University of Bakht Alruda Quarterly Refereed Scientific Journal :(UOBQRSJ)
Issue No. (37), Volume (1), June, 2023
Arab Impact Factor: (AIF: 1.12)
ISSN: 1858-6139, Online

State Liability Arising from the use of Algorithmic Decision Making and Artificial Intelligence in Areas Affecting Human Rights: Border Control and Predictive Policing

Almoatuz A. Munsoor[1]
[1]Associate Professor of Public International Law University of Jeddah, KSA
Email: aaadam@uj.edu.sa
Orcid: 0009-0004-5816-7142

Received: 26/04/2023 Published: 30/06/2023

Abstract
This study explores the evolving legal landscape of state accountability when algorithmic decision-making and AI systems infringe upon fundamental human rights in highly sensitive areas like border control and predictive policing. While governments increasingly turn to automated tools in pursuit of efficiency and security, these technologies introduce profound accountability gaps, especially when they yield discriminatory results, violate privacy. The Study evaluates whether existing international human rights law, domestic legal frameworks, and tort principles designed for an era of human decision makers can adequately respond to the novel harms wrought by algorithms. It grapples with core challenges such as proving causation, pinpointing responsibility, and demonstrating discriminatory intent in the face of opaque machine learning models. Through a comparative analysis of responses from the European Union, the United States, and other jurisdictions, the research argues for a fundamental rethinking of state liability one that honestly confronts the unique nature of AI, including its lack of transparency, its capacity for scale, and its tendency to embed systemic bias. At its heart, the paper contends that states must shoulder greater responsibility for algorithmic harms, given their duty to protect rights, their control over deployment, and their unique ability to understand and mitigate technological risks. It concludes that building robust, thoughtful liability frameworks is not just beneficial but essential to safeguarding human dignity in an age of growing automation.
Keywords: State liability, Algorithmic decision-making, Artificial intelligence, Human rights, Discrimination.

1. Introduction

We stand at a legal crossroads in the digital age, confronting a pressing and difficult question: how should states be held accountable when the artificial intelligence systems they employ violate fundamental human rights? Across the globe, governments have eagerly adopted algorithmic decision-making to handle complex public tasks, from managing immigration flows to anticipating criminal activity. The appeal is clear: promises of streamlined efficiency, consistent outputs, and significant cost savings (Završnik, 2020). At international borders, algorithms now screen visa applicants, flag security threats, and assess asylum claims. In some cities, law enforcement agencies use predictive tools to direct patrols, identify potential suspects, and forecast where crime might occur next (Ferguson, 2017).

Yet, a troubling body of evidence is emerging. Far from being neutral arbiters, these technologies often mirror and magnify our deepest societal flaws. They routinely produce discriminatory outcomes, violate personal privacy, and undermine individual liberty and dignity—frequently without meaningful transparency or avenues for redress (Eubanks, 2018). This reality forces us to re-examine the very foundations of legal responsibility. The traditional concepts of state liability were built around the human decision-maker: a person whose reasoning could be uncovered, whose intent could be probed, and whose mistakes could be individually corrected. Algorithmic systems shatter this paradigm. They are often inscrutable "black boxes," their complexity diffusing responsibility among designers, vendors, procurement officers, and end-users (Ananny & Crawford, 2018). Liability becomes a ghost in the machine when no one can explain why an algorithm made a particular choice, or when proprietary claims wall off its inner workings from scrutiny (Pasquale, 2015). This paper focuses intently on two critical domains—border control and predictive policing—because they represent the sharp edge of this challenge, where human rights are most acutely at stake. The algorithms governing borders hold immense power: they decide who may enter a country, seek refuge from persecution, or reunite with loved ones, directly impacting rights to movement, family life, and safety (Molnar & Gill, 2018). Similarly, predictive policing tools influence which neighborhoods are surveilled, who is stopped and searched, and who is deemed a threat, raising alarming concerns about racial profiling and the erosion

of privacy and equal protection (Richardson et al., 2019). In both arenas, state power is concentrated, vulnerable populations are affected, and we see with stark clarity how technical design choices carry profound ethical and human consequences. To navigate this complex terrain, the paper unfolds in several connected stages. First, it examines the current state of AI deployment in border control and policing to ground the discussion in real-world practices and documented harms. Second, it critically analyzes existing frameworks of state responsibility, testing their fitness for purpose in the algorithmic age. Third, it dissects the specific legal hurdles these systems create, from establishing causation to proving discrimination. Fourth, it surveys the innovative, though often fragmented, legal and regulatory responses beginning to take shape around the world. Finally, drawing on these insights, it proposes principles for a reconceived model of state liability—one designed specifically for the realities of AI, ensuring that the protection of human rights remains paramount.

2. Algorithmic Systems in Border Control and Predictive Policing

2.1Border Control Applications

The landscape of immigration and border management is being radically reshaped by automated systems that claim to bring scientific objectivity to deeply human processes. Governments now deploy algorithms that assess risk, verify identity, and make consequential recommendations about who can move across borders. This technological transformation manifests in several concrete ways. Visa applications, for instance, are no longer assessed by officers alone; they are fed into systems that analyze biographical details, travel history, social media footprints, and behavioral patterns to generate a "risk score" that heavily sways approval or denial (Amoore, 2020). At the physical border, biometric checkpoints use facial recognition, iris scanning, and fingerprint matching to verify identity, creating sprawling databases that allow states to track individuals across countless interactions and over time (Molnar & Gill, 2018). Furthermore, predictive analytics are employed to flag travelers deemed likely to overstay visas or pose security threats, often basing these judgments on correlations between demographic data and historical patterns (Beduschi, 2021). The pursuit of a perfectly secure, automated border is pushing systems into even more novel and intimate territory. We now see the development of tools that claim to detect deception by analyzing an asylum seeker's facial expressions, vocal stress, and physiological responses during automated interviews (Metcalf et al., 2021). Machine learning models are being trained on past asylum decisions to predict future outcomes, a practice that risks hardcoding the biases of yesterday into the jurisprudence of tomorrow (Raso et al., 2018). Perhaps most expansive are the integrated platforms that fuse data from criminal records, intelligence reports, financial transactions, and commercial data brokers to assemble detailed risk profiles, creating a digital dossier that follows a traveler long before and after they reach a checkpoint (Aradau & Blanke, 2017). Beneath the veneer of efficiency and innovation, however, lie serious and growing concerns about fairness, privacy, and due process. Mounting evidence reveals that these systems are far from infallible. Facial recognition technology, a cornerstone of many biometric systems, has been shown to fail significantly more often for people of color, women, and the elderly, leading to real-world consequences like wrongful detentions and denied entry (Buolamwini & Gebru, 2018). More insidiously, algorithms trained on historical enforcement data inevitably learn and reproduce past discriminatory practices, creating a pernicious feedback loop. They may disproportionately target individuals from specific nationalities or religious backgrounds, not because they pose a greater risk, but because biased past data tells the algorithm they should (Završnik, 2020). This problem is compounded by a profound lack of transparency. When a traveler is denied entry based on an algorithmic score, they are often unable to learn the "why" behind the decision, stripping them of any meaningful ability to appeal or identify potential discrimination (Molnar & Gill, 2018).

