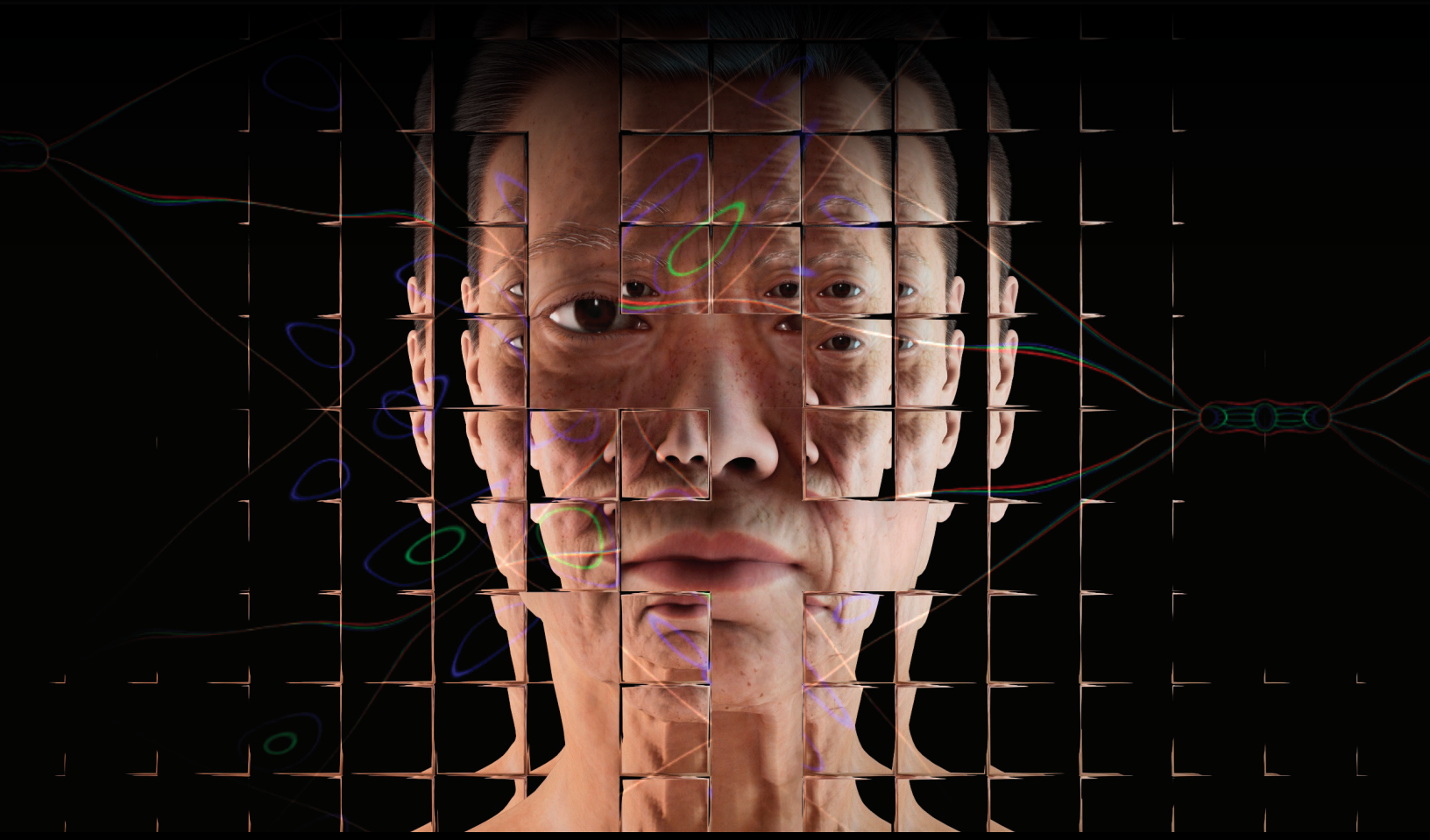
UC BERKELEY

CENTER FOR LONG-TERM CYBERSECURITY



CLTC WHITE PAPER SERIES

AI's Redress Problem

Recommendations to Improve Consumer Protection from Artificial Intelligence

IFEJESU OGUNLEYE

Cover image: Virtual Human, by Alan Warburton, © BBC. CC-BY 4.0. From "Better Images of AI," (https://betterimagesofai.org), a collaboration between BBC Research & Development, We and AI, and the Leverhulme Centre for the Future of Intelligence. The image shows a portrait of a simulated person against a black background, refracted in different ways by a fragmented glass grid. This grid is a visual metaphor for the way that artificial intelligence (AI) and machine learning technologies can be used to simulate and reflect the human experience in unexpected ways. A distorted neural network diagram is overlaid, familiarizing the viewer with the formal architecture of AI systems.

CLTC WHITE PAPER SERIES

AI's Redress Problem

Recommendations to Improve Consumer Protection from Artificial Intelligence

IFEJESU OGUNLEYE

JULY 2022

CENTER FOR LONG-TERM CYBERSECURITY

University of California, Berkeley

Contents

Executive Summary

This paper provides recommendations to regulators, corporations, and civil society organizations to facilitate greater accountability for the use of AI through redress mechanisms, including establishing an AI ombudsman service, an independent body established to investigate and resolve complaints. With AI systems increasingly being deployed across vital sectors such as finance, healthcare, criminal justice, and recruitment, it is important that redress mechanisms are established and maintained to ensure that consumers, data subjects, or users of AI systems have access to a range of effective redress options in the event that they suffer harm.

RECOMMENDATIONS FOR REGULATORS

- Ensure that individuals harmed by the deployment of AI systems are able to make a regulatory complaint or pursue legal action in court.
- Establish a dedicated AI ombudsman service that reviews disputes or complaints between individuals and companies in an independent and impartial manner.
- Empower groups or communities of people who have suffered systemic or widespread harm from the development and/or deployment of AI systems to collectively seek redress for such harms.
- Empower civil society organizations to represent consumers in seeking redress or making general interest complaints against companies using AI systems that are harmful.

RECOMMENDATIONS FOR COMPANIES

- Establish internal ombudsman services to receive and review complaints from stakeholders, including employees or consumers.
- Engage with external stakeholders, such as academic researchers or consumer advocacy groups, to identify and address issues of bias, discrimination, or unfairness that may exist in AI models.

