in order to address a party's strengths and weaknesses with respect to each issue'.[161] Similarly, Askeland suggests that in law, abductive reasoning merges knowledge, facts and law rather than explaining an observed factual phenomenon by combining new information with background knowledge. He argues that abductive reasoning can be adapted to legal problem-solving by inferring which rule a given fact can be subordinated to within the realm of legal norms:[162] 'In order to adapt abduction to the field of law, one must integrate the desired effect of a given rule to the set of data to begin with.'[163] For RSD, the issues need to be framed in relation to the definition of a refugee in Article 1A(2) of the Refugee Convention, which is the background knowledge. From there, different hypotheses are formed and the best explanation of the case is selected.

4.4 Bias, Fairness Metrics and Feedback Loops

The discussion in this article is not intended as a criticism of any of the predictive analytics models discussed. Each has been designed to predict future asylum cases, and overall they do so well. The potential future use of these models as guides and templates for decision-making prompts the waving of red flags here, particularly because of their potential impact on refugees and asylum seekers.

One such impact is the risk of data bias. Data bias in RSD is highly probable if insufficient consideration is given to whether chosen variables represent previous decisions rather than legal standards in the Refugee Convention and case law. *Data-based or statistical bias* results from how data has been selected to exclude parts of the population it is supposed to represent. Alternatively, the model can perpetuate and reinforce the bias in the data. *Modelling or 'method' bias* arises when the model

durable protection.' *Gavryushenko v Canada (Minister of Citizenship and Immigration)*, (2000) 194 FTR 161 (Can. FCTD, Jul. 26, 2000) at 156 [10] per Lufty CJ.

[156] See, for example, Dublin Regulation.

[157] UNHCR, Interpreting Article 1 of the 1951 Convention Relating to the Status of Refugees.

[158] Hathaway, The Law of Refugee Status, 91–92.

[159] Piccolo, "On Predicting and Explaining Asylum Adjudication."

[160] Aleven, "Using Background Knowledge in Case-based Legal Reasoning," 184.

[161] Aleven, "Using Background Knowledge in Case-based Legal Reasoning," 184.

[162] Askeland, "The Potential of Abductive Legal Reasoning," 68.

[163] Askeland, "The Potential of Abductive Legal Reasoning," 71.

reproduces unfair patterns encoded in the data or reinforces or creates patterns of social advantage. Another form of bias is *societal bias*, which reflects 'longstanding structural inequalities' in society and arises when existing bias infiltrates the data-collection process. A deeper discussion of the various types of algorithmic and data bias,[164] and how they may impact refugees in the RSD process, is outside the scope of this article, but two examples may help demonstrate the potentially harmful impact on refugees.

Measurement bias occurs when data features and labels are used as 'imperfect proxies' for the 'real values of interest'.[165] For example, arrest rates are often used as proxies for crime rates, but because 'minority communities are more highly policed', arrest rate proxies are differentially mismeasured from the crime rate (i.e. the real value of interest).[166] McNamara and Tikka suggest that in the context of RSD, measurement bias may emerge in relation to the assessment of plausibility in credibility assessments, 'wherein a sovereign's "assumption or inference as to how people would behave in certain situations" discounts actual realities of lived fear and persecution'.[167] The authors give the example of an arrest record from an asylum seeker's country of origin. An arrest might be indicative of 'persecution' if it were based on 'the grounds in the Refugee Convention, but should it be considered a proxy for persecution or count against granting asylum because of the potential that the claimant is a criminal? The authors question 'whether it is rational to expect an ML model to contextually determine, based on an asylum seeker's narrative, whether an interviewee was a protestor, a rioter, or a looter, let alone what contextual weight such situational identity might have for a legal determination'.[168]

Fairness metrics, which are measures designed to address algorithmic fairness and bias, may be designed into predictive analytics to help address data bias caused by the way data is chosen and applied in RSD. However, as the following discussion demonstrates, they are not a panacea and may not adequately address the specific conditions and processes of RSD. 'Fairness through unawareness' is a method that works by removing any data considered *prima facie* to be unfair. For example, for an algorithm used by judges making parole decisions, fairness through unawareness could dictate that data on ethnic origin should be removed when training the algorithm. Data on the number of previous offences, however, could be used.[169] The difficulty with this approach is that ignoring protected attributes such as ethnicity also ignores the fact that most data tend to be biased and can bear the stamp of historical racial bias.[170] 'Colour blind' ideologies may seem to promote fairness in the short term but are ineffective in 'rooting out perceived bias'.[171] In RSD, removing protected attributes would be counterproductive as these characteristics (i.e. nationality, ethnicity, religion) are central to determining a well-founded fear of being persecuted on 'convention grounds'.