2.2 Predictive Policing Applications

In the realm of law enforcement, a parallel revolution is underway, driven by the promise of data-driven crime prevention. Predictive policing algorithms aim to forecast criminal activity by analyzing past data to guide police resources. This takes two primary forms. The first is geographic: systems analyze historical crime data to generate maps highlighting "hot spots" where crime is deemed statistically likely to occur, thereby directing patrols to specific streets or neighborhoods (Perry et al., 2013). The second is person-based: algorithms assign risk scores to individuals based on a mix of factors like criminal history, social associations, location, and behavior, placing those with high scores on watchlists as people likely to be involved in violence, either as perpetrators or victims (Ferguson, 2017). A related tactic involves social network analysis, which flags individuals for increased scrutiny simply based on their associations with known offenders (Brayne, 2017). The technological toolkit is expanding into increasingly pervasive surveillance. Automated license plate readers silently log the

movements of vehicles city-wide, building massive databases of travel patterns that police can data-mine for leads (Brayne, 2017). Body cameras and street-level surveillance networks are increasingly integrated with facial recognition software, enabling real-time identification and tracking of individuals deemed suspicious (Garvie et al., 2016). Perhaps most controversial is the mining of social media posts, online purchases, and digital behavior with machine learning tools to try to identify individuals who might be planning crimes—a form of pre-crime surveillance that tests the boundaries of constitutional protection (Ferguson, 2017). A wealth of academic and investigative work now shows that these tools frequently perpetuate systemic injustice rather than curbing it. A core finding is that predictive policing systems often direct more police to minority neighborhoods. Crucially, this is not necessarily because crime is inherently higher there, but because historical policing patterns—which themselves may have been biased—have produced more crime data from those areas. The algorithm learns from this skewed data and recommends even more policing, which in turn generates more data, creating a self-fulfilling prophecy of over-policing (Lum & Isaac, 2016). This creates a destructive feedback loop: algorithms send police to certain communities, heightened surveillance produces more arrests, and this new data trains the algorithm to see those communities as even higher risk, cementing cycles of discrimination (Richardson et al., 2019). As with border systems, opacity shrouds the process. Individuals have no way of knowing if or why they are on a risk list, and no clear path to challenge their inclusion, leaving them subject to heightened police attention without recourse (Brayne, 2017).

3. Existing Frameworks of State Liability

3.1International Human Rights Law

International human rights law provides a crucial, principled starting point for this discussion. At its core lies a fundamental premise: states are responsible for upholding rights within their jurisdiction, and this duty cannot be outsourced or automated away. This principle is enshrined in foundational texts like the International Covenant on Civil and Political Rights, which obliges states not only to respect but also to ensure the rights of all individuals under their authority (United Nations, 1966). This responsibility is active and continuous. It requires states to take steps to prevent violations, to investigate them thoroughly when they occur, to provide real remedies to victims, and to reform systems to prevent repetition (UN Human Rights Committee, 2004). Critically, this accountability extends across the entire government—legislative, executive, and judicial—regardless of which branch's action or inaction led to the harm (International Law Commission, 2001). The content of these obligations is directly relevant to algorithmic systems. The law prohibits discrimination on a wide range of grounds including race, sex, religion, and national origin, and it demands that any differential treatment by the state must serve a legitimate goal, be proportionate, and not be arbitrary (United Nations, 1966; UN Human Rights Committee, 2018). Equally important are the guarantees of due process: when a state decision affects a person's rights, the individual is entitled to notice, a fair hearing, access to the reasons for the decision, and a chance to appeal (UN Human Rights Committee, 1992). These are not mere procedural formalities; they are the bedrock of a system that respects human dignity. There is a growing and conscious effort to apply these timeless principles to the novel context of algorithmic governance. United Nations experts and treaty bodies have clarified that a state cannot wash its hands of responsibility by pointing to a machine or a private contractor as the decision-maker; the buck stops with the state (UN Special Rapporteur on Extreme Poverty, 2019). They have emphasized that the quest for efficiency or budget savings can never justify discriminatory outcomes or strip away basic procedural protections (Nyst & Monaco, 2018). The emerging interpretation is clear: for algorithmic systems to be compatible with human rights, they must incorporate meaningful human oversight, provide intelligible explanations, and maintain accessible channels for individuals to seek correction and redress (Council of Europe, 2019). Despite this strong normative foundation, significant practical challenges arise when seeking to apply international human rights law to algorithmic harms. The traditional legal models assume a knowable human actor with discernible intent—an assumption that crumbles when facing an opaque model that may have "learned" its biases from flawed data (Hildebrandt, 2008). Proving a causal link between a specific algorithmic design flaw and a specific human injury is a formidable technical and legal hurdle. Furthermore, while international bodies can name, shame, and issue recommendations, their tools for compelling a state to dismantle or fundamentally redesign a harmful algorithmic system are often limited to diplomatic pressure, leaving a gap between recognition of a violation and effective enforcement (Alston, 2019).

