The Hooks Institute Policy Papers

Volume 2023 *Issue 1* Article 1

6-1-2023

The Promise and Peril: Unpacking the Impact of AI and Automation on Marginalized Communities

The Benjamin L. Hooks Institute for Social Change

Follow this and additional works at: https://digitalcommons.memphis.edu/hookspolicypapers

Recommended Citation

The Benjamin L. Hooks Institute for Social Change (2023) "The Promise and Peril: Unpacking the Impact of AI and Automation on Marginalized Communities," *The Hooks Institute Policy Papers*: Vol. 2023, Article 1.
Available at: https://digitalcommons.memphis.edu/hookspolicypapers/vol2023/iss1/1

This Full Issue is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in The Hooks Institute Policy Papers by an authorized editor of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.

HOOKS INSTITUTE POLICY PAPERS

The Promise and Peril: Unpacking the Impact of AI and Automation on Marginalized Communities

IN THIS ISSUE

JUNE 2023
EIGHTH EDITION

THE UNIVERSITY OF MEMPHIS | The Benjamin L. Hooks Institute for Social Change

TABLE OF CONTENTS

FOREWORD

Artificial Intelligence and Automation and Its Potential Impact on Race and Class

To say the world is in the throes of a technological revolution spearheaded by artificial intelligence ("AI"), and automation, may be one of the most understated observations of this century. While "Fake News" ran rampant on social and other media and influenced the November 2016 presidential election, that election provided ample warning of how media manipulated to mislead can have enormous negative consequences for every segment of life, including personal and employment relationships, national security, elections, media, etc.

However, something is intriguing about AI and automation. It gives us access to a futuristic society allowing us to explore unchartered waters. Bill Gates has argued for years that AI has its proven benefits. Potential uses of AI include creating personalized teaching models for students so that educators can maximize students' educational experiences (Gates, 2023). "AI can reduce some of the world inequities" (Gates, 2023) through its problem-solving capabilities, enhance worker productivity, and "[a]s computer power gets cheaper, GPT's ability to express ideas will increasingly be like having a white-collar worker available to help you with various tasks" (Gates, 2023).

As for the immediate future, AI may create as many casualties as opportunities. Undergirding the Writers Guild of America strike were Hollywood writers' concerns that AI, specifically the program ChatGPT (which can produce creative writing and audio in response to prompts), might reduce or eliminate the need for screenwriters in the future (Fortune 2023).

Individuals, governments, and organizations have used AI in insidious ways. In public housing complexes, surveillance cameras create over-policing of people of color. Despite the lack of evidence showing that Facial Recognition Technology (FRT) makes public housing complexes safer, "many of the 1.6 million Americans who live [in public housing] . . . are overwhelming people of color [who are subjected] to round-the-clock surveillance" (MacMillian, 2023). For example, in the small town of Rolette, North Dakota, the public housing complex has 100 residents under the surveillance of 107 cameras, "a number of cameras per capita approaching that found in New York's Riker Island complex" (MacMillian, 2023).

FRT has led to evictions for minor or alleged infractions that have uprooted lives. In Steubenville, Ohio, a resident was evicted for removing a laundry basket from the washing room of the complex, and another was threatened with eviction because she loaned her key fob to an authorized guest (MacMillian, 2023). The latter resident demonstrated that her vision loss required the help of her friend, who brought her groceries, thus successfully pleading her case against eviction (MacMillian, 2023). A single mother of two in New Bedford, Massachusetts, who received an eviction notice in 2021, stated that the public housing authority "made [her] life hell" when they alleged that her ex-husband - who was taking care of their children while his former wife worked during the day and attended school at night - was staying in the apartment without contributing rent in violation of the rules (MacMillian, 2023). Even Bill Gates acknowledges that the new frontier of AI is not without rugged and scorched terrain that produces inequities. Gates recognizes that "market forces won't naturally produce AI products and services that help the poorest. The opposite is more likely." He contends that "[w]ith reliable funding and the right policies, governments and philanthropy can ensure that AIs are used to reduce inequity" (Gates, 2023).

Automation

The speed with which technology and automation are transforming the landscape is taking place with unprecedented velocity, even outpacing the rate with which changes occurred during the industrial revolution. "The speed of current breakthroughs has no historical precedent. Compared with previous industrial revolutions, the [technological revolution] is evolving at an exponential rather than a linear pace. Moreover, it is disrupting almost every industry in every country. The breadth and depth of these changes herald the transformation of entire systems of production, management, and governance (Schwab, 2015).

This technological revolution has the potential to raise global income levels and improve the quality of life for populations around the world. To date, those who have gained the most from it have been consumers able to afford and access the digital world (Schwab, 2015). By contrast, African Americans, Hispanics, and marginalized people clustered in service,

warehouse, and other low skills occupations are the least likely beneficiaries of AI and automation gains because they are the most susceptible to job loss because of it. (McFerren & Delavega, 2018).