'Demographic parity' is where a positive prediction is assigned to two groups at a similar rate.[172] For example, a university admissions algorithm would satisfy demographic parity for gender if 50 per cent of its offers went to women and 50 per cent to men.[173] However, demographic parity does not effectively account for differences between groups. For example, it might not make sense to use demographic parity in specific settings, such as a fair arrest rate for violent crimes, because men are significantly more likely to commit acts of violence.[174] There is no protection against the possibility that such methods will exacerbate bias against specific gender-race pairings or other combinations of sensitive attributes,[175] trading one group's false positives for another's false negatives.[176] For RSD, signalling what makes a person *different* from members of the same group or nationality can be key to identifying a well-founded fear.

'Equality of opportunity' means that the same beneficial predictions are assigned to individuals in each group. Accordingly, protected and unprotected groups should have equal true positive rates.[177] Consider a predictive algorithm that grants loans only to individuals who have paid back previous loans. It satisfies 'disability-based equality of opportunity' if it grants loans to the same percentage of individuals who both pay back loans and have a disability as it does to those who pay back loans and

[164] For illuminating discussions on types of algorithmic bias, see Mehrabi, "A Survey on Bias and Fairness in Machine Learning."
[165] Van Giffen, "Overcoming the Pitfalls and Perils of Algorithms."
[166] Suresh, "A Framework for Understanding Sources of Harm."
[167] McNamara, "Well-founded Fear of Algorithms?"
[168] McNamara, "Well-founded Fear of Algorithms?" 262.
[169] Kusner, "The Long Road to Fairer Algorithms."
[170] Kusner, "The Long Road to Fairer Algorithms"; Morse, "Do the Ends Justify the Means?"
[171] Ely, "Cultural Diversity at Work."
[172] Pessach, "Algorithmic Fairness."
[173] Kusner, "The Long Road to Fairer Algorithms."
[174] Kusner, "The Long Road to Fairer Algorithms."
[175] Teodorescu, "Failures of Fairness in Automation Require a Deeper Understanding of Human–ML Augmentation."
[176] Morse, "Do the Ends Justify the Means?"
[177] Mehrabi, "A Survey on Bias and Fairness in Machine Learning."

who do not have a disability.[178] This approach allows for demographic differences but levels the playing field by requiring the unfair or erroneous judgements or false-positive rates to be equally distributed.[179] However, societal unfairness is not captured by equality of opportunity.[180] Being able to pay back a loan in the first place can be affected by bias: discriminatory employers might be less likely to hire a person with a disability, which can make it harder for that person to pay back a loan. Similarly, group-focused equity approaches require equality of opportunity and equalised odds. However, ensuring false positives are equitably distributed is like ‘plugging one leak may worsen another’.[181] Group fairness denies social, economic and political complexity as well as the unique intersectional discrimination for some groups, such as Black women.[182] As Hana argues, race is often understood as a fixed attribute rather than an institutional and relational phenomenon.[183] Race has many variables (e.g. self-classification, observed race, phenotype), but is perceived as a single-dimensional variable that can take on a handful of values.

A consequential risk of bias and insufficient fairness metrics is the potential for the creation of harmful feedback loops. A feedback loop is created when the output produced by the algorithm is fed back into the system as ‘input’ or training data. Previous outputs may contain bias and discrimination, which will be reinforced in the processing of future applications. A biased system used to make a calculation of probabilities will always produce a biased result, ‘which will be applied in the world and will create more inequalities, generating a feedback loop that is quite problematic’.[184] Feedback loops are particularly concerning for predictive analytics in RSD as they have the potential to amplify the discrimination experienced by asylum seekers and refugees.

For example, feedback loops may cause applications from a particular nationality to be rejected at a higher rate if previous applications from the same nationality were previously rejected at a higher rate than others, analogous to forming a culture of disbelief.[185] In 2020, the United Kingdom was forced to abandon a visa-processing program because of the way bias and discrimination were amplified via a feedback loop. In the ‘traffic light’ visa program, applicants from nationalities that were identified as ‘suspect nationalities’ received a higher risk score, and thus a higher level of scrutiny by officers. Advocates argued that the practice was discriminatory by design and that the algorithm suffered from a ‘feedback loop’ problem. Previous applications from the same nationality were previously rejected at a higher rate than others, creating discrimination and unfairness for future applicants on the grounds of nationality.[186]

6. Conclusion

I recently attended a ‘scenario lab’, which was delivered by the Centre for Digital Cultures and Societies at the University of Queensland. The scenario lab brought together academics from diverse disciplines to ruminate on and explore the future of technology, borders, asylum and migration. In one of the activities, which was led by science fiction writers, we were broken into groups to create stories that responded to future-focused socio-political scenarios. Before we began, we received some sage advice from one of the writers: ‘Avoid the temptation to be dystopian.”