3.2 Domestic Legal Frameworks

Within domestic legal systems, the picture is more varied but faces analogous hurdles. Most countries possess a suite of legal tools—constitutional, administrative, and tort-based—designed to check state

power and compensate individuals for its abuse. Constitutions commonly guarantee equality, due process, and privacy (Alston & Goodman, 2012). Administrative law requires government decisions to be transparent, reasoned, and subject to challenge (Rose, 2015). Tort law provides a path to sue the government for negligence or intentional harms (Steele, 2020). In theory, this arsenal should protect against algorithmic injustice. In practice, it often stumbles on the realities of the technology. Several legal doctrines, designed for another era, now act as formidable barriers. The ancient principle of sovereign immunity, still potent in many jurisdictions, shields the government from lawsuit unless the plaintiff can clear a high bar, such as proving an official violated a "clearly established" right (Fallon et al., 2015). Standing doctrines, which require a plaintiff to show a direct, personal injury, make it exceptionally difficult to challenge systemic algorithmic bias that diffusely harms an entire community (Citron & Pasquale, 2014). Perhaps most critically, the legal process of "discovery," which allows parties to access relevant evidence, is often blocked. Governments and vendors routinely invoke trade secrecy, national security, or proprietary rights to withhold the source code, training data, and validation studies that are the only keys to proving an algorithm is faulty or biased (Wexler, 2018). One cannot challenge what one cannot see. Even if a case proceeds, the substantive legal elements become difficult to satisfy. Proving discriminatory "intent" is a central requirement in many discrimination laws, but how does one prove the intent of an algorithm? Courts struggle when a system uses facially neutral factors (like postal code or social network patterns) that are perfect proxies for race or religion (Barocas & Selbst, 2016). Causation is another thorny issue: if an algorithm is merely an "advisory tool" and a human officer makes the final choice, is it the algorithm or the human who is legally responsible for the harm? This ambiguity lets responsibility fall into the gap between them (Selbst, 2017). The sheer number of actors involved from software developers and procurement officers to agency heads and field personnel makes pinpointing the responsible party a dizzying task (Engstrom & Gelbach, 2020). Recognizing these obstacles, some jurisdictions are beginning to innovate. New laws, most notably the European Union's General Data Protection Regulation (GDPR), are creating positive rights tailored to the automated age, such as the right to an explanation and the right to human review of significant automated decisions (European Parliament and Council, 2016). Lawmakers are experimenting with mandatory algorithmic impact assessments and bias audits before high-stakes systems can be deployed (Reisman et al., 2018). In a significant shift, some courts and statutes are moving towards recognizing "disparate impact" that is, discriminatory outcomes alone—as sufficient evidence of a violation, thereby placing the burden on the government to justify its algorithm's skewed results (Hellman, 2020). These are promising, if still nascent, steps toward bending traditional legal frameworks to meet a new technological reality.

4. Challenges Algorithmic Systems Pose for Liability Frameworks

4.1 Opacity and the Problem of Explanation

Perhaps the most immediate challenge algorithmic systems pose is their sheer opacity—the "black box" problem. This isn't merely a technical inconvenience; it strikes at the heart of legal accountability. When neither the person affected nor an external reviewer can understand how a decision was reached, the very possibility of meaningful challenge evaporates. This opacity springs from multiple sources: the dizzying complexity of machine learning models, corporate assertions of trade secrecy, and the fundamental fact that some algorithms arrive at conclusions through logic paths even their designers cannot fully trace (Burrell, 2016). The result is a deeply problematic legal reality: a traveler can be denied entry or a person placed on a watchlist based on a conclusion that no human can actually articulate or justify, rendering procedural fairness an empty promise (Wachter et al., 2017). To understand why this is so intractable, we must look at the technology itself. Modern machine learning, especially deep learning via neural networks, doesn't follow a programmer's simple "if-then" rules. Instead, it identifies patterns across millions of data points and connections, creating a model whose decision-making process is often a form of high-dimensional statistics that resists straightforward translation into human reasoning (Lipton, 2018). Furthermore, these systems are trained on colossal datasets with countless variables, making it impossible to pinpoint how any single piece of information—a person's neighborhood, a purchase, a social connection—tipped the scale toward a particular outcome (Doshi-Velez & Kim, 2017). Compounding this, many models continuously evolve, meaning the system that made a decision yesterday might operate differently today, all without a human changing a line of code (Rahwan et al., 2019). This lack of transparency is often a deliberate, legally enforced choice. Technology vendors fiercely protect their algorithms as intellectual property and trade secrets, arguing that disclosure would compromise their competitive edge and invite manipulation (Brauneis & Goodman, 2018). Governments, in turn, frequently sign contracts that lock

away source code and training data under clauses of confidentiality, effectively placing crucial public functions beyond the reach of legislative oversight, judicial review, and independent academic audit (Eubanks, 2018). This creates a perverse legal shield: an individual harmed by a system cannot prove discrimination or error because the evidence needed to do so is itself deemed a protected secret (Citron & Pasquale, 2014). The clash with foundational legal principles is direct and profound. Administrative law is built on the requirement for agencies to provide reasoned explanations for their actions so courts can review them for arbitrariness—a requirement an algorithmic score cannot satisfy (Coglianese & Lehr, 2017). Due process hinges on notice and a meaningful opportunity to be heard; how can one respond to a risk score one cannot comprehend? (Citron, 2008). Even constitutional equal protection analysis, which evaluates whether government classifications serve a legitimate purpose, is paralyzed if the government cannot explain what classifications its algorithm is actually using (Kroll et al., 2017). In short, opacity doesn't just make liability hard to prove; it undermines the procedural architecture that makes liability possible in the first place.

4.2 Diffusion of Responsibility

If opacity obscures the "how," the diffusion of responsibility obscures the "who." Algorithmic decision-making sprawls across a long, fragmented chain of actors. Political leaders set the policy; procurement officers choose the vendor; engineers design the model; civil servants input the data and operate the system; and front-line workers interpret and act on its outputs. In this sprawling ecosystem, when harm occurs, every link in the chain can point elsewhere. The software company blames the government for misusing its tool; the agency blames the vendor for a faulty product; the official blames the algorithm's inscrutable logic (Nissenbaum, 1996; Katyal, 2019). Accountability dissipates into a fog of plausible deniability. For a victim seeking justice, this diffusion creates a maze with no clear exit. Who do you sue? The government official who claims was just following the system's recommendation? The vendor whose terms of service disclaim all liability for downstream use? The result is a legal game of hot potato, where both sides deflect blame to the other, often citing the technical complexity as an ultimate, unfathomable cause (Mulligan & Bamberger, 2019). Organizationally, decision-making is scattered across departments, and temporally, the key choices—funding, procurement, deployment—may have been made years apart by different administrations, making it impossible to isolate a single responsible human decision (Green & Chen, 2019; Raji et al., 2020). This problem is frequently cemented by the very contracts that govern these systems. It is common for government contracts to include clauses absolving vendors of liability for errors or discriminatory outcomes, while vendor contracts often require governments to indemnify them against lawsuits (Coglianese & Lehr, 2019; Engstrom et al., 2020). These agreements create a no-liability zone, stripping away the financial and legal incentives for either party to invest in safety, audit for bias, or care deeply about accuracy. When no one bears the full cost of failure, the economic rationale for preventing harm vanishes (Citron & Pasquale, 2014). The broader consequence is the erosion of deterrence, a core goal of liability law. Traditional tort law aims to prevent future harm by making wrongdoers internalize its costs (Shavell, 1987). But when responsibility—and thus cost—is splintered among a vendor, an agency, and various officials, no single actor has a strong incentive to invest in the costly work of bias mitigation, rigorous testing, or robust oversight (Ben-Shahar & Porat, 2016). This allows harmful systems to persist and scale, not out of malice, but because the diffuse structure of accountability insulates every participant from the pressure to demand change (Selbst, 2017).

