

Course Name – Artificial Intelligence, Law and Justice  
Professor Name – Dr. Krishna Ravi Srinivas  
Department Name – Center of Excellence in Artificial Intelligence and Law  
Institute Name – NALSAR University of Law  
Week – 07  
Lecture – 33



# Artificial Intelligence, Law and Justice

## Session 33

### AI and Human Rights

Dr. Krishna Ravi Srinivas  
Adjunct Professor of Law &  
Director, Center of Excellence in Artificial Intelligence and  
Law  
NALSAR University of Law



Artificial Intelligence, Law, and Justice Course Session 33: AI and Human Rights.



## Recap

- In the last session we discussed how AI will impact the roles and functioning of Judges, and, looked at whether current guidelines are enough
- We discussed issues in 'human in the loop' concept's application to Judiciary.
- Moreover the need for capacity building and more use of EXAI in Judiciary was emphasized.



As usual, we will do a recap of the previous session. In the last session, we discussed how AI will impact the roles and functioning of judges and looked for whether the current deadlines are adequate enough. Further, we discussed the whole idea of humans in the loop, particularly when we applied this to the judiciary. Moreover, we stressed the need

for capacity building and more use of explainable AI in the judiciary. We reviewed how AI is likely to impact the functioning of the judges and what sorts of new skills and levels of capacity building are required among the judges to come to grips with the developments in AI. In this session, we will be dealing with an important topic on AI, which is AI and human rights, although this topic is not much discussed in the literature or often discussed as a specialized topic. This is equally important because we can easily connect previous sessions, such as explainable AI, algorithmic accountability, and some of the issues related to algorithmic decision-making and how AI-made decisions are affecting the rights and interests of people.



## Key Human Rights and Ethical Challenges of AI



- **Deliberate Use of AI for Suppression**
  - Mass surveillance of minorities, e.g., Uyghur in China
  - Limiting freedom of expression and assembly
  - Monitoring public compliance with behavioral rules
- **Collateral Consequences of AI Operation**
  - Embedding and exaggerating bias and discrimination
  - Invasion of privacy and reduction of personal autonomy
  - Potential harm in healthcare and welfare decisions
- **Detrimental Outputs of AI**
  - Manipulating audience views and violating freedom of thought
  - Prioritizing content that incites hatred and violence
- **Exacerbation of Social Divides**



So, to begin with, what are the major issues in human rights and the ethical challenges of AI? First of all, AI can be used for deliberate misuse, to suppress, in the sense of mass surveillance, limiting freedom of expression, and monitoring public compliance with behavioural rules. Because AI is a technology that can be implemented as part of many surveillance technologies, including facial recognition, gathering information about various people's social media usage, and their behaviour online through different tools, as well as through spyware and other technologies.

Moreover, the collateral consequence of AI's operations is that AI can exaggerate the bias and discrimination, and can create them anew. The question of privacy relating to AI, particularly the question of personal privacy, is also equally important. So, AI can harm personal privacy or can directly or indirectly restrain personal privacy. And then, of course, in the case of certain things like health care and welfare, although AI per se may not be the one that should be blamed, AI-based systems can impact human rights and the entitlements of many people. Then the different outputs of AI can be manufactured or manipulated to suit audience views, in the sense that they can provide disinformation, misinformation, and create deepfakes, thereby ensuring that individuals do not fully exercise their freedom of expression and violate their freedom of thought. So, when AI can be used to prioritize content that includes hate and incites violence, these are some of the human rights issues that we need to come to an understanding of as to why AI is something like a double-edged sword.



## Governing AI: Why Human Rights?



- Human rights overlooked
  - Issues and challenges in AI governance
- Myths about human rights
  - Common misconceptions
- What human rights have to offer
  - Benefits and protections provided by human rights



So, what is happening is that often the governance of AI principles like transparency and accountability is mentioned, and even ethics is also brought in as a key theme. But then human rights are not given that much importance. So, there is a need to look into the issues and challenges of AI governance from a human rights lens. Then there are many myths about human rights in the sense that human rights are often discussed only by specific people; human rights should not be something that is really taken into account because there are other things, like explainable AI and responsible AI, which can take care of all those issues. And then there are myths about human rights themselves, that human rights are not necessary for the well-being of society or that they could be misused, or that those who talk of human rights often end up supporting causes that are not lawfully correct. But human rights have to offer a lot to AI because, as a fundamental right, human rights cover a whole lot of things. And in the Indian context, the Indian constitution protects and guarantees human rights through various articles in the constitution. So, if we look at the governing of AI from a human rights perspective, we would find that there are a lot of things where human rights concern about AI can be expressed and explored.