This is more challenging than it seems, particularly in the context of forced displacement and border control. One only has to read Petra Molnar’s recent book *The Walls Have Eyes: Surviving Migration in the Age of Artificial Intelligence*[187] to understand that technology experimentation has very real consequences for our poorest and most vulnerable people. Increased AI regulation doesn’t always have the impact it should, and migration and asylum technology can evade scrutiny. For instance, the recent EU *Artificial Intelligence Act*’s[188] transparency requirements do not apply to all AI systems that could be considered high-risk. Although registration in a public database is required for high-risk systems, Article 49 excludes systems used in law enforcement and asylum and border control management. Instead, these systems are to be registered in a ‘secure non-public section of the EU database’.

[178] Kusner, “The Long Road to Fairer Algorithms.”
[179] Morse, “Do the Ends Justify the Means?”
[180] Kusner, “The Long Road to Fairer Algorithms.”
[181] Morse, “Do the Ends Justify the Means?”
[182] Morse, “Do the Ends Justify the Means?”
[183] Hanna, “Towards a Critical Race Methodology.”
[184] Mansoury, “Feedback Loop and Bias Amplification in Recommender Systems.”
[185] McNamara, “Well-founded Fear of Algorithms?” 238.
[186] BBC, “Home Office Drops ‘Racist’ Algorithm.”
[187] Molnar, *The Walls Have Eyes.*
[188] *EU AI Act*, Article 49.

At the same time, we must acknowledge that 'the horse has bolted'. AI is already entrenched in many facets of private and public life, and predictive analytics holds great promise for efficient, effective and consistent decision-making, and to identify bias and discrimination encoded in decision-making. This article does not seek to stem the AI tide; rather, it asks for a pause to consider whether predictive analytics can meaningfully translate to RSD.

Predictive analytics, traditionally built upon inductive decision-making processes like supervised ML and decision trees, risks compromising the abductive reasoning processes upon which RSD relies. Even if models are built to effectively navigate this hurdle, problems with data remain. Insufficient data is an intractable element of forced displacement, which means inaccuracies, or at the very least uncertainty, must also be. The prospective nature of a well-founded fear undermines the way algorithms are traditionally trained from historical data. And what of subjectivity? The inability of predictive analytics to measure subjective fear may mean that RSD credibility assessments are pushed to 'pseudo-scientific' tools such as lie detectors and emotion-recognition technology. Yet, findings of well-founded fear do not delineate between its subjective and objective elements. A well-founded fear is a well-founded fear. Without the ability to remove subjectivity, the training data will capture subjective fear from previous claimants that will go on to inform another completely different case. Case characteristics must be carefully chosen in close consultation with legal experts. Without considering how case characteristics reflect legal standards, they risk subverting hard-fought and won international legal standards. Finally, the more commonly acknowledged issues of algorithmic bias and harmful feedback loops must not be forgotten.

Hopefully, this discussion will encourage and facilitate further interdisciplinary dialogues about predictive analytics in RSD. The timing is nigh because such technology remains largely experimental and has yet to be integrated into state and UNHCR processes in a scalable way. Two considerations must be front and centre of any such conversation. First, there is a human in the feedback loop. Second, legal standards were designed for the ultimate goal of international protection for forced displacement. We must be careful not to be enraptured by shiny new things that 'tinker at the edges' of administrative and judicial decision-making. One need only look to RSD to see that the consequences may be more negatively transformative than they first seem.

Bibliography

Primary Legal Material

International

Convention Relating to the Status of Refugees ('Refugee Convention'), opened for signature 28 July 1951, 189 UNTS 150 (entered into force 22 April 1954), art 1A(2) read in conjunction with The Protocol relating to the status of refugees, opened for signature 31 January 1967, 606 UNTS 267 (entered into force 4 October 1967) ('The Protocol').

European Convention on Human Rights, formally known as the Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), opened for signature on 4 November 1950, 213 UNTS 222 (entered into force 3 September 1953).