4.3 Proof and Causation Difficulties

Even if one overcomes the barriers of opacity and diffusion, the core legal task of proving causation remains a monumental challenge. A plaintiff must demonstrate, often by a preponderance of the evidence, that "but for" the algorithmic error, they would not have suffered the harm (Hart & Honoré, 1985). This is fiendishly difficult when the tool is a black box and its output is filtered through human judgment. Did the algorithm cause the wrongful arrest, or was it the officer who independently agreed with its high-risk flag? The interplay between human and machine creates a causal fog that courts are ill-equipped to penetrate (Kleinberg et al., 2018). From the very start, plaintiffs face an evidentiary brick wall. The documents necessary to prove their case—the source code, the training data, the validation reports—are typically locked away, claimed as trade secrets by vendors or state secrets by governments (Wexler, 2018; Brauneis & Goodman, 2018). Governments may also invoke deliberative process privilege or national security to withhold information (Freeman & Rossi, 2012). This asymmetrical access to information means the party with all the evidence (the state and its vendor) can prevent the injured party from even gathering the facts needed to build a claim. Proving algorithmic discrimination adds another layer of complexity. Many systems use facially neutral proxies—like "risk

density in zip code" or "social network centrality"—that are deeply correlated with race or religion. Under legal standards that require proof of discriminatory intent, statistical evidence of stark racial disparity may not be enough (Washington v. Davis, 1976). The government can simply argue that the algorithm is blind to race and is merely optimizing for public safety, a legitimate state interest (Barocas & Selbst, 2016). Without access to the model's inner workings, disproving this claim is nearly impossible, even when the outcomes scream bias (Sweeney, 2013). Finally, the harm from these systems is often systemic, not individual, which fits poorly with a legal system built for discrete wrongs. A predictive policing algorithm may lead to the over-policing of an entire neighborhood, increasing the statistical likelihood that any resident will be arrested. But how does a single arrestee prove the algorithm specifically targeted them? (Lum & Isaac, 2016). Similarly, a border algorithm may deny thousands of visas based on a flawed criterion, but each applicant faces the immense hurdle of proving that, absent that flaw, they personally would have been approved (Molnar & Gill, 2018). Class action lawsuits offer a potential path, but they too face stiff legal requirements to certify a class, often stumbling on the very individuated nature of harm that these systems create (Wal-Mart Stores v. Dukes, 2011).

5. Emerging Legal and Regulatory Responses

5.1 European Union Approaches

The European Union has positioned itself at the forefront of the global effort to govern algorithmic accountability, constructing a multi-layered regulatory architecture. Its approach weaves together powerful data protection law, established anti-discrimination principles, and pioneering legislation specifically targeting artificial intelligence. The cornerstone of this effort is the General Data Protection Regulation (GDPR), which broke new ground by granting individuals tangible rights against automated systems. It establishes a right to meaningful information about the logic of automated decisions that significantly affect people, and a right to obtain human intervention and challenge those outcomes (European Parliament and Council, 2016; Wachter et al., 2017). This legally embeds a crucial check: the individual is no longer a passive subject of algorithmic judgment but an active participant entitled to question and appeal (Selbst & Powles, 2017). Building on this foundation, the EU's proposed AI Act represents the world's first comprehensive attempt at horizontal AI regulation. It adopts a risk-based taxonomy, outright prohibiting certain applications deemed unacceptable—like social scoring and intrusive real-time biometric surveillance in public spaces (European Commission, 2021). For high-risk systems, including those used in migration, asylum, and law enforcement, it imposes a stringent set of obligations. Developers and deployers must conduct conformity assessments, implement robust risk management and data governance, ensure transparency and human oversight, and meet strict accuracy standards (European Commission, 2021). The Act also creates ongoing duties for post-market monitoring and incident reporting, aiming for lifecycle accountability (Veale & Borgesius, 2021).

5.2 United States Developments

In contrast to the EU's comprehensive strategy, the United States pursues algorithmic accountability through a more fragmented and litigious path. Efforts are dispersed across constitutional challenges, sector-specific statutes, and a patchwork of state and local regulations. Advocates have brought cases under the Fourteenth Amendment's Equal Protection and Due Process clauses, and the Fourth Amendment's protection against unreasonable searches, arguing algorithmic tools violate these constitutional guarantees (Chander, 2017; Ferguson, 2017). Existing laws like those governing fair housing and equal credit opportunity are also being stretched to cover algorithmic discrimination (Barocas & Selbst, 2016). These legal battles, however, constantly crash against doctrines like qualified immunity for officials and restrictive standing rules, which often prevent cases from even being heard (Citron & Pasquale, 2014). This fragmentation is mirrored in the legislative landscape. While a federal Algorithmic Accountability Act has been proposed to mandate impact assessments, it has not yet become law (2022). Instead, innovation is happening locally. Cities like San Francisco and states like Massachusetts have enacted bans or strict moratoriums on government use of facial recognition (Georgetown Law Center on Privacy & Technology, 2021). This creates a complex, uneven regulatory terrain for companies and governments alike. The absence of an overarching federal privacy or AI law remains the single biggest obstacle to a coherent national approach, leaving a vacuum that the tech industry often fills with self-regulatory measures (Solove & Hartzog, 2022). The U.S. judiciary has become a critical arena, particularly in criminal justice. Landmark cases like State v. Loomis (2016) have grappled with the use of algorithmic risk assessments in sentencing, highlighting due process concerns about transparency and challenge. Scholars and litigators are actively debating whether defendants have a constitutional right to examine and contest the proprietary algorithms used against

them (Slobogin, 2019). These court battles are slowly drawing the boundaries of acceptable state use of algorithmic power. The path forward in the U.S. is fraught with political and structural challenges. Deep political polarization and powerful technology industry lobbying make consensus on strong federal regulation difficult (Citron, 2019). The federalist system allows for state-level innovation but also creates a compliance labyrinth that can stifle broader progress and cement industry-preferred, weaker standards (Hirsch, 2020). The ongoing struggle is between a desire for innovative leadership and the imperative to establish guardrails that protect civil liberties.