As the nation and world grapple with the societal impact of AI and Automation, the Hooks Institute remains focused on a core question central to promoting justice and equality: what policies and practices will prevent AI and automation from discriminating against people of color and other marginalized groups? How can AI and automation aid our nation in eliminating racial, economic, health, educational, and other disparities?

The policy papers in this edition analyze the impact of AI and automation in three crucial areas. Khortlan Becton, JD, MTS, explores the urgent need to regulate AI to eradicate existing and potential policies and practices that disproportionately discriminate against African Americans and minorities. Becton proposes the creation of a new federal agency to regulate AI.

Susan Elswick, EdD, LCSW, a faculty member in the University of Memphis School of Social Work, seeks a path to using AI and Automation to provide social work counseling to those in need. Elswick not only explores how effective client counseling is dependent upon access and ability to use technology by clients but also argues that social workers require formal training from institutions of higher learning on how to use AI and automation to benefit their clients.

Meka Egwuekwe, MS, founder and executive director of Code Crew, approaches AI and automation from the perspective of a practitioner who teaches others to write computer code. Recognizing that the world is experiencing a revolution in how work is performed, Egwuekwe proposes recommendations that reskill or upskill the workforce, increased support for startups and small businesses, and a societal framework that will embrace universal basic income as a resource to aid those displaced by AI and Automation.

The world has entered the frontier of AI and Automation. Let's ensure everyone has an equitable opportunity for life, liberty, and the pursuit of happiness as we embark on this evolving and transformative frontier.

Daphene R McFerren, JD
Executive Director, Benjamin L. Hooks Institute for Social Change

Elena Delavega, PhD
Professor, Department of Social Work

Daniel Kiel, JD
FedEx Professor of Law, Cecil C. Humphreys School of Law

Editors

June 30, 2023

REFERENCES

Cole, J. (2023, May 5). Chat GPT is the 'terrifying' subtext of the writers' strike that is reshaping Hollywood. *Fortune*. Retrieved from https://fortune.com/2023/05/05/hollywood-writers-strike-wga-chatgpt-ai-terrifying-replace-workers/

Gates, B. (2023, March 21). The age of AI has begun: Artificial Intelligence is as revolutionary as mobile phones and the internet. *GatesNotes*. Retrieved from https://www.gatesnotes.com/The-Age-of-AI-Has-Begun

MacMillian, Douglas (2023, May 16). Eyes on the poor: Cameras, facial recognition watch over public housing. *The Washington Post*. Retrieved from https://www.washingtonpost.com/business/2023/05/16/surveillance-cameras-public-housing/

McFerren, D. & Delavega, E. (2018) The robots are ready! Are we? Automation, Race and the Workforce. *Hooks Policy Papers*. Retrieved from https://www.memphis.edu/benhooks/pdfs/the_robots_are_ready.pdf

Schwab, K. (2015, December 12). The fourth industrial revolution: What it means and how to respond. *Foreign Affairs*. Retrieved from https://www.foreignaffairs.com/world/fourth-industrial-revolution?fa_anthology=1116078

World Economic Forum. (May 2023). The future of jobs report. Retrieved from https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf

REGULATING FACIAL RECOGNITION TECHNOLOGY TO REDUCE BIAS

Khortlan Becton, JD, MTS[1]
The Restorative Education Institute

I. Introduction

From unlocking a phone to identifying shoplifters in real-time, facial recognition technology ("FRT") use is increasing among private companies and having an increasingly large impact on the public. According to one study, the global FRT market is expected to grow from $3.8 billion in 2020 to $8.5 billion by 2025 (MarketsandMarkets, 2020). Domestically, FRTs are a central aspect of artificial intelligence ("AI") use and development in the U.S. private sector. The following statistics demonstrate the growing prevalence of FRT use in the U.S. private sector: 72% of hotel operators are expected to deploy FRTs by 2025 to identify and interact with guests; by 2023, 97% of airports will roll out FRTs; excluding Southwest Airlines, most major US airlines currently use FRTs (Calvello, 2019).

This explosion of private FRT use has prompted many professional organizations and community organizers to call for a moratorium on FRT use until the enactment of state and federal regulatory actions. One such group noted that industry and government have adopted FRTs "ahead of the development of principles and regulations to reliably assure their consistently appropriate and non-prejudicial use" (Association for Computing Machinery [ACM], 2020). Among the stakeholders calling for such moratoriums is a concern over the alarming level of bias present within commercial FRT systems. Given the widespread integration of FRTs throughout society, both presently and to come, the presence of bias in FRTs is particularly troublesome as decision-making driven by biased FRT can lead to significant physical and legal injuries. For example, self-driving cars are more likely to hit dark-skinned pedestrians (Samuel, 2019). Biased FRTs also have the likelihood of producing discriminatory hiring decisions, credit approvals, or mortgage approvals.