# Principles of AI Governance: Human Rights Contribution



- **Three Dimensions of AI Governance**
  - Substantive standards for AI developers and implementers
  - Processes to ensure standards are met
  - Accountability and remedies for breaches
- **Principles: Human Rights Law**
  - Integration of human rights law into AI governance principles



There are three dimensions of AI governance that we need to take into account when we talk about AI. First, there should be substantive standards for AI development implementers to ensure that these standards, which meet the minimum human rights standards, are met. Then, accountabilities and remedies for breaches of human rights should also be incorporated into that. So, the first idea would be to integrate or bring in human rights law or at least the human rights articles into the AI governance principles, which means going beyond transparency and accountability and bringing fundamental core human rights as core governance values in AI.

# Ethical values in AI systems



- **AI Ethical Risk Assessment Processes**
  - Involves developers, providers, and users
  - Focuses on refining ethical values and assessing AI products
  - Identifies and mitigates risks
- **Data Governance Tools**
  - Includes data sheets or 'nutrition labels'
  - Summarizes characteristics and intended uses of data sets
- **Government Impact Assessments**
  - Canada's Directive on Automated Decision-Making
  - US's draft Algorithmic Accountability Act
- **Challenges in Ethical Risk Assessment**
- **Audit Process Challenges**



The AI system should be assessed in terms of AI ethics. AI ethical risk assessment should also include human rights values. Particularly when it comes to data governance, we should consider the privacy angle. We should examine some of the other related aspects,

such as whether data governance norms can be used or misused not only to violate privacy but also to spread false information. Automated decision-making, when assessed through a human rights lens, will make certain things very clear to us. For example, the Canadian directive on automated decision-making brings a human rights perspective to that. Similarly, in the USA, there's a draft act on algorithmic accountability that has a human rights dimension. So, the ethical risk assessment of AI can benefit greatly if human rights are also made a part of that.



## Prohibition of certain AI forms

- **Inconsistent Prohibitions**
  - Governments and companies are prohibiting AI forms with serious ethical concerns
  - Lack of consistency in prohibitions
  - Rationale behind prohibitions often not openly acknowledged
- **Examples of Prohibitions**
  - US states banning certain uses of facial recognition technology
  - EU's Artificial Intelligence Act prohibiting manipulative AI practices and biometric identification systems in public spaces
  - Twitter banning political advertising in 2019



On the other hand, it is necessary, if not downright essential, to claim that some forms of AI or some systems of AI with specific innovations should not be developed at all or should be prohibited. For example, AI systems that can raise serious ethical concerns in the sense that they can spread malware, tamper with other systems, or be used for malicious purposes should be banned; however, the prohibition lacks consistency due to a lack of measurement, and there are many dual-use concerns with AI because it can be used for good purposes as well as for bad ones. Dual use is present when we know that certain dangers can occur due to the use of certain technologies, and the rationale behind specific provisions is not explicitly acknowledged; however, it should be made transparent. For example, in some US states, facial recognition technology has been banned, if not fully, then at least some of its uses. Then the EU AI Act has a detailed explanation on using AI for manipulating human behaviour, human consumers, and, more importantly, biometric identification is heavily regulated through the EU AI Act. Of course, the EU AI Act also does a whole lot of risk analysis and looks into various potential harms that could arise on account of AI. And more importantly, it has a human rights dimension built into it. So, some sort of prohibition of certain AI forms. For example, in recent days, there have been a lot of demands to ban deepfakes and AI systems or AI-related misinformation and disinformation.



# Transparency measures in AI



- **Public Transparency Measures**
  - Registries and release of source code or algorithmic logic
  - Required in France under the Digital Republic Law
- **UK Government's Algorithmic Transparency Standard**
  - Launched in November 2021
  - Public sector organizations provide information on their use of algorithmic tools
  - Information published online in a standardized format
  - Several government algorithms made public



But then these are not only legally challenging, but there are also a lot of definitional issues as to how we define them. And more importantly, are there any available technical means that could be used to prohibit them right from the beginning? But that again is a very ticklish issue because defining them itself is a big problem. Technical issues can be overcome. But those with malicious intentions often find ways and means to truly overcome some of these barriers. Then the need for transparency measures in AI is what we have talked about again and again. But algorithmic audits should also look into the potential human rights violations that can happen when algorithmic decision-making is unchecked. The Digital Republic Act of France, for example, mentions maintaining registers and the release of source code or algorithmic routing as a public transparency measure. The UK government launched an algorithmic transparency standard in November 2021. These when implemented also take into account some of the human rights concerns that arise when AI systems are deployed publicly.