5.3 Comparative Insights

Looking globally reveals a spectrum of regulatory philosophies. Canada emphasizes human rights-centric review, requiring algorithmic impact assessments for government systems to ensure alignment with Charter rights (Treasury Board of Canada Secretariat, 2019). Australia leans more on adapting its strong administrative law framework, using existing mechanisms for review of government decisions to check algorithmic outputs (Australian Government, 2019). In Latin America, countries like Brazil are weaving algorithmic transparency and data protection directly into constitutional amendments and legal frameworks, reflecting a distinct regional approach (Dos Santos & Ramos, 2020). Despite different starting points, common challenges emerge everywhere. Governments universally struggle with technical capacity, lacking the in-house expertise to audit complex AI systems (Engstrom et al., 2020). Resource constraints limit effective oversight of rapidly evolving technologies (Ranchordás, 2020). And all jurisdictions wrestle with the core tension between fostering innovation and preventing harm (Calo, 2017). This suggests there is no perfect, one-size-fits-all model. The emerging global lesson is that effective governance requires a multi-layered, holistic approach. No single law or tool is sufficient. Instead, accountability is best achieved by combining: ex-ante measures like transparency and impact assessments; procedural rights like explanation and appeal; substantive standards against discrimination; independent oversight bodies; and robust remedial mechanisms (Kaminski, 2019; Raji et al., 2020). The future of algorithmic accountability lies not in a silver bullet, but in a carefully constructed ecosystem of checks and balances.

6. Toward a Reconceptualized Framework of State Liability

6.1 Foundational Principles

Any meaningful reconceptualization of state liability must begin with a clear-eyed acknowledgment: states bear a non-delegable duty to protect human rights, and this duty travels with them into the digital realm. When a government chooses to deploy algorithmic systems in sensitive domains like border control or policing, it does not outsource its constitutional or human rights obligations. Rather, it assumes a heightened responsibility—not just for the policy goals, but for the technological means used to achieve them. This responsibility is rooted in the state's unique position: it controls the decision to deploy, it funds and operates the system, and it alone possesses the authority and resources to understand and mitigate the complex risks these technologies introduce (Alston, 2019). Arguments of efficiency or cost-saving cannot exonerate the state from this fundamental duty; if anything, the pursuit of such benefits demands greater vigilance to ensure they are not achieved at the expense of justice and equity (UN Special Rapporteur on Extreme Poverty, 2019). Traditional legal frameworks for liability are built for a world of human actors. They assume we can identify a decision-maker, scrutinize their reasoning, and assign fault for intentional or negligent actions (Hildebrandt, 2008). Algorithmic systems unravel these assumptions. We are dealing with tools characterized by opacity (we cannot see how they reason), scale (a single flaw can affect thousands instantly), and adaptability (they change over time in unpredictable ways) (Rahwan et al., 2019). Therefore, a new liability model cannot simply retrofit old doctrines. It must be designed from the ground up to address systemic harm and collective responsibility, moving beyond the narrow focus on individual fault to encompass the entire socio-technical system of design, procurement, and deployment (Green, 2021). Given this reality, there is a compelling argument for moving toward a form of strict or enhanced liability for states using algorithmic systems in rights-affecting contexts. This is not about assigning blame without fault, but about creating the right incentives and acknowledging practical realities. The state is the party best positioned to prevent harm—through rigorous vetting, ongoing monitoring, and bias mitigation. Strict liability creates a powerful incentive for such precautionary measures by ensuring the state internalizes the costs of failure (Shavell, 1987). Furthermore, it recognizes a simple truth of the algorithmic age: the profound information asymmetry and technical complexity mean that individuals harmed by these systems will almost never be able to meet the traditional burden of proving specific negligence or intent (Selbst, 2017). The state chose the tool; the state must answer for its consequences.

University of Bakht Alruda Quarterly Refereed Scientific Journal :(UOBQRSJ)
Issue No. (37), Volume (1), June, 2023
Arab Impact Factor: (AIF: 1.12)
ISSN: 1858-6139, Online

6.2 Procedural Requirements

A robust liability framework must be proactive, embedding safeguards before harm occurs. This begins with mandatory transparency and assessment. Prior to deployment, the public has a right to know that an algorithmic system will be used, for what purpose, and based on what general logic and data (Kroll et al., 2017). More than disclosure, the state should be required to conduct and publish a rigorous algorithmic impact assessment. This is not a box-ticking exercise, but a thorough evaluation of potential effects on protected groups, privacy, due process, and other rights, with genuine consultation of affected communities and independent experts (Reisman et al., 2018; Green & Viljoen, 2020). Transparency builds public trust and enables informed democratic debate about the adoption of such powerful tools. When an individual faces an adverse decision from an algorithmic system, the right to an understandable explanation is paramount. This goes beyond a generic privacy policy. It requires clear notification that an automated process was involved and a meaningful account of the key factors leading to the decision—in terms the person can reasonably understand and use to mount a challenge (Wachter et al., 2017; Selbst & Barocas, 2018). While complete technical transparency may be unfeasible, "explanations" must be functional, not illusory. They must empower the individual, not placate them with corporate jargon (Kaminski & Malgieri, 2020). Furthermore, we must reject the myth of meaningless human review. Having a person "in the loop" is only valuable if that person has the training, authority, and independence to question and override the algorithmic recommendation (Green & Chen, 2019). Systems must be designed to avoid "automation bias," where humans become mere rubber stamps for the machine's output (Citron, 2008). Reviewers must be free from perverse performance metrics that reward uncritical acceptance and must have access to the information needed to make an independent judgment (Eubanks, 2018). Finally, the framework must guarantee accessible and effective challenge mechanisms. The path to contest an algorithmic decision cannot require a PhD in computer science or a team of lawyers. Streamlined administrative review processes are essential (Engstrom et al., 2020). Critically, the legal burden should shift once a challenger demonstrates a prima facie case of error or disparate impact. It should then fall to the state to justify the system's decision, proving its fairness, accuracy, and compliance with procedure (Hellman, 2020). To enable fair challenges, individuals must have access to relevant evidence, with legitimate state secrets protected through secure, supervised procedures like in-camera review (Brauneis & Goodman, 2018).