Though the observable and conceivable consequences of bias in FRTs are virtually boundless, state and federal regulatory schemes have not adapted to the growth of FRTs. A continuing lag in regulations designed to address bias in FRTs will likely lead to a range of discriminatory effects that existing agencies do not have the capacity to prevent or redress. Therefore, a federal regulatory scheme propagated by a new agency specifically authorized to regulate AI technologies will better ensure the governance of private entities' use of facial recognition technologies to address bias than the current regulatory scheme.

---

1 Khortlan Becton graduated summa cum laude from the University of Alabama with Bachelor of Arts degrees in Religious Studies and African American Studies, received a Master of Theological Studies from Vanderbilt Divinity School, and a Juris Doctor from Temple School of Law. Khortlan is a Truman Scholar Finalist, a member of Phi Beta Kappa, and recipient of numerous awards from the various academic institutions that she has attended.

While attending law school, Khortlan began studying and advocating for the regulation of artificial intelligence technologies, including facial recognition technology. She served as lead author for a summary of recent literature on algorithmic bias in decision-making and related legal implications. That paper's co-author, Professor Erika Douglas (Temple School of Law), presented the literature review at the American Bar Association's Antitrust Spring Meeting in 2022.

Through holding various service and leadership roles, Becton has developed a deep appreciation for creative and collaborative problem-solving to address intergenerational issues of poverty and systemic inequality. Becton has continued to pursue her passion for education and justice by launching The Restorative Education Institute. The Institute is a non-profit organization purposed to equip youth and adults to practice anti-racism through historical education and substantive reflection.

A. The Relationship Between AI and FRTs

In popular usage, AI refers to the ability of a computer or machine to mimic the capabilities of the human mind and combining these and other capabilities to perform functions a human might perform (IBM, 2020). AI-powered machines are usually classified into two groups—general and narrow (Towards Data Science, 2018). Narrow AI, which drives most of the AI that surrounds us today, is trained and focused to perform specific tasks. (IBM, 2020). General AI is AI that more fully replicates the autonomy of the human brain—AI that can solve many types of problems and even choose the problems it wants to solve without human intervention (IBM, 2020).

Machine learning is a subset of AI application that enables an application to progressively reprogram itself, digesting data input by human users, to perform the specific task the application is designed to perform with increasingly greater accuracy (IBM, 2020). Deep learning, a subset of machine learning, allows applications to automatically identify the features to be used for classification, without human intervention (IBM, 2020).

Facial recognition technologies are artificial intelligence systems programmed to identify or verify the identity of a person using their face (Thales Group, 2021). "A general statement of the problem of machine recognition of faces can be formulated as follows: given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces" (Chellappa et al., 2003). Face recognition is often described as a process that first involves four steps: face detection, face alignment, feature extraction, and face recognition (Brownlee, 2019).

1. Face Detection. Locate one or more faces in the image with a bounding box.
2. Face Alignment. Normalize the face to be consistent with the database, such as geometry and photometrics.
3. Feature Extraction. Extract features from the face that can be used for the recognition task.
4. Face Recognition. Perform matching of the face against one or more known faces in a prepared database (Brownlee, 2019).

Companies are developing and implementing FRTs in new and potentially beneficial ways, such as: helping news organizations identify celebrities in their coverage of significant events, providing secondary authentication for mobile applications, automatically indexing image and video files for media and entertainment companies, and allowing humanitarian groups to identify and rescue human trafficking victims (Amazon Web Services [AWS], 2021). Recently, FRT has been in the news for its application in the investigation of the Jan. 6, 2021, Capital riot (Sakin, 2021). Other news stories about facial recognition have centered on the coronavirus pandemic. One business proposed creating immunity passports for those who are no longer at risk of contracting or spreading COVID-19 and to use FRTs to identify the immunity passport holder (Sakin, 2021). A MarketsandMarkets (2020) study estimates that the global facial recognition market is expected to grow from $3.8 billion in 2020 to $8.5 billion by 2025.