RECOMMENDATIONS FOR CIVIL SOCIETY

- Engage with underserved or marginalized individuals or communities to identify harmful impacts and seek redress.
- Ensure that findings from engagement with communities, audits, or research are made publicly available.

Introduction

The development of artificial intelligence (AI) technologies continues to advance in a wide range of sectors and industries.[1] In tandem with the potential benefits of innovation in this area, there have been numerous discoveries about AI systems being developed and deployed with inherent flaws, causing negative and harmful impacts on various marginalized or vulnerable groups, communities, and individuals.[2] Attempts are underway across the globe to create effective governance mechanisms that minimize the risks associated with AI technologies while encouraging innovation. These attempts include general governance frameworks, such as the European Union's AI Act,[3] as well as the United States' AI Initiative, which seeks to encourage research and development and remove barriers to AI innovation.[4] As several institutions and organizations begin to develop governance and regulatory frameworks for the use of AI technologies in various jurisdictions, a key gap is beginning to emerge in the area of remediation or redress processes in the event of harmful impacts.

For regulations to be truly effective, they have to provide methods of enforcement. The right to enforce provisions of a regulation is typically held by the state through a specific regulatory or criminal agency. However, in instances where the conduct of the regulated entities has a direct negative impact on individuals, private redress mechanisms are often established to allow affected persons to seek redress and obtain remedies. Allowing private redress mechanisms against corporations can help ensure effective enforcement of rules by providing individuals with the right and power to monitor and challenge harmful or negative conduct. It is also important to provide redress opportunities because AI technologies can make the wrong decisions even when they are functioning as designed, and often operate at a much larger scale than human systems, thereby impacting a larger proportion of the population. Algorithmic systems used to determine people's entitlement to benefits,

1 AI is defined by the National Artificial Intelligence Act 2020 as a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine- and human-based inputs to: (a) perceive real and virtual environments; (b) abstract such perceptions into models through analysis in an automated manner; and (c) use model inference to formulate options for information or action.