6.3 Substantive Standards

Procedural rights must be anchored in clear substantive standards that govern how these systems are built and used. The foremost standard is a stringent prohibition on discrimination. This must encompass both intentional discrimination and disparate impact—discriminatory outcomes that result from seemingly neutral criteria. When a system produces statistically significant disparities along racial, gender, or other protected lines, the burden must shift decisively to the state to prove the differential treatment is necessary to achieve a compelling interest and is narrowly tailored (Barocas & Selbst, 2016; Hellman, 2020). This requires ongoing audit trails and continuous monitoring to detect and correct bias that emerges in real-world operation (Raji & Buolamwini, 2019). Accuracy and reliability are not just technical metrics; they are legal imperatives. The required standard must be proportionate to the stakes. An algorithm that influences parole decisions or asylum claims demands a far higher threshold of validation than one recommending movie choices (Coglianese & Lehr, 2019). Validation must use representative data, test performance across all demographic subgroups, and be based on real-world outcomes, not just theoretical models (Raji et al., 2020). States must be compelled to publicly disclose key performance indicators, including error rates and confidence scores, so their claims of accuracy can be scrutinized (Selbst, 2017). The principle of proportionality must be rigorously applied. The state must demonstrate that the benefits of deploying an algorithmic system in a given context genuinely justify its intrusiveness. This involves considering less invasive alternatives and ensuring that efficiency gains never become an excuse to override fundamental rights (Zarsky, 2016; Ferguson, 2017; Alston, 2019). The use of a predictive policing algorithm in a low-crime neighborhood, for instance, may fail this test entirely. To combat technological lock-in and institutional inertia, sunset clauses and periodic review should be mandatory. No algorithmic system should be "set and forget." Governments should be required to periodically re-justify the system's continued use, reassess its effectiveness and fairness, and formally re-authorize its operation (Ranchordás, 2020). This creates a crucial off-ramp for systems that are ineffective, outdated, or harmful, and prevents them from becoming permanent, unexamined fixtures of state power (Coglianese & Lehr, 2017).

University of Bakht Alruda Quarterly Refereed Scientific Journal :(UOBQRSJ)
Issue No. (37), Volume (1), June, 2023
Arab Impact Factor: (AIF: 1.12)
ISSN: 1858-6139, Online

6.4 Remedial Mechanisms

For a liability framework to have teeth, it must be paired with robust, accessible, and multi-faceted remedies. A key innovation necessary in the algorithmic age is the strategic reversal of evidentiary burdens. Given the extreme information asymmetry where the state controls the algorithm and all its data—the law cannot place the full burden of proof on the injured individual. When a claimant demonstrates a prima facie case of harm (e.g., a statistically aberrant outcome or a procedural violation), the burden should shift to the state to prove that its system functioned lawfully, fairly, and accurately (Hellman, 2020). Statistical evidence of discriminatory patterns should be accepted as sufficient to establish a violation, absent compelling and transparent rebuttal evidence from the government (Barocas & Selbst, 2016). This acknowledges that the party with sole access to the "black box" must be responsible for opening it when challenged. Individual remedies must be comprehensive. At a minimum, they should include the reversal or reconsideration of the adverse decision through a fair process, aiming to place the individual in the position they would have been in absent the error (Green, 2021). Compensation must cover all tangible harms: lost liberty from wrongful detention, economic losses from denied opportunities, and non-pecuniary damages for emotional distress and dignitary harm (Eubanks, 2018; Hellman, 2020). The law must recognize that being subjected to a biased, inscrutable machine process is, in itself, a profound injury to personal autonomy and equal standing. However, individual compensation alone is insufficient for systemic failures. The framework must empower courts to order structural remedies, such as injunctions mandating the modification, suspension, or complete dismantling of harmful systems (Richardson et al., 2019). To ensure compliance, courts could appoint special masters or independent monitors with technical expertise to oversee reform implementation (Goel et al., 2016). Ongoing judicial supervision may be necessary until the state demonstrates sustained, verifiable compliance, moving beyond promises to measurable change (Green & Chen, 2019). Finally, proactive and independent oversight bodies are essential. Specialized agencies, insulated from political pressure and equipped with technical audit capabilities, should be empowered to investigate systems proactively, compel disclosure of code and data, and impose significant sanctions—including fines, public naming, and operational suspension—for non-compliance (Kaminski, 2019; Raji et al., 2020; Yeung, 2018). This creates a continuous layer of accountability beyond sporadic litigation.

6.5 Implementation Challenges

Translating this principled framework into practice will confront significant real-world constraints. Comprehensive oversight of all government algorithms is likely infeasible, necessitating a strategic prioritization of resources on high-stakes domains like criminal justice, immigration, and social welfare (Reisman et al., 2018). Building this oversight capacity requires substantial investment in expertise within judiciary, legislative, and civil society institutions, along with dedicated training for legal professionals (Engstrom et al., 2020; Kaminski, 2019). A persistent tension will exist between transparency for accountability and legitimate state interests in confidentiality (e.g., national security, genuine trade secrets). The framework must therefore develop nuanced access mechanisms, such as secure in-camera review by cleared experts or trusted third-party auditors, to balance these competing demands (Brauneis & Goodman, 2018; Wexler, 2018). Transparency must also be designed to avoid enabling system gaming or exposing security vulnerabilities (Kroll et al., 2017). Algorithmic accountability is inherently transnational. Governments use tools developed by foreign vendors and process data across borders, creating jurisdictional complexities (Bradford, 2020). Effective governance thus requires international cooperation to harmonize standards, share best practices, and manage conflicts of law (Bietti, 2020). The extraterritorial reach of regulations like the GDPR presents both a model and a challenge for global coordination (Selbst, 2021). Perhaps the greatest challenge is the pace of technological change. A regulatory framework must be inherently adaptable to avoid rapid obsolescence. This argues for principles-based standards (e.g., "fairness," "accountability") over prescriptive technical rules, and for built-in periodic review and revision mechanisms to incorporate new learnings and address emerging technologies like affective computing or autonomous systems (Coglianese & Lehr, 2019; Ranchordás, 2020; Crawford et al., 2019).