The Federal Trade Commission's ("FTC") recent settlement with Everalbum, a California-based developer of a photo storage app, exemplifies the growth of FRT use in the commercial sector and the liabilities companies may face for implementing the technology. In its complaint, the FTC alleged that Everalbum, which offered an app that allowed users to upload photos and videos to be stored and organized, launched a new feature that, by default, used face recognition to group users' photos by faces of the people who appear in the photos (Everalbum, Inc., n.d.). Everalbum also allegedly used, without affirmative express consent, users' uploaded photos to train and develop its own FRT (Everalbum, Inc., n.d.). Regarding its implementation of FRTs, the FTC charged Everalbum for engaging in unfair or deceptive acts or practices, in violation of Section 5(a) of the Federal Trade Commission

Act, by misrepresenting that it was not using facial recognition unless the user enabled it or turned it on (Everalbum, Inc., n.d.). In January 2021, Everalbum settled the FTC allegations concerning its deceptive use of FRTs. The proposed settlement requires Everalbum to delete all face embeddings the company derived from photos of users who did not give their express consent to their use and any facial recognition models or algorithms developed with users' photos or videos (Everalbum, Inc., n.d.). The company must also obtain a user's express consent before using biometric information it collected from the user to create face embeddings or develop FRTs (Everalbum, Inc., n.d.). Everalbum's recent settlement with the FTC underscores the nascency of federal governance of FRTs, as the Everalbum settlement is among the first of few federal agency enforcements targeting commercial use of FRTs (Federal Trade Commission [FTC], 2019).[2] Signaling the potential for increasing regulation and enforcement in this area, FTC Commissioner Rohit Chopra noted that FRT "is fundamentally flawed and reinforces harmful biases" while highlighting the importance of "efforts to enact moratoria or otherwise severely restrict its use" (Federal Trade Commission [FTC], (2021).

B. Bias in Facial Recognition Technologies

Although proponents of FRTs boast high accuracy rates, a growing body of research exposes divergent error rates in FRT use across demographic groups (Najibi, 2020). In the landmark 2018 "Gender Shades" report, MIT and Microsoft researchers applied an intersectional approach to test three commercial gender classification algorithms (Buolamwini & Gebru, 2018). The researchers provided skin type annotations for unique subjects in two datasets and built a new facial image dataset that is balanced by gender and skin type (Buolamwini & Gebru, 2018). Analysis of the dataset benchmarks revealed that all three algorithms performed the worst on darker-skinned females, with error rates up to 34.7% higher than for lighter-skinned males (Buolamwini & Gebru, 2018). The classifiers also performed more effectively on male faces (Buolamwini & Gebru, 2018). The researchers suggested that darker skin may not be the only factor responsible for misclassification and that darker skin may instead be highly correlated with facial geometrics or gender presentation standards (Buolamwini & Gebru, 2018). Noting that default camera settings are often optimized to better expose lighter skin than darker skin, the researchers concluded that under-and overexposed images lose crucial information making them inaccurate measures of classification within artificial intelligence systems (Buolamwini & Gebru, 2018). The report also emphasizes the need for increased diversity of phenotypic and demographic representation in face datasets and algorithmic evaluations since "[i]nclusive benchmark datasets and subgroup accuracy reports will be necessary to increase transparency and accountability in artificial intelligence" (Buolamwini & Gebru, 2018).

In 2019, the National Institute of Standards and Technology ("NIST") released a series of reports on ongoing face recognition vendor tests ("FRVT"). Using both one-to-one verification algorithms and one-to-many identification search algorithms submitted to the FRVT by corporate research and development laboratories and a few universities, the NIST Information Technology Laboratory quantified the accuracy of face recognition algorithms for demographic groups defined by sex, age, and race or country of origin (Natl. Inst. of Stand. & Technol. [NIST], 2018). The NIST used these algorithms with four large datasets of photographs collected in U.S. governmental applications[3] (Natl. Inst. of Stand. & Technol. [NIST], 2018), which allowed researchers to process a total of 18.27 million images of 8.49 million people through 189 mostly commercial algorithms from 99 developers (Natl. Inst. of Stand. & Technol. [NIST], 2018).

---

2 The FTC, which is playing an active role in the misuse of facial recognition, previously imposed a $5 billion penalty and new privacy restrictions on Facebook in 2019. Similar to the allegations against Everalbum, the complaint against Facebook alleged that Facebook's data policy was deceptive to users who have Facebook's facial recognition setting because that setting was turned on by default, while the updated data policy suggested that users would need to opt-in to having facial recognition enabled.

3 The four large datasets of photographs are: (1) Domestic mugshots collected in the U.S.; (2) Application photographs from a global population of applicants for Immigration benefits; (3) Visa photographs submitted in support of visa applicants; and (4) Border crossing photographs of travelers entering the U.S.