2 "5 Examples of Biased Artificial Intelligence," *Logically*, last modified July 30, 2019, https://www.logically.ai/articles/5-examples-of-biased-ai

3 European Commission. Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, *COM/2021/206*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

4 National Artificial Intelligence Initiative, Overseeing And Implementing The United States National AI Strategy, *National Artificial Intelligence Initiative Office*, accessed July 4, 2022, https://www.ai.gov/

payments, or other state support have been shown on many occasions to make inaccurate and discriminatory decisions. In France, an investigation by journalists found that the government's adoption of an automated system for social security benefit distribution led to delays and errors in between one to two percent of applications, affecting between 60,000 and 120,000 people.[5]

In the US, AI systems are increasingly used by government entities to manage social and welfare programs, improve efficiency, and detect fraud. However, cases of discrimination, bias, or error that have had severe effects on the subjects or beneficiaries of these systems have come to light in the last few years. In 2013, Michigan adopted a simple algorithmic system to reduce the incidence of fraud in unemployment insurance claims, but deployed it without human oversight.[6] The system increased accusations of fraud against claimants by five times, leading to demands of repayment as well as criminal and civil actions.[7] A review of the system later found that over 90 percent of the allegations of fraud were made incorrectly.[8] However, the error led the state to take action against thousands of citizens, many of whose wages and bank accounts were garnished.[9] Similar algorithmic systems instituted in other states for fraud detection and increased bureaucratic efficiency have suffered similar errors.

Establishing redress mechanisms can provide general accountability by allowing consumers and consumer organizations to submit complaints and challenges to AI systems. Such mechanisms can also empower regulators to identify and collate trends in negative impacts from AI systems. This information can create awareness about common problems among users of AI systems and provide learning opportunities for continued development in the field.[10]

5 "How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers", *Human Rights Watch*, last modified November 10, 2021, https://www.hrw.org/news/2021/11/10/how-eus-flawed-artificial-intelligence-regulation-endangers-social-safety-net#

6 Felton, Ryan, "Criminalizing the Unemployed", Detroit Metro Times, July 1, 2015, https://www.metrotimes.com/detroit/criminalizing-the-unemployed/Content?oid=2353533

7 The Conversation, "States Increasingly Turn to Machine Learning and Algorithms to Detect Fraud", *US News*, February 14, 2020, https://www.usnews.com/news/best-states/articles/2020-02-14/ai-algorithms-intended-to-detect-welfare-fraud-often-punish-the-poor-instead

8 Ibid

9 Ibid

10 In this instance, users of AI systems include parties involved in the development and deployment of AI systems.

There are a variety of redress mechanisms available across sectors and industries, such as private or representative complaints processes to supervisory authorities such as regulators, consumer bodies, tribunals, and judicial review. Although some of these mechanisms have been incorporated in recent AI regulations, including transparency obligations in the EU Artificial Intelligence Act and the Federal Trade Commission AI and Algorithms Guidance, there has been no adoption of extensive mechanisms to provide effective redress for harms caused by deployed AI systems.

Redress Mechanisms in AI Regulations

EUROPEAN UNION ARTIFICIAL INTELLIGENCE ACT

Harms or negative impacts occasioned by the development and/or deployment of AI systems are most often suffered by individuals or communities, as seen in the deployment of criminal justice AI applications. Despite this, the EU AI Act (the "AI Act") does not contemplate or provide private rights of action to individuals or groups negatively impacted by a deployed AI system (whether or not such a system is considered high risk). The failure to establish private rights of action means that individuals are unable to seek specific recourse against companies that have developed or deployed AI systems that have negatively impacted them in some way.