European Union

Commission 'Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts' OJ L 2024/1689, 12.7.2024. https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf

Directive 2011/95/EU of the European Parliament and of the Council of 13 December 2011 on standards for the qualification of third-country nationals or stateless persons as beneficiaries of international protection, for a uniform status for refugees or for persons eligible for subsidiary protection, and for the content of the protection granted (recast).

Regulation (EU) No 604/2013 of the European Parliament and of the Council of 26 June 2013 establishing the criteria and mechanisms for determining the Member State responsible for examining an application for international protection lodged in one of the Member States by a third-country national or a stateless person (recast) ('Dublin Regulation').

Australia

Chan v MIEA (1989) 169 CLR 379.
Migration Legislation Amendment (Electronic Transactions and Methods of Notification) Act 2001 (Cth) (Australia).
Minister for Immigration and Ethnic Affairs v Wu Shan Liang [1996] HCA.
Re Minister for Immigration and Multicultural Affairs; Ex parte Miah [2001] HCA 22.
S152/2003 v. MIMA [2004] HCA 18 (Aus. HC, Apr. 21, 2004).

Canada

Adjei v MEI, [1989] 57 DLR 4th 153 (Can. FC A.): SCC (1995).
Gzim Bela v Canada [2013] FC 784 (Can. FC, July 12, 2013).
Gavryushenko v Canada (Minister of Citizenship and Immigration), (2000) 194 FTR 161 (Can. FCTD, Jul. 26, 2000).
Ward v. Canada (Attorney General) [1993]2 SCR 689 (Can. SC, Jun. 30, 1993.
Yusuf v Canada (1992) 1 FC 629 (Can. FCA, Oct. 24, 1991).

Ireland

Zgnat'ev v Minister for Justice, Equality and Law Reform [2001] IEHC 70 (Ir. HC, Mar. 29, 2001).

Denmark

Aliens (Consolidation) Act.

UK

Karanakaran v Secretary of State for the Home Department [2000] EWCA Civ 11, [2000] 3 ER 449.
Katrinak v SSHD, [2001] EWCA Civ 832 (Eng. CA, May 10, 2001).
R (Saber) v Secretary of State for the Home Department [2007] UKHL 57 (UKHL, Dec. 12, 2007).
R v. SSHD ex parte Sivakumaran [1988] 1 All E.R. 193 (UKHL).
SM (Section 8: Judge's process) Iran [2005] UKAIT 00116 (5 July 2005).
USA

INS v Cardoza-Fonseca (1987) 467 U.S. 407 (USSC).
Marcos v Attorney General 2005 US App., Lexis 10698 (USCA9, Jun. 9, 2005).
Ravichandran v Secretary of State for the Home Department [1996] Imm AR 97.
Tatsiana Boika v Attorney General 727 F.3d 735 (USCA7, Aug. 16, 2013).

Secondary Sources

Ahani, Narges, Tommy Andersson, Alessandro Martinello, Alexander Teytelboym, and Andrew C Trapp. "Placement Optimization in Refugee Resettlement." *Operations Research* 69, no 5 (2021): 1468–1486. https://doi.org/10.1287/opre.2020.2093.

Ahmed, Nahian, Nazmul Alam Diptu, M Sakil Khan Shadhin, M Abrar Fahim Jaki, M Ferdous Hasan, M Naimul Islam, and Rashedur M Rahman. "Artificial Neural Network and Machine Learning Based Methods for Population Estimation of Rohingya Refugees: Comparing Data-driven and Satellite Image-driven Approaches." *Vietnam Journal of Computer Science* 6, no 4 (2019): 439–455. https://doi.org/10.1142/S2196888819500246.

Aleven, Vincent. "Using Background Knowledge in Case-based Legal reasoning: A Computational Model and an Intelligent Learning Environment." *Artificial Intelligence* 150, nos 1–2 (2003): 183–237. https://doi.org/10.1016/S0004-3702(03)00105-X.

Aletras, Nikolaos, Dimitrios Tsarapatsanis, Daniel Preoţiuc-Pietro, and Vasileios Lampos. "Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective." *PeerJ Computer Science* 2 (2016): e93. https://doi.org/10.7717/PEERJ-CS.93.

The Algorithm Register of the Dutch Government. https://algoritmes.overheid.nl/en/algoritme/36268942 .

Anderson, Adrienne, Jane McAdam, Hélène Lambert, and Michelle Foster. "A Well-founded Fear of Being Persecuted … But When?" *Sydney Law Review* 42, no 2 (2020): 155–180.

Arvidsson, Matilda, and Gregor Noll. "Decision Making in Asylum Law and Machine Learning: Autoethnographic Lessons Learned on Data Wrangling and Human Discretion." *Nordic Journal of International Law* 92, no 1 (2023): 56–92.