The FRVT report confirms that a majority of the face recognition algorithms tested exhibited demographic differentials of various magnitudes in both false negative results (rejecting a correct match) and false positive results (matching to the wrong person) (Crumpler, 2020). In regard to false positives, the NIST found: (1) that false positive rates are highest in West and East African and East Asian people, and lowest in Eastern European individuals (Natl. Inst. of Stand. & Technol. [NIST], 2018)[4], (2) that, with respect to a number of algorithms developed in China, this effect is reversed, with low false positives rates on East Asian faces; (3) that, with respect to domestic law enforcement images, the highest false positive rates are in American Indians, with elevated rates in African American and Asian populations; (4) and that false positives are higher in women than men, and this is consistent across algorithms and datasets (Natl. Inst. of Stand. & Technol. [NIST], 2018). In regard to false negatives, the NIST found: (1) that false negatives are higher in Asian and American Indian people in domestic mugshots; (2) that false negatives are generally higher in people born in Africa and the Caribbean, the effect being stronger in older individuals[5] (Natl. Inst. of Stand. & Technol. [NIST], 2018).

Encouragingly, the NIST concluded that the differences between demographic groups were far lower in algorithms that were more accurate overall (Natl. Inst. of Stand. & Technol. [NIST], 2018). This conclusion signals that as FRTs continue to evolve, the effects of bias can be reduced (Crumpler, 2020). Based on its finding that the algorithms developed in the U.S. performed worse on East Asian faces than did those developed in China, the NIST theorized that the Chinese teams likely used training datasets with greater representation of Asian faces, improving their performance on that group (Natl. Inst. of Stand. & Technol. [NIST], 2018). Thus, the selection of training data used to build algorithmic models appears to be the most important factor in reducing bias (Crumpler, 2020).

Although both the “Gender Shades” and FRVT reports identify under-representative training sets as major sources of algorithmic bias, another recent study of commercial facial algorithms led by Mei Wang showed that “[a]ll algorithms . . . perform the best on Caucasian testing subsets, followed by Indians from Asia, and the worst on Asians and Africans. This is because the learned representations predominately trained on Caucasians will discard useful information for discerning non-Caucasian faces” (Wang, 2019). Furthermore, “[e]ven with balanced training, we see that non-Caucasians still perform more poorly than Caucasians. The reason may be that faces of coloured skins are more difficult to extract and pre-process feature information, especially in dark situations” (Wang, 2019).

Between 2014 and 2018, the accuracy of facial recognition technology has increased 20-fold (Natl. Inst. of Stand. & Technol. [NIST], 2018). However, further applications of FRT will almost certainly bring new challenges if the prevalence of bias remains unchecked. According to Jan Lunter, co-founder and CEO of Innovatrics, facial recognition companies can approach the issue of bias using the insights that the biometrics industry has gained over the past two decades. “Any failure to use these techniques,” Lunter warns, “will not only fan public mistrust, but also inhibit the iterative pace of improvement shown over the past five years” (Natl. Inst. of Stand. & Technol. [NIST], 2018).

II. Current State and Federal Regulatory Schemes

Against a backdrop of scant federal regulation of commercial AI use, including FRTs, several states have adopted their own regulatory schemes to govern the emergent technology. Illinois (740 Ill Comp. Stat), Washington (Wash. Rev. Code), California (Cal. Civ. Code), and Texas (11 Tex. Bus. & Com. Code) have each enacted legislation that targets private sector use of biometric information, including facial images. The states’ legislative schemes

4 This effect is generally large, with a factor of 100 more false positives between countries.

5 These differing results relate to image quality: The mugshots were collected with a photographic setup specifically standardized to produce high-quality images across races; the border crossing images deviate from face image quality standards.

Published by University of Memphis Digital Commons, 2023

commonly define biometric identifiers that encompass facial images by describing them as “face geometry” or unique biological patterns that identify a person (Yeung et al, 2020). However, the states each employ vastly different methods of enforcement. In Texas and Washington, only the state attorney general has enforcement power (11 Tex. Bus. & Com. Code). In California, the state attorney general and the consumer share responsibility for taking action against entities that violate privacy protections (Cal. Civ. Code). While, in Illinois, any person has the right to pursue action against firms and obtain damages between $1,000 and $5,000 per violation (740 Ill. Comp. Stat). Consequently, companies such as Google, Shutterfly, and Facebook have been sued in Illinois for collecting and tagging consumers’ facial information (Yeung et al., 2020).

Facial recognition bans, which range in scope, are on the rise at the municipal level. In September 2020, Portland, Oregon, banned facial recognition use by both public and private entities, including in places of “public accommodation,” such as restaurants, retail stores and public gathering spaces (Metz, 2020). The Portland, Oregon ban does allow private entities’ use of FRTs (1) to the extent necessary to comply with federal, state, or local laws; (2) for user verification purposes to access the user’s own personal or employer-used communication and electronic devices; or (3) in automatic face detection services in social media apps (Hunton Andrews Kurth LLP, 2020). Similarly, Portland, Maine passed an ordinance in November 2020 banning both the city and its departments and officials from “using or authorizing the use of any facial surveillance software on any groups or members of the public” (Heater, 2020). The ordinance allows members of the public to sue if “facial surveillance data is illegally gathered and/or used” (Heater, 2020). Importantly, the Portland, Maine ban does not apply to private companies.