Article 3.5 of the Explanatory Memorandum to the AI Act does contemplate a need for individuals to take redress, but notes that in the event of infringements by AI systems on the fundamental rights enshrined in the EU Charter of Fundamental Rights, effective redress for affected parties will be brought about by requiring transparency and traceability of AI systems and other *ex post* controls.[11] The transparency obligations under the Act are only to the minimum extent necessary for individuals to exercise their rights to effective redress.[12]

This oversight has been criticized by civil society and human rights organizations, which have noted that the failure to confer either individual or collective rights to redress in respect of high-risk AI systems prevents the AI Act from effectively addressing the harms that arise from the deployment of AI systems.[13]

EU DIGITAL SERVICES ACT

The EU Digital Services Act (DSA), set to come into force in 2024, covers narrow uses of algorithmic systems by requiring online platforms that use algorithms for content recommendation, moderation, or advertising to provide clear, understandable, and easily

11 European Commission. Artificial Intelligence Act, Article 3.5

12 The Conversation, "States Increasingly Turn to Machine Learning and Algorithms to Detect Fraud"

13 European Digital Rights et al. An EU Artificial Intelligence Act for Fundamental Rights; A Civil Society Statement. *EDRI*, last modified December, 2021, https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf

accessed information about such algorithms and establish complaint and redress frameworks for customers.[14]

Article 17 of the DSA requires online platforms to establish and provide access to "effective internal complaint-handling" systems that can be accessed electronically and without charge. Such complaint systems should be user-friendly and easy to access, allowing customers to submit substantiated complaints about the removal or disabling of information provided by them on a platform or suspension or removal of their account.[15] In effect, consumers are allowed to provide evidence alongside complaints about the wrongful or mistaken removal of their content or account on online platforms. Platforms are required to address such complaints in a timely manner and provide information to complainants without undue delay.

In addition to internal complaints systems, the DSA also empowers EU member-states to establish "out-of-court" dispute settlement bodies — i.e., ombudsman services — to resolve disputes related to decisions made by online platforms, as well as complaints that remain unresolved after going through internal complaints processes.[16] Such ombudsman services are to be impartial and independent of all parties and capable of settling disputes in an efficient and cost-effective manner.[17]

FTC GUIDANCE

The Federal Trade Commission (FTC) published "Using Artificial Intelligence and Algorithms," which provides guidance on how companies can use AI technologies without contravening FTC regulations such as the Fair Credit Reporting Act and the Equal Credit Opportunity Act.[18] The guidance highlighted four principles in particular:

14 European Commission. Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC. *COM(2020) 825 final*, https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A825%3AFIN

15 European Commission. Digital Services Act, Article 17, Section 3

16 European Commission. Digital Services Act, Article 18, Section 3

17 Ibid

18 Andrew Smith. "Using Artificial Intelligence and Algorithms". *Federal Trade Commission* (blog), April 8, 2020, https://www.ftc.gov/news-events/blogs/business-blog/2020/04/using-artificial-intelligence-algorithms

- *Transparency*, to ensure that consumers are aware of the collection of data and the use of automated tools by a company.
- *Explainability* of algorithmic decisions to consumers, including the principal reasons why a decision was reached, factors that affected the decision and their importance, and any changes to the terms of contracts or deals.
- *Fair decisions*, to prevent discrimination against protected classes and achieve outcomes that do not have disparate impacts on certain groups or people. Fairness also includes providing opportunities for consumers to access information held about them and dispute it if they believe it to be incorrect.
- *Robust and empirically sound data and models,* as results from implementing reasonable procedures to ensure the accuracy of consumer data or information held.

The FTC guidance also notes that any notice given to consumers about information held about them that serves as the basis of algorithm decisions should include the source of the information, and should also advise consumers of their access and dispute rights, which are limited to the right to see the data about them that is reported or held and correct any inaccurate information.

Redress Mechanisms in Data Protection Regulations

Data protection regulations, especially the General Data Protection Regulation (GDPR), contain redress mechanisms that allow individuals to take action personally, or through designated representatives, against organizations that breach their privacy rights or cause harm as a result of their collection, processing, storage, or transmission of personal data. These mechanisms provide an option for seeking redress or compensation for harms caused by AI systems.