Askeland, Bjarte. "The Potential of Abductive Legal Reasoning." *Ratio Juris* 33, no 1 (2020): 66–81. https://doi.org/10.1111/raju.12268.

Ashley, Kevin D. *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge: Cambridge University Press, 2017.

Barale, Claire, Mark Klaisoongnoen, Pasquale Minervini, Michael Rovatsos, and Nehal Bhuta. "AsyLex: A Dataset for Legal Language Processing of Refugee Claims." In *Proceedings of the Natural Legal Language Processing Workshop* 2023, 244–257. 2023.

Barocas, Solon, and Andrew D Selbst. "Big Data's Disparate Impact." *California Law Review* 104 (2016): 671.

Benesty, Michaël. "NLP applied to French legal decisions." Slideshare. https://www.slideshare.net/MichaelBENESTY/supra-legem-ml-meetup-feb-2016.

Benesty, Michaël, "The Judge Statistical Data Ban – My Story." *Artificial Lawyer*, 7 June 2019. https://www.artificiallawyer.com/2019/06/07/the-judge-statistical-data-ban-my-story-michael-benesty.

Bossin, Michael, and Laila Demirdache. "A Canadian Perspective on the Subjective Component of the Bipartite Test for 'Persecution': Time for Re-evaluation." *Refuge* 22 (2004): 108. https://doi.org/10.25071/1920-7336.21322.

Branting, L. Karl, Craig Pfeifer, Bradford Brown, Lisa Ferro, John Aberdeen, Brandy Weiss, Mark Pfaff, and Bill Liao. "Scalable and Explainable Legal Prediction." *Artificial Intelligence and Law* 29 (2021): 213–238. https://doi.org/10.1007/s10506-020-09273-1.

Bringas Colmenarejo, Alejandra. *Artificial Intelligence and Big Data Analytics in Decision Making: The Migration and Asylum Systems.* Glasgow: IMSISS, 2020.

Cameron, Hilary Evans. *Refugee Law's Fact-finding Crisis: Truth, Risk, and the Wrong Mistake.* Cambridge: Cambridge University Press, 2018.

Cameron, Hilary Evans, Avi Goldfarb, and Leah Morris. "Artificial Intelligence for a Reduction of False Denials in Refugee Claims." *Journal of Refugee Studies* 35, no. 1 (2022): 493–510. https://doi.org/10.1093/jrs/feab054.

Chen, Daniel L, and Jess Eagel. "Can Machine Learning Help Predict the Outcome of Asylum Adjudications?" *In Proceedings of the 16th Edition of the International Conference on Artificial Intelligence and Law*, pp. 237–240. 2017.

Citron, Danielle Keats, and Frank Pasquale. "The Scored Society: Due Process for Automated Predictions." *Washington Law Review* 89 (2014): 1.

Collenette, Joe, Katie Atkinson, and Trevor Bench-Capon. "An Explainable Approach to Deducing Outcomes in European Court of Human Rights Cases Using ADFs." In *Computational Models of Argument*, 21–32. Amsterdam: IOS Press, 2020. https://doi.org/10.3233/FAIA200488.

Commonwealth Ombudsman, *Automated Decision-making Better Practice Guide*. Last updated September 17, 2024. https://www.ombudsman.gov.au/__data/assets/pdf_file/0029/288236/OMB1188-Automated-Decision-Making-Report_Final-A1898885.pdf.

De Ville, Barry. "Decision Trees." *Wiley Interdisciplinary Reviews: Computational Statistics* 5, no 6 (2013): 448–455.

Dunn, Matt, Levent Sagun, Hale Şirin, and Daniel Chen. "Early Predictability of Asylum Court Decisions." *In Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*. 2017.

El Ghosh, Mirna, Hala Naja, Habib Abdulrab, and Mohamad Khalil. "Towards a Legal Rule-based System Grounded on the Integration of Criminal Domain Ontology and Rules." *Procedia Computer Science* 112 (2017): 632–642. https://doi.org/10.1016/j.procs.2017.08.109.

European Commission for the Efficiency of Justice, European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment, adopted at the 31st plenary meeting of the CEPE (Strasbourg, 3–4 December 2018. https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c.

European Commission for the Efficiency of Justice, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment*. 2018. https://shorturl.at/r7vLT.