## Processes: human rights law



- **Government Duty to Comply with Human Rights**
  - Ensure AI usage in public decision-making respects human rights
- **Protect Individuals from Human Rights Abuses**
  - Prevent abuses by companies and non-state actors
- **State Obligations**
  - Take appropriate steps to prevent abuse
  - Investigate and punish human rights violations
  - Redress abuses through effective policies and regulations



The government, by default and by its own statute, and then by its own constitutional undertaking, is bound to honour human rights. And there can be no exceptions to AI on this ground. So, governments have to ensure that AI usage in public decision-making respects and honours human rights and does not violate them either in part or in full. Moreover, it is equally important that AI does not abuse human rights or promote human rights abuses. So, it is important that non-state actors and private companies do not use AI for the purpose of human rights abuse. Then there are state obligations under various constitutional provisions, as well as the state's obligations under different international governance treaties and agreements, to protect human rights. So those should be fully adhered to when we talk about AI systems.



## Governmental duty to protect



- **Appropriate Mix of Laws, Policies, and Incentives**
  - National and international measures
  - Mandatory and voluntary measures
- **Fostering Business Respect for Human Rights**
  - Requiring suitable corporate structures
  - Identifying and addressing human rights risks
  - Engaging with external stakeholders
- **Additional Steps for State-Owned or Public Sector Businesses**
  - Management control
  - Contractual control



Now the government's duty to protect can take different forms. It can come in nationally

mandatory institutional forms. It can be achieved by setting up specific institutions like human rights commissions, human rights embassies, or human rights agencies to monitor and then reach a conclusion by examining specific cases when complaints are brought to them. Some of these measures can be voluntary, while others can be mandatory. And then there is also a respect for human rights in business. Often businesses have some compact or companies come up with voluntary guidelines to respect human rights, partially on account of the initiative by the UN to bring together business entities or corporations as well as human rights activists to come to an understanding of how businesses can adopt human rights and thereby contribute to societal good. But then the state can do much more than that, particularly when it brings AI systems into public use or when it deploys AI for larger purposes.

The slide features a header with the NPTEL logo on the left and the NALSAR logo on the right. The title "Compliance with human rights" is centered in red. Below the title is a bulleted list of points. To the right of the list is an illustration of a scale of justice with two figures. At the bottom right is a video inset of a man in a blue shirt speaking.

## Compliance with human rights

- **Governments' Obligations**
  - Cannot wait to see how AI develops
  - Must engage in proactive governance activities
- **Actions Required**
  - Regulation and impact assessments
  - Audits to ensure AI does not infringe human rights
- **Understanding Human Rights Implications**
  - Deploy dedicated capacity-building efforts
  - Establish technology and human rights office

There are two things about which we need to be very clear. The obligations of the states under the different conventions, treaties, covenants, or protocols do not hinge upon how AI is going to develop or how AI has developed. It is independent of that, and the government should not wait and see how AI will harm human rights. Rather, they should take proactive steps to ensure that harms are eliminated or, if not fully eliminated, minimized. The second point here is that compliance with human rights and promoting innovation in AI should not be seen as opposites. They should go hand in hand, or only innovations that promote and respect human rights should be given importance. So, the governance perspective in human rights when it comes to AI should involve proactive engagement rather than reactive measures or something that happens after some damage has been done.

So, we would need a lot of actions from the governments, either through impact assessments or by setting up specific mechanisms and institutions to probe potential and actual human rights violations, and also to do so on a continuous basis; capacity building here is very important. And more importantly, technology should be audited for its potential human rights violations or for potential human rights harm. It is also possible that technology can be used to promote, harness, and fully utilize human rights. So, when it comes to compliance with human rights, the government should also give importance

to, promote, and incentivize the development of technologies that can further the cause of human rights.



## Smart mix of laws and policies

- **Need for Effective Regulation**
  - Ensure companies do not infringe human rights
  - Provide effective remedies for infringements
- **Limitations of Voluntary Approaches**
  - Ambiguity of ethical commitments
  - Strong commercial considerations
- **Obligation of States**
  - Enact legally binding norms
  - Protect human rights against AI challenges
- **Regulation of AI Applications**
  - Prohibit or constrain risky AI applications
  - Focus on biometric technologies



So, what do we need? We need a smart, good mix of laws and policies to promote human rights and ensure that AI does not end up violating or negating human rights. So, it should start with companies. It can be voluntary code, or it could be something like companies coming together, or it could be through different legal means, different acts, or binding commitments. Some of the ethical commitments should be very clear, with no ambiguity in them, or they should be very transparent in the sense that they are clearly binding, and there should be no way of treating, reading, or misreading them in two or three more ways to suggest that this does not mean that this is true, or an ambiguity in that should be avoided. How should the state take this into account? As we said, they can enact specific acts to deal with AI-related human rights violations and harms and establish specific mechanisms or integrate them into the current mechanisms. But states are also obligated, particularly under the EU AI Act, to prohibit certain risky AI applications because they could violate human rights. Similarly, different states and different countries will have governance norms to regulate certain risky technologies or innovations. Some specific applications, like biometric technologies or the use of genetic evidence, or some of the technologies that could be used for surveillance, when viewed from a human rights perspective, could be either negated in the sense that they are banned or very strictly regulated in the sense that their use is allowed only when there are specific contexts and when there is a specific need to use them. Often the courts decide under what context some of these tools can be used and then who will be the one supervising it.