GENERAL DATA PROTECTION REGULATION

The General Data Protection Regulation (GDPR) contains various mechanisms for individuals whose data has been collected, processed, or stored in manners not in compliance with the provisions of the regulation to seek redress. This includes:

Articles 77–79: Right to Lodge a Complaint and Effective Judicial Remedy
". . . Every data subject shall have the right to lodge a complaint with a supervisory authority . . . if the data subject considers that the processing of personal data relating to him or her infringes this Regulation."

Article 80: Representation of Data Subjects
"The data subject shall have the right to mandate a not-for-profit body, organization or association . . . to lodge the complaint on his or her behalf."

"Member States may provide that any body, organization or association . . . , independently of a data subject's mandate, has the right to lodge, in that Member State, a complaint with the supervisory authority which is competent . . . if it considers that the rights of a data subject under this Regulation have been infringed as a result of the processing."

Article 82: Right to Compensation
Any person who has suffered material or non-material damage as a result of an infringement of this Regulation shall have the right to receive compensation from the controller or processor for the damage suffered.

These rights provided by the GDPR, particularly the right provided to non-profit organizations to lodge complaints on behalf of data subjects, have led to regulatory actions being taken against companies in various EU States. In 2021, the privacy group "None of Your Business" filed over 400 complaints against companies across the EU for their breach of GDPR obligations in relation to cookie banners on their websites.[19] Another privacy non-profit organization, La Quadrature du Net, filed a joint complaint against Amazon for unlawful use of customer data for advertising purposes in 2018, which ultimately led to the imposition of a fine of almost 750 million euros by the Luxembourg National Commission for Data Protection.[20] While most collective claims brought by consumer advocacy groups under the GDPR have been with the mandate of consumers (i.e., representing particular consumers), the Advocate General of the Court of Justice of the European Union recently published an opinion noting that consumer protection associations are allowed under the GDPR to institute legal action against companies without the authorization of affected consumers where the objective of such action is to protect the rights of consumers.[21] However, the ability to bring such representative action must be provided for by national law in the member state.

CCPA AND CPRA

The California Privacy Rights Act (CPRA) and the soon-to-be-sunsetted California Consumer Privacy Act (CCPA) provide limited rights to consumers to bring legal private action against eligible companies in the event of a data breach for damages or injunctive relief.[22]

Section 1798.150(a) (1) of the CCPA provides that:
"*Any consumer whose nonencrypted or nonredacted personal information is subject to the unauthorized access and exfiltration, theft or disclosure as a result of the business' violation of the duty to implement and maintain reasonable security procedures and practices appropriate to the nature of the information to protect the personal information may institute a civil action. . . .*"

19 *Noyb* files 422 formal GDPR complaints on nerve-wrecking "Cookie Banners". *None of Your Business*, last modified August 10, 2021, https://noyb.eu/en/noyb-files-422-formal-gdpr-complaints-nerve-wrecking-cookie-banners

20 Amazon fined 746 million euros following our collective legal action. *La Quadrature du Net*, last modified July 30, 2021, https://www.laquadrature.net/en/2021/07/30/amazon-fined-746-million-euros-following-our-collective-legal-action/

21 Court of Justice of the European Union. *Advocate General's Opinion in Case C-319/20. Facebook Ireland*. Press Release No 216/21 (2021) https://curia.europa.eu/jcms/upload/docs/application/pdf/2021-12/cp210216en.pdf

22 Eligible companies under the Act are for-profit companies that have annual gross revenue of over $25 million, or annually buy, sell, or share the personal information of 100,000 or more consumers or households, or derive 50% or more of annual revenue from selling or sharing consumers' personal information.

For the private right of action granted by the CCPA to vest, a data breach must have led to the unauthorized theft or exfiltration of personal information such as names, social security numbers, government identification numbers, financial or health records, and biometric data; the mere access to, or loss of, such information does not provide legal recourse for individuals. In addition, companies must have failed to institute appropriate security measures in order to be held liable under this provision.

Consumers are also required, prior to taking legal action against companies, to give written notice of the breach of the CCPA and wait 30 days for the business to rectify such breach before undertaking legal action.