The federal government’s national AI strategy continues to take shape with constant new developments. On November 17, 2020, the Director of the Office of Management and Budget (“OMB”), pursuant to Executive Order 13859, issued a memorandum addressed to the heads of executive departments and agencies that provided guidance for the regulation of non-governmental applications of “narrow” or “weak” AI[6] (The White House, 2020). The OMB’s memo briefly recognized the potential issues of bias and discrimination in AI applications and recommended that agencies “consider in a transparent manner the impacts that AI applications may have on discrimination.” Specifically, the OMB recommended that when considering regulatory or non-regulatory approaches related to AI applications, “agencies should consider, in accordance with law, issues of fairness and non-discrimination with respect to outcomes and decision produced by the AI application at issue, as well as whether the AI application at issue may reduce levels of unlawful, unfair, or otherwise unintended discrimination as compared to existing processes.”

Pursuant to the National AI Initiative Act of 2020 (The White House, 2020), the Director of the Office of Science and Technology Policy (“OSTP”) formally established the National AI Initiative Office (the “Office”) on January 12, 2021. The Office is responsible for overseeing and implementing a national AI strategy and acting as a central hub for coordination and collaboration for federal agencies and outside stakeholders across government, industry and academia in AI research and policymaking (The White House, 2020). On October 4, 2022, the OSTP released the Blueprint for an AI Bill of Rights (the “Blueprint”), which “identified five principles that should guide the design, use, and deployment of automated systems to protect the American public in the age of artificial intelligence” (The White House Office of Science and Tech. Policy [OSTP], 2022a). The five guiding principles are: 1. Safe and Effective Systems; 2. Algorithmic Discrimination Protections; 3. Data Privacy; 4. Notice and Explanation; and 5. Human Alternatives, Consideration, and Fallback.” (The White House Office of Science and Tech. Policy [OSTP], 2022a). The AI Bill of Rights further provides recommendations for designers, developers,

6 The OMB memorandum defines “narrow” AI as “go[ing] beyond advanced conventional computing to learn and perform domain-specific or specialized tasks by extracting information from data sets, or other structured or unstructured sources of information.”

and deployers of automated systems to put these guiding principles into practice for more equitable systems. The Biden-Harris administration has also announced progress across the Federal government that has advanced the Blueprint's guiding principles, including actions from the Department of Labor, the Equal Employment Opportunity Commission, the Consumer Financial Protection Bureau, and the Federal Trade Commission ("FTC") (The White House Office of Science and Tech. Policy [OSTP], 2022b).

Most recently, U.S. Senate Majority Leader Charles Schumer has spearheaded efforts to manage AI by circulating a framework that outlines a proposed regulatory regime for AI technologies. Schumer declared on the Senate floor, "Congress must move quickly. Many AI experts have pointed out that the government must have a role in how this technology enters our lives. Even leaders of the industry say they welcome regulation." Schumer's nod towards industry leaders is likely in reference to the several congressional panels that held hearings on AI with industry experts during the week of May 16, 2023. Most notably, Sam Altman, the CEO of OpenAI, the company known for promulgating ChatGPT, testified before a Senate committee on May 16, 2023, imploring legislators to regulate the fast-growing AI industry. Altman proposed a three-point plan for regulation that called for: 1. A new government agency with AI licensing authority, 2. The creation of safety standards and evaluations, and 3. Required independent audits. In response to Altman's plea, Senator Schumer met with a group of bipartisan legislators to begin drafting comprehensive legislation for AI regulation.

The FTC has already taken an active role in regulating private sector development and use of FRT, as evidenced by its recent settlements with Facebook and Everalbum. Further solidifying the FTC's regulatory stance, acting FTC Chairwoman Rebecca Kelly Slaughter made remarks at the Future of Privacy Forum specifically tying the FTC's role in addressing systemic racism to the digital divide, AI and algorithmic decision-making, and FRTs (Federal Trade Commission [FTC], 2019).

On April 19, 2021, the FTC published a blog post announcing the Commission's intent to bring enforcement actions related to "biased algorithms" under section 5 of the FTC Act, the Fair Credit Reporting Act, and the Equal Credit Opportunity Act (Federal Trade Commission [FTC], 2021). Importantly, the statement expressly notes that, "the sale or use of—for example—racially biased algorithms" falls within the scope of the FTC's prohibition of unfair or deceptive business practices (Federal Trade Commission [FTC], 2021). The FTC also provides guidance on how companies can "do more good than harm" in developing and using AI algorithms by auditing its training data and, if necessary, "limit[ing] where or how [they] use the model;" testing its algorithms for improper bias before and during deployment; employing transparency frameworks and independent standards; and being transparent with consumers and seeking appropriate consent to use consumer data (Federal Trade Commission [FTC], 2021).