## Urgent need for regulation



- **Systematic AIA and Audit Processes**
  - Employ rigorous standards and due process
  - Consider potential human rights impacts
- **Human Rights Risk Assessment**
  - Make assessment of human rights risks explicit
- **Incentivizing Corporate Good Practice**
  - Demonstrate respect for human rights
  - Facilitate remedy
- **Public Reporting Requirements**
  - Require companies to report on due diligence
  - Report on human rights impacts identified and addressed



So, to begin with, there's an urgent need to integrate human rights into regulation. Regulations, rights, standards, and due process should be brought in first, and then a human rights audit or impact assessment of human rights should be conducted to address AI innovations where there is a good potential for violation. Then the assessment of human rights should be made very explicit. This is not something that should be done stealthily or in hiding. It should be something very concrete and very prominent. And then also incentivize corporates through good practices, give them some incentives, and then show that they demonstrate full respect for human rights in letter and spirit. Public reporting on algorithms and AI innovation should be made mandatory. For example, as we said, the French Digital Act also mandates certain measures like these. Companies, as part of their due diligence report, should mention whether the AI systems they bought, acquired, or put into place are fully compliant with human rights or violate human rights. And more importantly, human rights impacts should be assessed, addressed, and then ensured that human rights do not negatively impact the innovations that are permissible.



## Systematic AIA and audit processes



- **Importance of Supervision**
  - Conducted by regulatory and administrative authorities
  - Ensures accountability for compliance with human rights responsibilities
- **Legal Liability**
  - Parallel to supervision
  - Addresses harms and violations



The systematic AI audit and audit process should be made part of the regular AI audit. To make these things happen, there should be clear-cut legal liability and responsibility for them. And then they should be supervised by an institution either set up by the state, by the higher judiciary, or as a combination, or as an autonomous organization. So, the systematic AI audit, or auditing of algorithms, and then the audit processes should be an ongoing, regular routine practice.



## Supervision and accountability



- **Importance of Supervision by Authorities**
  - Ensures accountability for human rights responsibilities
  - Works alongside legal liability for harms
- **Implementation in Europe and EU**
  - Mandatory human rights and environmental due diligence for larger businesses
- **Role of Human Rights Experts**
  - Exploring administrative supervision of corporate duties
  - Complementing liability for harms in courts



Then, of course, supervision and accountability should be built in as a process of that.



## Government procurement obligations



- **Legal Obligations**
  - Governments must not breach human rights with AI systems
- **Knowledge and Information**
  - Understanding AI's capacity and implications is crucial
  - Ensuring AI meets standards on equality, privacy, and other rights
- **Public-Private Contracts**
  - Negotiating terms to ensure AI aligns with human rights
  - Deploying procurement conditions for compliance
- **Encouraging Improvements**
  - Public procurement can enhance human rights standards in AI industry
- **Compliance of Adopted Systems**
  - Ensuring existing AI systems comply with human rights standards



Then, more importantly, often what happens is that governments are the major buyers of the AI systems in the market; they also commission AI systems through innovation, in the sense that they are the ones who buy and the ones who commission. So, when the government procurement obligations are there, the government should ensure that an AI human rights audit is made part of that. So, the implications of AI systems in terms of human rights should be made upfront and then made available. And then AI meets the minimum standards of human rights in terms of equality, non-violation of privacy, and other rights, that should be mentioned. Or in other words, the government should not pursue AI systems that prima facie and upfront violate equality and then privacy and other fundamental rights. We have also seen this earlier when we talked about transparency and accountability, more from an algorithmic responsibility perspective. And then, when governments enter into public partnerships, private partnerships, or similar initiatives, they should ensure that wherever AI is deployed, human rights are not violated. More importantly, governments should ensure that public procurement systems enhance human rights standards in the AI industry or incorporate the minimum human rights standards when they are developed. So, the compliance of the adopted system should be regularly reviewed, and there should be an annual review of various systems where there is a potential for harm or abuse of human rights.



## UN Guiding Principles on Business



- **Respect for Human Rights**
  - Companies should avoid infringing on human rights
  - Address any adverse human rights impacts from their activities
- **Policy Commitment**
  - Approved at senior level
  - Publicly available
  - Embedded in the business culture
- **Due Diligence Process**
  - Ongoing human rights impact assessment
  - Tracked for responsiveness
  - Reported externally
- **Benefits of Responsible Business Agenda**



The UN has come up with a compact for business and human rights. This compact has got some guiding principles: respect for human rights, policy commitment, due diligence process, and benefits of a responsible business agenda. So, when private businesses adhere to this compact or adhere to the UN norms, they should ensure that AI does not result in human rights violations.