The CPRA expands the category of data for which breaches can create a private right of action to include email addresses, passwords, or security questions and answers that will allow access to consumers' accounts. Breach of this information as a result of companies' failure to institute and maintain appropriate security measures will lead to a right of action. The CPRA also states that the "*implementation and maintenance of reasonable security procedures and practices . . . following a breach does not constitute a cure with respect to that breach.*"[23]

Consumers can therefore maintain private rights of action against companies for breaches with respect to their personal information even after the implementation of security measures to the extent that such measures do not rectify a breach that previously occurred.

23 California Privacy Rights Act 2020, Section 16

Creating Effective Redress Mechanisms for AI Systems

As the development and deployment of AI systems continue to advance, existing data protection redress mechanisms do not appear to be directly transferable to all AI systems or effectively compensate individuals for harms suffered due to certain design or operational features contained in these systems. Examples of these features include:

- **Potential uncoupling of capacity and moral culpability**
  Several parties are often involved in the development and deployment of AI systems, including developers and users. Where harms are occasioned by an AI system, it may also be difficult to determine which aspect of the system is responsible for such harm. Some AI systems operate entirely autonomously while others require human intervention or control. In addition, companies or government agencies that deploy AI systems may often procure such systems from developers and be unable to exercise agency or control over the behavior of the technology. It can therefore be difficult to determine moral culpability in these instances.

- **Lack of knowledge of the existence of the system**
  The ability to seek redress for harms caused by AI systems is necessarily predicated on the knowledge of the existence of such systems. AI systems often operate in the background of interactions between companies and consumers. In addition, algorithms and other technologies are often deployed on the basis of data or other inputs that may be indistinguishable, opaque, or unknown. As such, individuals may be unaware that they are the subjects of an AI system or have been interacting with one.

- **Lack of specific intent**
  It is typically unlikely that harm occasioned by a deployed AI system is the result of an intent to discriminate against or exclude a particular group of people. In fact, efforts to introduce AI systems are often well-intentioned and may seek to remove human bias from decision-making. This lack of intent to harm does not, however, preclude the occurrence of harmful impact, and may in fact limit the ability to identify or recognize such impact.

- **Lack of interpretability or explainability**
  Dynamic AI systems and models often make decisions that are somewhat or completely

unexplainable. The existence of unexplainable AI systems inhibits effective human oversight and monitoring. Even where the decisions made by such systems are accurate, it is undesirable that decisions or recommendations made by such systems are not traceable or understandable. The lack of understanding of how an AI system reaches a decision can affect the ability of affected consumers, developers, and other users of the system to identify what party is responsible for the harm caused.

- **Scale of potential impact**
  AI systems are often implemented to scale up processes to be more efficient and widespread. While this can allow for positive benefits to be quickly dispersed to various groups and communities, it also creates a risk of amplification of harmful impacts. In addition, although individual harm from bias, discrimination, or error contained in AI systems may be minimal, the reach of such systems means that they can have substantial collective impact. While the harm caused to individuals might seem minor or insignificant, the impact at a community level may be aggravated and serious, and may only be effectively redressed through collective action.

- **Exacerbation of systemic inequities**
  Although most of the bias or discrimination issues highlighted by AI systems (such as differentiated access to financial services or healthcare due to racial or gender identity) are long-standing societal issues and are not created by AI technology, the operation of AI systems results in the amplification or exacerbation of these issues. AI systems often incorporate, directly or indirectly, factors or issues that serve as proxies for bias.[24]

Given the challenges of ensuring appropriate redress and compensation for harms occasioned by AI systems, it is important that effective avenues for recourse are provided to individuals and communities harmed by the development and deployment of such systems. It is also important that developers or deployers of AI systems that cause harm are not able to avoid liability or leave individuals or communities unable to obtain redress for harms suffered as a result of those systems. Existing rules that set out liability as a result of negligence or design defect may prove hard to enforce in complex situations where, for instance, it is difficult to ascertain what design feature specifically led to the harm or who was responsible for it. As such, we set out recommendations below to ensure that effective redress mechanisms are available to affected individuals and groups.