Conclusion

This research has argued that the accelerating deployment of algorithmic systems in areas of core state power represents not merely a technological shift, but a fundamental challenge to our legal conceptions of accountability. Traditional frameworks of state liability, sculpted in an era of tangible human actors and discernible decision trails, are profoundly ill-equipped to address harms born from opacity, scaled by automation, and diffused across networks of actors. In response, we have proposed a path toward a

reconceptualized model of state liability. This model is rooted in the non-negotiable principle that the state's duty to protect rights travels with it into the digital sphere. It accepts that the unique attributes of AI—its complexity, scalability, and adaptive nature—demand new legal tools. The proposed framework intertwines enhanced procedural safeguards (like mandatory impact assessments and a meaningful right to explanation) with rigorous substantive standards (prohibiting discriminatory impact and mandating proportional use), all backed by recalibrated remedial mechanisms that shift burdens where appropriate and enable both individual and systemic relief. The journey from principle to practice will be difficult. It will require navigating resource limitations, balancing transparency with other legitimate interests, fostering international collaboration, and maintaining regulatory agility in the face of relentless innovation. These are not reasons for inaction, but parameters for thoughtful implementation. Ultimately, the project of governing algorithmic power is not a technical sidebar to modern governance; it is central to the preservation of democratic values and human dignity in the 21st century. States cannot be permitted to embrace the efficiency of automation while disavowing the responsibility it entails. Our legal and political institutions must now evolve with purpose and clarity. They must ensure that the formidable power of algorithmic systems, however opaque their workings, remains firmly subject to the constraints of law, the scrutiny of justice, and the enduring demands of human rights. The task is urgent, complex, and essential.

References

Algorithmic Accountability Act of 2022, H.R. 6580, 117th Cong. (2022).

Alston, P. (2019). *Report of the Special Rapporteur on extreme poverty and human rights*. United Nations General Assembly. A/74/493.

Amoore, L. (2020). Cloud geographies: Computing, data, sovereignty. *Progress in Human Geography*, 42(1), 4-24.

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973-989. https://doi.org/10.1177/1461444816676645

Aradau, C., & Blanke, T. (2017). Politics of prediction: Security and the time/space of governmentality in the age of big data. *European Journal of Social Theory*, 20(3), 373-391. https://doi.org/10.1177/1368431016667623

Australian Government. (2019). *Artificial intelligence: Australia's ethics framework*. Department of Industry, Science, Energy and Resources.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671-732. https://doi.org/10.15779/Z38BG31

Beduschi, A. (2021). International migration management in the age of artificial intelligence. *Migration Studies*, 9(3), 576-596.

Ben-Shahar, O., & Porat, A. (2016). Personalizing negligence law. *New York University Law Review*, 91(3), 627-688.

Bietti, E. (2020). From ethics washing to ethics bashin: A view on tech ethics from within moral philosophy. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 210-219. https://doi.org/10.1145/3351095.3372860

Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press.

Brauneis, R., & Goodman, E. P. (2018). Algorithmic transparency for the smart city. *Yale Journal of Law & Technology*, 20(1), 103-176.

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, 82(5), 977-1008.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 77-91.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*.

Calo, R. (2017). Artificial intelligence policy: A primer and roadmap. *UC Davis Law Review*, 51(2), 399-435.

Chander, A. (2017). The racist algorithm? *Michigan Law Review*, 115(6), 1023-1045. https://doi.org/10.36644/mlr.115.6.racist

Citron, D. K. (2008). Technological due process. *Washington University Law Review*, 85(6), 1249-1313.

Citron, D. K. (2019). *Hate crimes in cyberspace*. Harvard University Press.

Citron, D. K., &amp (2014). The scored society: Due process for automated predictions. *Washington Law Review*, 89(1), 1-33.

Coglianese, C., & Lehr, D. (2017). Regulating by robot: Administrative decision making in the machine-learning era. *Georgetown Law Journal*, 105(5), 1147-1223.

Coglianese, C., & Lehr, D. (2019). Transparency and algorithmic governance. *Administrative Law Review*, 71(1), 1-56.

Council of Europe. (2019). *Declaration by the Committee of Ministers on the manipulative capabilities of algorithmic processes*. Decl(13/02/2019)1.

Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., Kak, A., Mathur, V., McElroy, E., Sánchez, A. N., Raji, D., Rankin, J. L., Richardson, R., Schultz, J., Myers West, S (2019). *AI Now 2019 Report*. AI Now Institute.
Doshi-Velez, F., & Kim, B. (2017) Towards a rigorous science of interpretable machine learning. *arXiv preprint* arXiv:1702.08608.
Dos Santos, T. B., & Ramos, P. H. (2020). Artificial intelligence and data protection in Latin America: Regulatory challenges and opportunities. *International Data Privacy Law*, 10(3), 228-246. https://doi.org/10.1093/idpl/ipaa009
Engstrom, D. F., & Gelbach, J. B. (2020). Legal tech, civil procedure, and the future of adversarialism. *University of Pennsylvania Law Review*, 169(6), 1001-1094.
Engstrom, D. F., Ho, D. E., Sharkey, C. M., & Cuéllar, M.-F. (2020). *Government by algorithm: Artificial intelligence in federal administrative agencies*. Stanford Administrative Law Lab.
Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
European Commission. (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence*. COM(2021) 206 final.
European Parliament and Council. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union*, L119/1.
Fallon, R. H., Meltzer, D. J., & Shapiro, D. L. (2015) *Hart and Wechsler's the federal courts and the federal system* (7th ed.). Foundation Press.
Ferguson, A. G. (2017). *The rise of big data policing: Surveillance, race, and the future of law enforcement*. New York.
Freeman, J., & Rossi, J. (2012). Agency coordination in shared regulatory space. *Harvard Law Review*, 125(5), 1131-1211.
Garvie, C., Bedoya, A., & Frankle, J. (2016). *The perpetual line-up: Unregulated police face recognition in America*. Georgetown Law Center on Privacy & Technology.
Georgetown Law Center on Privacy & Technology. (2021). *Banning face surveillance*.
Goel, S., Rao, J. M., & Shroff, R. (2016). Precinct or prejudice? Understanding racial disparities in New York City's stop-and-frisk policy. *Annals of Applied Statistics*, 10(1), 365-394. https://doi.org/10.1214/15-AOAS897
Green, B. (2021). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, 45, 105681. https://doi.org/10.1016/j.clsr.2021.105681
Green, B., & Chen, Y. (2019). Disparate interactions: An algorithm-in-the-loop analysis of fairness in risk assessments. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 90-99. https://doi.org/10.1145/3287560.3287563
Green, B., & Viljoen, S. (2020). Algorithmic realism: Expanding the boundaries of algorithmic thought. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 19-31. https://doi.org/10.1145/3351095.3372840
Hart, H. L. A., & Honoré, T. (1985). *Causation in the law* (2nd ed.). Oxford University Press.
Hellman, D. (2020). Measuring algorithmic fairness. *Virginia Law Review*, 106(4), 811-866.
Hildebrandt, M. (2008). Defining profiling: A new type of knowledge? In M. Hildebrandt & S. Gutwirth (Eds.), *Profiling the European citizen: Cross-disciplinary perspectives* (pp. 17-45). Springer. https://doi.org/10.1007/978-1-4020-6914-7_2
Hirsch, D. D. (2020). The glass house effect: Big Tech, big data, and the future of privacy regulation. *Vanderbilt Journal of Entertainment & Technology Law*, 22(3), 775-823.
International Law Commission. (2001). *Draft articles on responsibility of states for internationally wrongful acts*. United Nations General Assembly. A/RES/56/83.
Kaminski, M. E. (2019). The right to explanation, explained. *Berkeley Technology Law Journal*, 34(1), 189-218.
Kaminski, M. E., & Malgieri, G. (2020). Algorithmic impact assessments under the GDPR: Producing multi-layered explanations. *International Data Privacy Law*, 11(2), 125-144. https://doi.org/10.1093/idpl/ipaa020
Katyal, S. K. (2019). Private accountability in the age of artificial intelligence. *UCLA Law Review*, 66(1), 54-141.
Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., & Mullainathan, S. (2018). Human decisions and machine predictions. *Quarterly Journal of Economics*, 133(1), 237-293. https://doi.org/10.1093/qje/qjx032
Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review*, 165(3), 633-705.
Lipton, Z. C. (2018). The mythos of model interpretability. *Queue*, 16(3), 31-57. https://doi.org/10.1145/3236386.3241340
Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, 13(5), 14-19.
Metcalf, J., Moss, E., & boyd, d. (2021). Algorithmic impact assessments and accountability: The co-construction of impacts. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 735-746.
Molnar, P., & Gill, L. (2018). *Bots at the gate: A human rights analysis of automated decision-making in Canada's immigration and refugee system*. International Human Rights Program and Citizen Lab.