## Challenges in AI due diligence



- **Distinguishing Features of AI**
  - AI's capacity for self-improvement makes predicting consequences difficult
  - Human rights impact depends on technology and deployment context
- **Need for Extensive Due Diligence**
  - Involves a wide set of stakeholders
  - Must be extensive to address human rights impacts
- **Regular Review of AI Systems**
  - AI must be reviewed regularly once in operation
  - Comprehensive due diligence throughout AI system's life cycle
- **Current Gaps in Company Structures**
  - Many companies lack processes to detect and act on human rights issues
  - Results of due diligence should be made public



But then there are a whole lot of issues in the sense that these are easy to talk about and then easy to even say that this should be done, that should be done. This is what you should be mandated to do. But in real practice, translating them into actionable points or actionable agenda items is not all that simple. Why? AI is not a dumb system. AI is not a blind system. AI systems often have the capacity to learn, be flexible, adapt, and adopt. So, when they have the capacity for self-improvement, learning, and then putting that into

practice in a loop, it is very difficult to predict what exactly the consequences will be, however much the initial assessments would have said that no violation of human rights would be possible. This is all the truer when we talk about bias, discrimination, and the violation of equality. Then identifying the specific human rights violation also depends on the type of technology used, the sort of users being envisaged, and the kinds of uses the system is actually put into practice. So, the excessive due diligence should be part of that, and this should involve a wide range of stakeholders.

This is not something where the company is given a form to fill out and then tick the boxes; it should be much more serious and sensitive to that. More importantly, the regular review of AI systems, however cumbersome and time-consuming it is, should be done from a human rights perspective. So, the problem is also there because not many companies, even today, have a regular AI governance structure in place where some of these things can be incorporated to fit well within the AI governance framework. On the other hand, due diligence often ignores some of the critical AI-related matters in the sense that what the stock market authorities need, what the government authorities need, or what different regulators need often will not have the AI human rights impact assessment as part of that. So, some of these gaps should be assessed by the regulators as well as by the government.

**NPTEL**

## Examples of corporate AIAs

**NALSAR**

- **AIAs Labelled as Human Rights Assessment**
  - Verizon's ongoing human rights due diligence
- **AI Ethics Assessments Similar to Human Rights Due Diligence**
  - IEEE's adopted AI ethics assessment
- **Google's AI Deployment Review Process**
  - References AI Principles
  - Includes consultation with human rights experts



Then there are lots and lots of good corporate AI audits or AI assessments, known as AIAs. For example, AIAs are labelled as human rights assessments as well. Verizon has an ongoing human rights due diligence process where it undertakes this on a voluntary basis. IEEE's adopted AI ethics assessment has incorporated human rights due diligence as part of it. Similarly, Google's deployment of AI reinforces AI principles, references, and includes consultations with human rights experts. These are just some examples because many of the bigger corporates, particularly the multinationals, are very sensitive to accusations of human rights violations, particularly bias, discrimination, and gender discrimination. So, they are trying to ensure that their corporate names are not spoiled or their corporate image is not damaged on account of something going wrong when AI

violates human rights, so they're on their own incorporating or building tools that facilitate corporate AI audits.



## Fostering a pro-human rights culture



- **General Corporate Statements**
  - Intentions and activities are publicly available
- **Human Rights Risks**
  - Identification and mitigation through due diligence
  - Less public information available



Generally, a pro-human rights culture should be adopted because treating human rights as just another item to be considered or something that is not of much importance to the business is not the right attitude. So, the due diligence process should incorporate the human rights angle, and then the intentions and activities related to human rights should be publicly available; companies themselves should also make this information available to their own employees, because nobody should think that this is something that can be ignored.



## Remedies in AI governance



- **Scope of Corporate Processes**
  - Focus on specific issues like bias and privacy
  - Brief mention of other human rights
- **Effect of Impact Assessments**
  - Unclear effects on company activities
  - Human rights risks need mitigation
  - Business processes balance risks and benefits
- **Access to Remedy for AI Failures**
  - Effective reparation and accountability
  - Measures to prevent recurrences



So, as we said, there are multiple remedies available in AI governance; whether it is bias or privacy, specific issues need specific action points. Of course, there are a lot of laws

that cover them, but there could be gaps in that. An AI human rights perspective, taking a holistic perspective will be able to address them and fill those gaps. Then human rights are not one or two; core human rights defined in the constitution could be important, but there are other human rights that governments would like to emphasize and enforce, so they should be brought into the picture. But often what happens is impact assessments are done as routine matters; what they state in them is not taken further to revise the AI systems, and then they remain as yet another box-ticking exercise, so this practice should stop. And then the human rights risks need mitigation strategies, but impact assessments do not advise any mitigation strategies or recommend any mitigation strategies; they stop short of that.