24 Human Rights in the Age of Artificial Intelligence. *Access Now*, last modified November, 2018, https://www.accessnow.org/cms/assets/uploads/2018/11/AI-and-Human-Rights.pdf

Redress Mechanism Recommendations

These recommendations are not intended to stand alone, but rather should be applied collectively to provide a range of options to consumers seeking redress. These recommendations include both individual and collective options, and are aimed at relevant parties, including regulatory bodies, companies developing and deploying AI applications, and civil society organizations working on societal impacts of AI systems.

RECOMMENDATIONS FOR REGULATORS

1. Allow Private Rights of Action

Individuals who have suffered harm or damage from the direct deployment of an AI system should be granted the right to complain to a public agency or department with regulatory oversight over the system, or to pursue legal action in a court of law. In order to be able to exercise this right, individuals will need to know when AI systems are in use. Therefore, a right to be informed is a necessary precursor to the right to private action for recourse.

2. Establish an AI Ombudsman Service

In spite of efforts to prevent negative impacts or consequences from the development and/or deployment of AI systems, AI tools and technologies will on occasion cause harm, and create issues or grievances. These grievances may be at the individual or community level, and may be one-off occurrences or systemic issues. It is therefore important to develop an easily accessible redress mechanism that provides individuals and communities the opportunity to raise their grievances, seek accountability, and obtain redress. A dedicated AI ombudsman could help fill this role by serving as an independent arbiter of disputes or complaints.

An ombudsman is an individual or body designated to investigate and resolve complaints made by consumers or citizens against public agencies, sectors, or industries. An effective ombudsman should be independent, impartial, and able to mandate remedial actions against the companies, agencies, or departments for which it mediates disputes, including by making recommendations for procedural changes or awarding financial compensation.

In addition to reviewing individual complaints or disputes about specific AI systems, the AI ombudsman should also serve as a public repository of AI incidents to allow for regulatory monitoring of trends or risks arising from specific use cases, technologies, or industries, providing an opportunity for developers and deployers of AI systems to learn from each other's mistakes.

3. Allow Collective Redress Mechanisms / Broaden Legal Standing for Redress Actions

Collective redress is considered to be "any mechanism that may accomplish the cessation or prevention of unlawful business practices which affect a multitude of claimants or the compensation for the harm caused by such practices."[25] This redress can take the form of judicial or administrative action.

Harmful impacts of deployed AI systems are typically never limited to individuals, but will often affect many people within one or more social groups or communities in similar (if not identical) ways. In other areas where systemic or widespread harm has been suffered by groups of people, such as data protection or anti-trust, it has been found that individuals face significant obstacles when seeking redress.[26] Redress mechanisms can be inaccessible, costly, and time-consuming for individual consumers, subjects, or beneficiaries. In addition, individuals may be unwilling to commence redress action on their own, but may prefer to join in with other harmed individuals to collectively seek redress. Collective redress therefore addresses some of the obstacles typically faced by individuals and helps ensure that harmed individuals or communities have access to justice.[27]

4. Empower Civil Society to Defend Rights

Civil society organizations (including consumer advocacy organizations, academia, and other research institutes) should be empowered to either represent individual consumers or bring "general interest" complaints against AI systems that have negative impacts on a significant portion of beneficiaries or subjects of such systems, or that have been shown to have discriminatory or biased societal impacts. General interest complaints can be brought before a regulatory oversight body or a court without mandate or permission from directly impacted individuals due to the societal impact of the harm or the dereliction of the conduct.

25 Reding, Viviane, Almunia, Joaquín and Dalli, John. Towards a Coherent European Approach to Collective Redress: Next Steps. *SEC* (2010) 1192. https://ec.europa.eu/competition/antitrust/actionsdamages/Commission_2010_information_towards_european_collective_redress.pdf

26 Lahuerta, Sara. Enforcing EU Equality Law through Collective Redress: Lagging Behind? *Common Market Law Review* 55 (2018) 783–818. https://eprints.soton.ac.uk/421297/2/COLA2018070.pdf

27 Ibid

Individuals affected by AI systems may often not be aware of the workings of the systems, the underlying data used, or the weight given to various data points or inputs in the decision-making process, or even that they were the subject of or interacted with an AI system. In fact, most of the harmful or biased applications of AI systems discovered in recent years were uncovered by civil society organizations conducting their own investigations or research. Such organizations should therefore be empowered to bring actions against systems that cause harm by encouraging and/or requiring stakeholder engagement or resource and intelligence sharing.