Ewert, Lina. "Displaced, Profiled, Protected? Humanitarian Surveillance and New Approaches to Refugee Protection." In *Postcoloniality and Forced Migration*, edited by Martin Lemberg-Pedersen, Sharla M Fett, Lucy Mayblin, Nina Sahraoui and Eva Magdalena Stambøl, 93–108. Bristol: Bristol University Press, 2022.

Fussell, Sidney. "Dystopian Robot Dogs are the Latest in a Long History of US–Mexico Border Surveillance." *The Guardian*, 16 February 2022. https://www.theguardian.com/us-news/2022/feb/16/robot-dogs-us-mexico-border-surveillance-technology.

Gigerenzer, Gerd, and Wolfgang Gaissmaier. "Heuristic Decision Making." *Annual Review of Psychology* 62, no 1 (2011): 451–482. https://doi.org/10.1146/annurev-psych-120709-145346.

Guarini, Marcello, Amy Butchart, Paul Simard Smith, and Andrei Moldovan. "Resources for Research on Analogy: A Multi-disciplinary Guide." *Informal Logic* 29, no 2 (2009): 84–197. http://dx.doi.org/10.22329/il.v29i2.1225.

Hall, Paula, and Debbie Ellis. "A Systematic Review of Socio-technical Gender Bias in AI Algorithms." *Online Information Review* 47, no 7 (2023): 1264–1279. https://doi.org/10.1108/OIR-08-2021-0452.

Hanna, Alex, Emily Denton, Andrew Smart, and Jamila Smith-Loud. "Towards a Critical Race Methodology in Algorithmic Fairness." *In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 501–512. 2020.

Harman, Gilbert H. "The Inference to the Best Explanation." *The Philosophical Review* 74, no 1 (1965): 88–95. https://doi.org/10.2307/2183532.

Harlow, Carol, and Richard Rawlings. "Proceduralism and Automation: Challenges to the Values of Administrative Law." In *The Foundations and Future of Public Law (in Honour of Paul Craig)*. Oxford: Oxford University Press, 2019.

Hathaway, James C, and Michelle Foster. *The Law of Refugee Status*. Cambridge: Cambridge University Press, 2014.

Hathaway, James C, and William S Hicks. "Is There a Subjective Element in the Refugee Convention's Requirement of Well-Founded Fear?" *Michigan Journal of International Law* 26 (2004): 505.

Hillier, Will. "What is a Decision Tree and How is It Used?" Career Foundry, 17 April 2023. https://careerfoundry.com/en/blog/data-analytics/what-is-a-decision-tree.

Jacobsen, Katja Lindskov. "Experimentation in Humanitarian Locations: UNHCR and Biometric Registration of Afghan Refugees." *Security Dialogue* 46, no 2 (2015): 144–164. https://doi.org/10.1177/0967010614552545.

Josephson, John R. "On the Proof Dynamics of Inference to the Best Explanation." In *The Dynamics of Judicial Proof: Computation, Logic, and Common Sense*, edited by Marilyn MacCrimmon, 287–305. Heidelberg: Physica-Verlag, 2002.

Katsikouli, Panagiota, William H Byrne, Thomas Gammeltoft-Hansen, Anna Højberg Høgenhaug, Naja Holten Møller, Trine Rask Nielsen, Henrik Palmer Olsen, and Tijs Slaats. "Machine Learning and Asylum Adjudications: From Analysis of Variations to Outcome Predictions." *IEEE Access* 10 (2022): 130955–130967. https://doi.org/10.1109/ACCESS.2022.3229053.

Katz, Daniel Martin, Michael J Bommarito, and Josh Blackman. "A General Approach for Predicting the Behavior of the Supreme Court of the United States." *PloS One* 12, no 4 (2017): e0174698. https://doi.org/10.1371/journal.pone.0174698.

Kinchin, Niamh. "Technology, Displaced? The Risks and Potential of Artificial Intelligence for Fair, Effective, and Efficient Refugee Status Determination." *Law in Context* 37, no 3 (2021): 45–73.

Kroll, JA. "Accountable Algorithms. PhD thesis, Princeton University, 2015.

Kusner, Matt J., and Joshua R. Loftus. "The Long Road to Fairer Algorithms." *Nature* 578, no 7793 (2020): 34–36. https://doi.org/10.1038/d41586-020-00274-3.

Law Council of Australia. "Positioning Australia as a Leader in Digital Economy Regulation: Automated Decision Making and AI Regulation – Issues Paper: Digital Technology Taskforce," June 3, 2022. https://lawcouncil.au/resources/submissions/positioning-australia-as-a-leader-in-digital-economy-regulation—automated-decision-making-and-ai-regulation.