III. Argument

The fledgling federal, state, and municipal AI and FRT regulations exist in a loose patchwork that will likely complicate enforcement and compliance for private companies. These complications could hamper or, in some cases, de-incentivize the reduction of bias in FRTs as companies could seek shelter in whichever jurisdiction is most permissive. The federal government's creation of the National AI Initiative Office does not ensure reductions in FRT bias because the Office is primarily authorized to facilitate AI innovation and cooperation between the government and private companies, rather than addressing any inherent biases present in the FRTs. The FTC's recent enforcements against private use of FRTs and its recent guidelines indicate that it has an interest in addressing the use of FRTs and FRT bias. However, the FTC possesses limited authority in this context and has historically struggled to compel compliance from large corporations. Thus, a new agency, specifically authorized to regulate and eliminate issues of bias that arise from commercial FRT applications, is needed to effectively address the presence and effect of bias within FRTs.

Published by University of Memphis Digital Commons, 2023

A. The States' privacy protections for consumers, comprised of a patchwork of state and municipal regulations, are inadequate to sufficiently address the issues of bias anticipated from the commercial use of FRTs.

In the absence of federal laws that regulate the commercial use of AI, much less FRTs, state and city laws have attempted to fill the regulatory gap. State governments may be regarded and valued as "living laboratories" in some respects, but their collective piecemeal legislation concerning commercial AI and FRT use may negatively impact the reduction of bias in FRTs and could likely lead to a deregulatory "race to the bottom."
Significantly, three states–Illinois, Texas and Washington—have recognized the urgent need to address the burgeoning use of AI in the private sector and put privacy protections in place for consumers. The Illinois Biometric Information Privacy Act, passed in 2008, requires commercial entities to obtain written consent in order to capture an individual's biometric identifiers (including face geometry) or sell or disclose a person's biometric identifier (740 Ill. Comp. Stat). The Illinois Act also places security and retention requirements on any collected biometric data (740 Ill. Comp. Stat). Although Texas and Washington have enacted similar laws, their laws vary significantly from Illinois' in that only the attorney generals are authorized to enforce the laws against commercial entities (11 Tex. Bus. & Com. Code). Illinois' law, on the other hand, includes a private right of action, which has led to several lawsuits against companies such as Clearview AI, Google, and Facebook (Greenberg, 2020; Yeung et al, 2020).

The variance among the entities empowered to enforce these states' laws will likely create enforcement and compliance difficulties, particularly as it pertains to bias, because AI and FRTs inherently transcend state borders. Based on recent studies of the presence of bias in commercial FRTs, the selection of training data used to build algorithmic models appears to be the most important factor in reducing bias (Crumpler, 2020). Thus, the reduction of bias in commercial FRTs would be significantly hindered if companies are unsure whether they have access to certain images based on specific state laws. For example, Everalbum's settlement with the FTC revealed that the international company compiled FRT training datasets by combining facial images it had extracted from Ever users' photos with facial images obtained from publicly available datasets (Everalbum, Inc., n.d.). Everalbum's FRT development was geographically constrained on a state-by-state basis to exclude images from users believed to be residents of Illinois, Texas, Washington, or the European Union (Everalbum, Inc., n.d.). From the perspective of increasing representative training datasets, the company's exclusion of facial images from users in Texas and Illinois, specifically, would have negatively impacted the representation of Latinx people and other racial minorities[7] (Krogstad, 2020).

Given uncertainty among AI and FRT developers within the patchwork state regulatory scheme, paired with researchers' recommendations to increase phenotypic and demographic representation in face datasets and algorithmic evaluations (Buolamwini & Gebru, 2018), companies will likely want to conduct business in locations that enable them to have access to large amounts of data. In response, states may avoid enacting AI and FRT regulations that deter companies from conducting business in those states, resulting in what is termed as a deregulatory "race to the bottom" (Chen, 2022). If a "race to the bottom" situation was to occur in response to the patchwork of state AI regulations, then companies would likely seek to build and train FRTs in those states where consumers had less rights to their biometric data since the companies would have access to more information to compile larger datasets.