RECOMMENDATIONS FOR CORPORATIONS

1. Establish an Internal Ombudsman

Private entities should establish internal ombudsman services to manage complaints from employees, customers, and contractors. Research has shown that where there is an internal ombudsman program in place, it will receive more complaints than other dispute processes.[28] Stakeholders such as employees and consumers may often find it easier and more accessible to approach an ombudsman than to use other redress mechanisms. The establishment of an internal ombudsman also provides the opportunity for companies to discover problems and take steps to rectify them without the need for regulatory action, increasing trust and buy-in from internal and external stakeholders.

2. Allow Meaningful External Engagement for Research and Audit Purposes

Companies typically view their models, data, and AI systems as proprietary information that should be kept confidential. As such, they may be reluctant to engage with and disclose information to stakeholders, such as academic researchers and consumer advocacy organizations working to identify bias, unfairness, or discrimination in developed or deployed AI systems or models. Secrecy as to the existence of AI systems or their operation makes such systems unaccountable or difficult to scrutinize. By allowing academic access to models or systems, particularly before deployment, companies can obtain valuable and diverse feedback, which can only serve to improve the performance of such systems and prevent the deployment of biased or inaccurate systems. Companies can also use bug bounty programs to identify harms caused (or likely to be caused) by their algorithms or AI systems. Companies that take these proactive steps to identify harms associated with AI systems and mitigate

28 Howard, Charles. *The Organizational Ombudsman: Origins, Roles and Operations - A Legal Guide* (ABA Book Publishing, 2009).

them before they are able to crystallize will be less likely to find themselves with unsatisfied consumers demanding redress.

RECOMMENDATIONS FOR CIVIL SOCIETY ORGANIZATIONS

1. Work with Communities and Individuals to Seek Redress

Civil society organizations and advocacy groups that work with underserved or marginalized groups and communities are often uniquely placed to be able to identify and collate negative impacts being suffered by individuals or groups and trace systemic issues arising as a result of deployed AI systems. They should therefore work with individuals and communities to collate evidence of harmful consequences and assist with efforts to seek redress.

2. Publish Findings on Deployed AI Systems

Organizations should endeavor to publish findings from research and audits conducted on the impact AI of systems deployed in various sectors and industries, and maintain a publicly available depository of such technologies, the harmful impact occasioned, groups affected, and scale of the impact.

Conclusion

As AI systems are increasingly being developed and deployed across key sectors, such as employment, health, finance, criminal justice, and social services administration, the need has become urgent to build effective regulatory and governance frameworks to ensure the equitable and inclusive use and integration of such systems. A key aspect of effective regulation is ensuring the ability of harmed individuals and stakeholders to seek and obtain redress in a manner that is transparent, easily accessible, and timely. Although various regulatory frameworks in effect include some redress mechanisms, the peculiarities of AI systems often reduce the effectiveness of such mechanisms and make them insufficient to address harms or risks caused by deployed AI systems. It is therefore important for AI regulatory frameworks to create redress mechanisms capable of addressing harms that arise from AI systems in an effective and consumer-centric manner. Failure to do so may further exacerbate issues of inequality and exclusion for certain demographic groups and individuals.

About the Author

Ifejesu Ogunleye is a graduate of the Master of Development Practice program at the University of California, Berkeley. She served as a Graduate Researcher at the Center for Long-Term Cybersecurity's AI Security Initiative, as a Project Policy Analyst at the CITRIS Policy Lab, and as a summer policy fellow at the AI for Good Foundation.

CLTC

Center for Long-Term Cybersecurity

UC Berkeley

Center for Long-Term Cybersecurity
cltc.berkeley.edu
@CLTCBerkeley