Mulligan, D. K., & Bamberger, K. A. (2019). Procurement as policy: Administrative process for machine learning. *Berkeley Technology Law Journal*, 34(3), 773-803. https://doi.org/10.15779/Z38C24QP4D
Nissenbaum, H. (1996). Accountability in a computerized society. *Science and Engineering Ethics*, 2(1), 25-42.
Nyst, C., & Monaco, N. (2018). *State-sponsored trolling: How governments are deploying disinformation as part of broader digital harassment campaigns*. Institute for the Future.
Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.
Perry, W. L., McInnis, B., Price, C. C., Smith, S. C., & Hollywood, J. S. (2013). *Predictive policing: The role of crime forecasting in law enforcement operations*. RAND Corporation.
Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J.-F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A., ... Wellman, M. (2019). Machine behaviour. *Nature*, 568, 477-486.
Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429-435.
Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33-44. https://doi.org/10.1145/3351095.3372873
Ranchordás, S. (2020). Experimental regulations and regulatory sandboxes: Law without order? *University of Richmond Law Review*, 55(1), 227-288.
Raso, F. A., Hilligoss, H., Krishnamurthy, V., Bavitz, C., & Kim, L. (2018). *Artificial intelligence & human rights: Opportunities & risks*. Berkman Klein Center for Internet & Society.
Reisman, D., Schultz, J., Crawford, K., & Whittaker, M. (2018). *Algorithmic impact assessments: A practical framework for public agency accountability*. AI Now Institute.
Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review Online*, 94, 192-233.
Rose, J. B. (2015). Toward a critical administrative law. *University of Pennsylvania Law Review*, 72(2), 829-932.
Selbst, A. D. (2017). Disparate impact in big data policing. *Georgia Law Review*, 52(1), 109-195.
Selbst, A. D. (2021). An institutional view of algorithmic impact assessments. *Harvard Journal of Law & Technology*, 35(1), 117-186.
Selbst, A. D., & Barocas, S. (2018). The intuitive appeal of explainable machines. *Fordham Law Review*, 87(3), 1085-1139.
Selbst, A. D., &amp (2017). Meaningful information and the right to explanation. *International Data Privacy Law*, 7(4) 233-242.
Shavell, S. (1987). *Economic analysis of accident law*. Harvard University Press.
Slobogin, C. (2019). The automation of criminal justice. *Virginia Journal of Criminal Law*, 7(1), 138-192.
Solove, D. J., & Hartzog, W. (2022). The FTC and the new common law of privacy. *Columbia Law Review*, 114(3), 583-676.
Steele, J. (2020). *Tort law: Text, cases, and materials* (4th ed.). Oxford University Press.
Sweeney, L. (2013). Discrimination in online ad delivery. *Communications of the ACM*, 56(5), 44-54.
Treasury Board of Canada Secretariat. (2019). *Directive on automated decision-making*. Government of Canada.
UN Human Rights Committee. (1992). *General Comment No. 20: Article 7 (Prohibition of torture, or other cruel, inhuman or degrading treatment or punishment)*. HRI/GEN/1/Rev.9.
UN Human Rights Committee. (2004). *General Comment No. 31: The nature of the general legal obligation imposed on States Parties to the Covenant*. CCPR/C/21/Rev.1/Add.13.
UN Human Rights Committee. (2018). *General Comment No. 36 on article 6 of the International Covenant on Civil and Political Rights, on the right to life*. CCPR/C/GC/36.
UN Special Rapporteur on Extreme Poverty. (2019). *Digital technology, social protection and human rights*. United Nations General Assembly. A/74/493.
United Nations. (1966). *International Covenant on Civil and Political Rights*. Treaty Series, vol. 999, p. 171.
Veale, M., & Borgesius, F. Z. (2021). Demystifying the Draft EU Artificial Intelligence Act. *Computer Law Review International*, 22(4), 97-112. https://doi.org/10.9785/cri-2021-220402
Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76-99. https://doi.org/10.1093/idpl/ipx005
Wal-Mart Stores, Inc. v. Dukes, 564 U.S. 338 (2011).
Wexler, R. (2018). Life, liberty, and trade secrets: Intellectual property in the criminal justice system. *Stanford Law Review*, 70(5), 1343-1429.
Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance*, 12(4), 505-523.
Zarsky, T. Z. (2016). The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science, Technology, & Human Values*, 41(1), 118-132.
Završnik, A. (2020). Criminal justice, artificial intelligence systems, and human rights. *ERA Forum*, 20, 567-583.