Lehr, David, and Paul Ohm. "Playing with the Data: What Legal Scholars Should Learn About Machine Learning." *University of California Davis Law Review* 51 (2017): 653.

Lex Machina. Website. https://lexmachina.com.

Lo, Dawn. "Can AI Replace a Judge in the Courtroom?" UNSW Newsroom, 1 October 2021. https://www.unsw.edu.au/newsroom/news/2021/10/can-ai-replace-judge-courtroom.

Linant de Bellefonds, Xavier. “L’utilisation d’un ‘système expert’ en droit comparé.” *Revue internationale de droit comparé* 46, no 2 (1994): 703–718.

Luker, Trish. “Decision Making Conditioned by Radical Uncertainty: Credibility Assessment at the Australian Refugee Review Tribunal.” *International Journal of Refugee Law* 25, no 3 (2013): 502–534. https://doi.org/10.1093/ijrl/eet043.

Macklin, Audrey. “Coming Between Law and the State: Church Sanctuary for Non-citizens.” *Nexus* 49 (2005): 51.

Mazzeo, Antonio, “Border Surveillance, Drones and Militarisation of the Mediterranean.” Statewatch, 6 May 2021. https://www.statewatch.org/analyses/2021/border-surveillance-drones-and-militarisation-of-the-mediterranean.

McCarroll, Estefania. “Weapons of Mass Deportation: Big Data and Automated Decision-making Systems in Immigration Law.” *Georgetown Immigration Law Journal* 34 (2019): 705.

McGill, Jena, and Amy Salyzyn. “Judging by Numbers: Judicial Analytics, the Justice System and Its Stakeholders.” *Dalhousie Law Journal* 44 (2021): 249.

McGregor, Lorna, Daragh Murray, and Vivian Ng. “International Human Rights Law as a Framework for Algorithmic Accountability.” *International and Comparative Law Quarterly* 68, no 2 (2019): 309–343.

McGregor Richmond, Karen. *Explainable Legal Case-Based Reasoning (Xl-Cbr) in Asylum Adjudication: Evidential Analysis and AI Prediction.* Copenhagen: Danish National Research Foundation Centre of Excellence in International Law, 2022.

McGregor Richmond, Karen, Satya M Muddamsetty, Thomas Gammeltoft-Hansen, Henrik Palmer Olsen, and Thomas B Moeslund. “Explainable AI and Law: An Evidential Survey.” *Digital Society* 3, no 1 (2024): 1. https://doi.org/10.1007/s44206-023-00081-z.

Meagher, Dan. *Law, Society and Civil Rights*. Geelong: Deakin University, 2011.

Medvedeva, Masha, Michel Vols, and Martijn Wieling. “Using Machine Learning to Predict Decisions of the European Court of Human Rights.” *Artificial Intelligence and Law* 28, no 2 (2020): 237–266. https://doi.org/10.1007/s10506-019-09255-y.

Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. “A Survey on Bias and Fairness in Machine Learning.” *ACM Computing Surveys (CSUR)* 54, no 6 (2021): 1–35.

Memon, Amina, Zoe Given-Wilson, Derya Ozkul, Karen McGregor Richmond, Julia Muraszkiewicz, Ella Weldon, and Cornelius Katona. “Artificial Intelligence (AI) in the Asylum System.” *Medicine, Science and the Law* 64, no 2 (2024): 87–90. https://doi.org/10.1177/00258024241227721.

Molnar, Petra. “Robots and Refugees: The Human Rights Impacts of Artificial Intelligence and Automated Decision-making in Migration.” In *Research Handbook on International Migration and Digital Technology*, edited by Marie McAuliffe, 134–151. Cheltenham: Edward Elgar, 2021.

Morris, Jason. “User-friendly Open-source Case-based Legal Reasoning.” *In Proceedings of the Seventeenth International Conference on Artificial Intelligence and Law*, 270–271. 2019. https://doi.org/10.1145/3322640.3326717.

Muddamsetty, Satya Mahesh, Mohammad Naser Sabet Jahromi, Thomas B Moeslund, and Thomas Gammeltoft-Hansen. “Danish Asylum Adjudication Using Deep Neural Networks and Natural Language Processing.” In *International Workshop on Juris-informatics (JURISIN 2023*), 92–105. 2023.

Murphy, Kevin P. *Machine Learning: A Probabilistic Perspective.* Cambridge, MA: MIT Press, 2012.

Nordic Asylum Law and Data Lab. Website. https://asylumdata.ku.dk/research/data-science-for-asylum-legal-landscaping-data4all.