They simply say that there is a potential human rights violation or that there are many human rights violations. Then the business practices balance risk and benefit. Often, human rights are not given that importance. So, compared to the benefits, this is a risk that we can afford to ignore, or we need not pay undue adherence or attention to it. That's the normal perspective. This should undergo change. More importantly, what happens when AI violates human rights, particularly due to failures of AI to do something positive, or when, by default, the system does something that is totally unanticipated in nature? If that unanticipated consequence occurs, what should be done, there should be clear-cut corporate policy on these matters. So how do we prevent them further should also be part of the corporate larger strategy on AI and human rights?

The slide features a title 'Finland Decision' in red. On the left is the NPTEL logo, and on the right is the NALSAR logo. The main content is a bulleted list:

- **Tribunal's Decision in March 2018**
  - Finland's National Non-Discrimination and Equality Tribunal ruled against a credit institution
- **Reason for the Decision**
  - Credit decision was based on assumptions from statistical data
  - Criteria included gender, first language, age, and residential area
- **Outcome**
  - Tribunal found the decision discriminatory
  - Prohibited the use of such decision-making methods

To the right of the text is a graphic titled 'FINLAND MAP' showing a map of Finland with a flag and a bar chart. In the bottom right corner, there is a video inset of a man in a blue shirt speaking.

Now we have seen a lot of things to do, but we will also look into some cases where courts and other institutions have taken the lead in identifying certain things for human rights violations and have dealt with them. Finland's National Non-Discrimination and Equality Tribunal ruled against a credit institution in March 2018 because the credit decision was based on assumptions from statistical data, which included criteria such as gender, first language, age, and residential area. These discriminations have already been discussed in detail in previous classes. So, the tribunal said discrimination was present. That the decision was discriminatory, and it stated that such decision-making tools or methods should not be used in the future.



## Practical actions for companies



- **Hague District Court's Decision**
  - Ordered Dutch government to stop using SyRI in February 2020
  - SyRI reviewed personal data to predict benefit or tax fraud
- **Lack of Transparency**
  - Government refused to reveal how SyRI used personal data
  - Difficult for individuals to challenge investigations or risk scores
- **Violation of Privacy Rights**
  - Legislation regulating SyRI did not comply with Article 8 ECHR
  - Failed to balance societal benefits with privacy violations
- **Discriminatory Practices**
  - SyRI used only in 'problem neighbourhoods'
  - Proxy for discrimination based on socio-economic background and immigration status



In the Hague, there was a system that was used in Siri by the government, and the idea was to review personal data to predict benefit or loan fraud and tax fraud. But then the government did not share adequate details about how it was used. So, individuals who were impacted by that found it very difficult to challenge it. They felt their privacy rights had been violated by Siri and that it was not in line with Article 8 of the European Charter for Human Rights. So they went to the court, and then the government said Siri was used only in specific neighbourhoods that were classified as problem neighbourhoods, and the proxy for discrimination was based on socioeconomic background and immigration status. So, they used certain things as a proxy, and this was found to be totally wrong.



## South Wales Case



- **Case Overview**
  - First challenge to AI invoking UK human rights law
  - South Wales Police trialled live automated facial recognition technology (AFR)
- **AFR Technology**
  - Compared CCTV images of public event attendees with a database
  - Non-matching images were immediately deleted
- **Legal Challenge**
  - Complainant challenged AFR's capture and comparison of his image
  - Referenced Article 8 ECHR and UK Data Protection Act



In South Wales, police trial automated facial technology that was put into automatic

operation in the sense that as you move, walk, run, or drive a vehicle, your face will be automatically tracked, and then they compared the public CCTV event attendance with the database they had earlier. They also looked into the non-matching items: 'oh, this place is not found here', so the complainant challenged that this sort of capturing their facial recognition or images without any prior informed consent or without informing them violated their fundamental rights. So, they referred to Article 8 of the European Convention on Human Rights as well as the UK Data Protection Act, because such capturing of images without prior informed consent and then using that without any prior notice is considered a serious violation.

 **NPTEL**

## South Wales Case



- **Improper Legal Basis for AFR Use**
  - AFR use breached the Data Protection Act
- **Balance Between Individual Rights and Community Interests**
  - Court did not find police use of AFR to strike the wrong balance
- **Failure to Discharge Public Sector Equality Duty**
  - South Wales Police did not ensure AFR software was free from racial or gender bias
  - No evidence of bias in the software
- **Temporary Halt of AFR Use**
  - South Wales Police's use of AFR was temporarily halted
  - Possibility of reintroduction with proper legal footing
- **Reintroduction of AFR**
  - South Wales Police has reintroduced facial recognition technology in certain circumstances



So, what did the court say? The court said this sort of use breached the Data Protection Act. You need to balance between individual rights and community interests in the sense that you cannot simply say that in the larger interest of the community, you can capture each and every image, compare that with the internal database, and then identify which images do not match. And then there was also the thing that the software that was used was not free from racial or gender bias. More importantly, they should have found evidence that there was no bias from the software. So, it was said to halt the temporary use of AFR for the time being, and then they were needed to reintroduce the same thing with proper legal footing and then with a differently trained set or modified software accordingly. So, in some circumstances, they reintroduced the revised facial recognition technology.



## Promote AI ethics



- **Case Overview**
  - First challenge to AI invoking UK human rights law
  - South Wales Police trialed live automated facial recognition technology (AFR)
- **AFR Technology**
  - Compared CCTV images of public event attendees with a database
  - Non-matching images were immediately deleted
- **Legal Challenge**
  - Complainant challenged AFR's capture and comparison of his image
  - Referenced Article 8 ECHR and UK Data Protection Act



So, the idea is that AI ethics should be promoted as a routine practice, not as a simple afterthought. This is the first case where the use of AI was challenged under the Human Rights Act. And then, when this was done, it was proven that Article 8 of the ECHR and the UK Data Protection Act were violated.



## Holistic commitment to human rights



- **Improper Legal Basis for AFR Use**
  - AFR use breached the Data Protection Act
- **Balance Between Individual Rights and Community Interests**
  - Court did not find police use of AFR to strike the wrong balance
- **Failure to Discharge Public Sector Equality Duty**
  - South Wales Police did not ensure AFR software was free from racial or gender bias
  - No evidence of bias in the software
- **Temporary Halt of AFR Use**
  - South Wales Police's use of AFR was temporarily halted
  - Possibility of reintroduction with proper legal footing
- **Reintroduction of AFR**
  - South Wales Police has reintroduced facial recognition technology in certain circumstances



So, what we need is a holistic commitment to human rights that ensures software is free from bias and that the respective Conventions on human rights and data protection rights are not violated. And more importantly, such a holistic commitment will also mean engaging with stakeholders in more than one way.



## Recruit human rights expertise



- **Initial Ruling in 2019**
  - Administrative decisions based on algorithms deemed illegitimate
- **Reversal in 2021**
  - Courts recognized the benefits of speed and efficiency
  - Algorithmic decisions must adhere to administrative review principles
- **Principles for Algorithmic Decision-Making**
  - Transparency
  - Effectiveness
  - Proportionality
  - Rationality
  - Non-discrimination
- **Rights of Complainants**



And it is also important that we bring in the human rights expertise for that. While AI systems can benefit more in terms of efficiency and speed, there is a balance that algorithmic decision-making may go wrong, may be biased, or may not be fully free from discrimination. So, algorithmic decision making has led us to administrative review principles which include transparency, effectiveness, proportionality, rationality and non-discrimination. So, the rights of the complainants should ensure that there is a mechanism that addresses their rights, and there should be a specific mechanism available when people find that their rights are being violated.



## Conduct human rights due diligence



- **Complaint Filed by Big Brother Watch**
  - Issued in July 2022
  - Filed to the British information commissioner
- **Alleged Use of Facial Recognition Technology**
  - Involves Facewatch and Southern Co-op
  - Used to scan, maintain, and assess profiles of supermarket visitors
- **Breach of Data Protection and Privacy Rights**



So human rights due diligence should be made a part of this exercise. The Big Brother Watch and NGO filed in July 2022, and then they filed with the British Information Commissioner, who was the Information Commissioner for Right Information. They said

that facial recognition technology had been used in certain circumstances without proper legal vetting or without proper permission. So, it seemed that this breached data protection and privacy rights.

 

## Establish remedy mechanisms

- **Requirement of Remedy in Human Rights Law**
  - Governments and companies must provide suitable remedies
  - Remedies are necessary for breaches of obligations and responsibilities
- **Components of Effective Remedy**
  - Effective reparation for victims
  - Accountability for those responsible
  - Measures to prevent future breaches
- **Significance of Remedy Availability**
  - Ensures human rights and ethical principles have real impact
  - Balances against commercial considerations



So, the effective government mechanism should be built not as an afterthought, but as part of the routine development, initialization, and digitization process. And then there should also be an effective remedy mechanism available so that human rights and ethical principles have a real impact and are not violated.

 

## Standards bodies and human rights

- **Accountability for AI-related Harm**
  - Businesses should pursue accountability against companies causing harm
  - Harm may result from malfunctioning AI systems
  - Interference from another company's AI can also cause harm



And then there is a need for standard bodies that will set the standards for human rights-related AI systems so that they do not violate human rights; more importantly, businesses should also work with public authorities and human rights advocates. The harms that could result from the malfunctioning of AI systems, or from a malfunctioning AI system

without any intention, might have harmed certain people. So, these things should be taken into account. So, the standard setting for AI human rights violations should be made as part of the regular routine AI audit.