NSW Parliament. *Automated Decision-Making in NSW: Mapping and Analysis of the Use of ADM Systems by State and Local Governments*. Sydney: NSW Government, 2024. https://www.parliament.nsw.gov.au/tp/files/187691/Executive-Report-ADMS.pdf.

Office of the High Commissioner for Human Rights, *Digital Border Governance: A Human Rights Based Approach*. Geneva: United Nations, 2023. https://www.ohchr.org/sites/default/files/2023-09/Digital-Border-Governance-A-Human-Rights-Based-Approach.pdf.

Parmar, Aakash, Rakesh Katariya, and Vatsal Patel. “A Review on Random Forest: An Ensemble Classifier.” In *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018*, 758–763. Dordrecht: Springer, 2019.

Republic of Estonia Ministry of Justice. “Estonia Does Not Develop AI Judge”. 16 February 2022. https://www.just.ee/en/news/estonia-does-not-develop-ai-judge.

Piccolo, Sebastiano Antonio, Panagiota Katsikouli, Thomas Gammeltoft-Hansen, and Tijs Slaats. “On Predicting and Explaining asylum Adjudication.” *In Proceedings of the Nineteenth International Conference on Artificial Intelligence and Law*, 217–226. 2023.

Sandvik, Kristin Bergtora, Katja Lindskov Jacobsen, and Sean Martin McDonald. “Do No Harm: A Taxonomy of the Challenges of Humanitarian Experimentation.” *International Review of the Red Cross* 99, no 904 (2017): 319–344.

Sourdin, Tania. “Judge v Robot? Artificial Intelligence and Judicial Decision-making.” *The University of New South Wales Law Journal* 41, no 4 (2018): 1114–1133.

Steponenaite, Vilte Kristina, and Peggy Valcke. "Judicial Analytics on Trial: An Assessment of Legal Analytics in Judicial Systems in Light of the Right to a Fair Trial." *Maastricht Journal of European and Comparative Law* 27, no 6 (2020): 759–773.

Stewart, Pamela, and Anita Stuhmcke. "Judicial Analytics and Australian Courts: A Call for National Ethical Guidelines." *Alternative Law Journal* 45, no 2 (2020): 82–87. https://doi.org/10.1177/1037969X19899674.

Surdeanu, Mihai, Ramesh Nallapati, George Gregory, Joshua Walker, and Christopher D Manning. "Risk Analysis for Intellectual Property Litigation." *In Proceedings of the 13th International Conference on Artificial Intelligence and Law*, 116–120. 2011.

Suresh, Harini, and John Guttag. "A Framework for Understanding Sources of Harm Throughout the Machine Learning Life Cycle." In *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, 1–9. 2021.

Tao, Will. "A Closer Look at How IRCC's Officer and Model Rules Advanced Analytics Triage Works." Vancouver Immigration Blog, 29 March 2022. https://vancouverimmigrationblog.com/a-closer-look-at-how-irccs-officer-and-model-rules-advanced-analytics-triage-works.

Undavia, Samir, Adam Meyers, and John E Ortega. "A comparative study of classifying legal documents with neural networks." In *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 515. 2019.

UNHCR. "Interpreting Article 1 of the 1951 Convention Relating to the Status of Refugees." https://www.refworld.org/reference/research/unhcr/2001/en/29198.

UNHCR. *Beyond Proof Credibility: Assessment in EU Asylum Systems*. 2013. https://www.unhcr.org/en-au/protection/operations/51a8a08a9/full-report-beyond-proof-credibility-assessment-eu-asylum-systems.html.

UNHCR. *Handbook on Procedures and Criteria for Determining Refugee Status Under the 1951 Convention and the 1967 Protocol Relating to the Status of Refugees*. UN Doc. HCR/IP/4/Eng/REV.3 (2011).

Walton, Douglas. *Abductive Reasoning*. Tuscaloosa, AL: University of Alabama Press, 2014.

Waterman, Donald A., Jody Paul, and Mark Peterson. "Expert Systems for Legal Decision Making." *Expert Systems* 3, no 4 (1986): 212–226. https://doi.org/10.1111/j.1468-0394.1986.tb00203.x.

Yearwood, John, A Stranieri, and J Zeleznikow. "Case-based Retrieval of Refugee Review Tribunal Text Cases." *In Legal Knowledge Based Systems. JURIX: The Tenth Conference*, 67–83. 1997.

Završnik, Aleš. "Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings." *European Journal of Criminology* 18, no 5 (2021): 623–642. https://doi.org/10.1177/1477370819876762.