 

## Understanding of human rights

- **Understanding the Complaint Process**
  - Complainants need to know how to complain and to whom
  - Confidence that their complaint will be addressed timely
- **Transparency and Explainability**
  - Complainants should understand how decisions about them were made
  - Information on the role and operation of AI in decision-making
- **Access to Data**
  - Details on AI design and testing
  - Information on intended and actual AI operation in specific cases
  - Role of human decision-making or oversight



So, we need to understand human rights. We need to bring in explainability and transparency. We need to put human rights into the code. And then we also need to link that with other data rights, such as access to data, privacy rights, and so on.

 

## Expertise in AI and human rights

- **Remedy Providers**
  - Courts
  - Governmental mechanisms (regulators, ombudspersons, complaints processes)
  - Non-governmental mechanisms (corporate remediation processes)
- **UN Guiding Principles Recommendations**
  - Businesses should establish or participate in effective operational-level grievance mechanisms
- **Characteristics of Effective Mechanisms**
  - Legitimate (enabling trust)
  - Accessible
  - Predictable
  - Equitable
  - Transparent



So, we need to build expertise in human rights at different levels for various purposes. And then there are a good number of recommendations from the UN, as well as from various other national human rights bodies and different civil society actors in this.



## Cross-cutting regulation



- **Expected Challenges in the Field**
  - Numerous challenges anticipated in the coming years
- **Guiding Principle**
  - Provision of an effective right to remedy
  - Includes addressing breaches of human rights responsibilities



But cross-cutting regulation will be required when AI systems are deployed on a massive scale. This is a real challenge, particularly when human rights can be impacted by AI deployment across sectors for different purposes in different contexts, where facial recognition technology, biometric technology, and AI mapping human behaviour on a continuous basis, linked up with datafication and data technologies, mean that human behaviour is monitored on a continuous basis. So, we need to look at the cross-cutting regulation for AI that violates human rights in more than one way.



## Resources for human rights bodies



- **Effective Practical Steps Needed**
  - Companies must prioritize human rights in AI governance
  - Governments should implement regulations to protect human rights
  - International organizations need to set global standards
  - Civil society must advocate for human rights in AI policies
  - Investors should support ethical AI initiatives



So, the human rights bodies also need some capacity building. They also need adequate resources in terms of finance and human resources.



## Incentivize beneficial AI development



- **Importance of Human Rights in AI**
  - AI is reshaping human experience
  - Human rights should be central to AI governance
- **Benefits of Human Rights-Based AI Governance**
  - Nothing to fear from this approach
  - Much to gain by taking human rights as the baseline



And then, benefiting AI development that respects and honours human rights should be made part of it.



## Harmonize international understanding



- **Ignoring Human Rights Undermines Established Norms**
  - Liberty, fairness, and equality are compromised
  - Accountability processes are disregarded
- **Creation of Confusing Alternatives**
  - Inadequate substitutes to existing norms
  - Duplication of efforts in norm development
- **Implementation and Remedy Issues**
  - Processes for norm implementation are duplicated
  - Remedies for breaches are inadequately addressed



And then globally, there should be a human rights understanding that harmonizes such understanding so that minimum standards are set, the relevant articles are interpreted in an unambiguous way, certain good practices are identified and promoted, and more importantly, new norms need to be developed. New risks, with their identified new norms, would be developed into actionable guidelines and then actionable points. If necessary, the existing Human Rights Act needs to be reformed or revised. Or a new act may be made part of those among various acts where there is a good scope for AI Human Rights being controlled.



## Multi-stakeholder forum



- **Promote AI Ethics**
  - Ensure ethical practices in AI development and deployment
- **Responsible Business Agendas**
  - Encourage businesses to adopt responsible practices
- **Complementary Role of Human Rights Frameworks**
  - Acknowledge the importance of existing human rights frameworks



A multi-stakeholder forum or stakeholder engagement on a continuous basis will be a good opportunity to do that.



## UN alignment with human rights



- **Champion holistic commitment to human rights**
  - Ensure top-level organizational support
  - Promote adherence to all human rights standards
- **Enable change in corporate mindset**
  - View human rights as a useful tool
  - Avoid seeing human rights as a constraint on innovation



UN alignment with human rights is something that the UN has been trying to advocate for years, of course with some success. So basically, when we say AI and human rights, we do not talk in terms of negative impacts alone. We talk in terms of proactive ideas, and more importantly, we ensure that AI does not violate human rights either by default or by accident. Taking these things into perspective, there is a strong need for understanding human rights in the context of AI so that we know AI does not end up harming or negating human rights.



# Next

- AI and Legal Education Session 34



In the next class, we will look into AI and legal education; that would be session 34 because AI, as it is impacting other sectors, is also impacting legal education in a huge way, so that will be the focus of the next class. Thank you.