On the one hand, enabling companies' ability to compile larger datasets seems like a great avenue to reduce bias in FRT applications, as the larger datasets would provide increased phenotypic and demographic diversity. However, a lack of state standards governing the quality and collection of biometric data could negatively impact FRT accuracy and, in turn, exacerbate the presence of biased results. According to one study, non-Caucasians may

---

7 According to Pew Research, Texas is one of two states with the most Latinx people at 11.5 million. Illinois' Latinx population increased from 2010 to 2019 by 185,000 people.

perform more poorly than Caucasians on FRTs, even with balanced training, because "faces of coloured skins are more difficult to extract and pre-process feature information, especially in dark situations" (Wang et al., 2019). Similarly, the "Gender Shade" researchers noted that default camera settings are often optimized to better expose lighter skin than darker skin (Buolamwini & Gebru, 2018). This observation led the researchers to conclude that under-and overexposed images lose crucial information making them inaccurate measures of classification within artificial intelligence systems (Buolamwini & Gebru, 2018). If biased FRT performance is linked to the difficulty of extracting and pre-processing feature information from non-Caucasian faces, especially in dark situations; and, if sub-optimal camera lightening of non-Caucasian faces often produces images that lack crucial information rendering them inaccurate datapoints; then, lax state regulations on the quality and collection of biometric data will likely widen the discrepancy between FRTs' performance on Caucasian and non-Caucasian faces, undermining efforts to reduce bias in commercial FRT use.

B. The current federal regulatory scheme lacks the scope and capacity to sufficiently address the issues of bias anticipated from the commercial use of FRTs.

The U.S. federal government, in passing the National AI Initiative Act of 2020 and creating the National AI Initiative Office (the "Office"), decided to primarily focus its resources on the support and growth of AI and its attendant technologies, including FRTs (Gibson, Dunn & Crutcher LLP, 2021). The Act also (1) expanded and made permanent the Select Committee on AI, which will serve as the senior interagency body responsible for overseeing the National AI Initiative; (2) codified the National AI Research Institutes and the National Sciences Foundation, collaborative institutes that will focus on a range of AI research and development areas, into law; (3) expanded AI technical standards to include an AI risk assessment framework; and (4) codified an annual AI budget rollup of Federal AI research and development investments (The White House Office of Science and Tech. Policy [OSTP], 2021). Further, on January 27, 2021, President Biden signed a memorandum titled, "Restoring trust in government through science and integrity and evidence-based policy making," setting in motion a broad review of federal scientific integrity policies and directing agencies to bolster their efforts to support evidence-based decisions making (The White House Office of Science and Tech. Policy [OSTP], 2021). In spite of these nascent attempts to federally regulate commercial use of FRTs, the existing commercial applications of FRTs and the instances of bias that arise from such use remain largely unregulated.

The National AI Initiative Office lacks the capacity and authority to regulate bias arising from current commercial FRT use since, according to its enabling statute, the Office is principally concerned with supporting public and private AI innovation. The National AI Initiative Act describes the Office's responsibilities as serving as a liaison between the government, industry, and academia; outreaching to the public, and promoting innovation (The White House, 2020).

None of the enumerated responsibilities described in the National AI Initiative Act authorize the Office to specifically regulate existing commercial AI use, let alone address any issues of bias. The first two responsibilities establish the Office's authority to "provide technical and administrative support" to other federal AI Initiative committees and serve as a liaison on federal AI activities between a broadly defined group of public and private entities. The last two responsibilities charge the Office with reaching out to "diverse stakeholders" and promoting interagency access to the AI Initiative's activities. The Office's enabling statute does not clearly indicate whether the regulatory body has enforcement authority on private actors as there is no provision that confers on the Office the ability to promulgate rules or regulations. Likewise, the Office does not seem to have the power to impose sanctions in order to ensure industry compliance. Instead, the Office is focused on building coordination between the private sector and governmental entities to promote further AI innovation. Thus, the Office does not have explicit regulatory authority over any existing private use of AI or FRTs.

Published by University of Memphis Digital Commons, 2023

Supporters of the National AI Initiative Act and the Office may argue that the Office is appropriately situated to address issues of bias arising from the commercial use of FRTs, however that argument is undermined by the express statutory language of the Act. A supporter of the Office may point to the entity's responsibility to serve as a liaison on federal AI activities between public and private entities to argue that, by facilitating the exchange of technical and programmatic information that could address bias in AI, the Office would help FRT developers reduce bias. However, the statute does not appear to enable the Office to influence or contribute to the substantive contents of the information shared between the public and private sectors about the AI Initiative activities. If the Office lacks the ability to influence the substance of information exchanged, then it also lacks the ability to specifically direct information sharing that could redress bias in commercial AI applications. A supporter of the Office may also point to its outreach responsibility to argue that the Office will work to address bias by reaching out to diverse stakeholders, including civil rights and disability rights organizations. Yet, the statutory language is vague as to the substance of this "regular public outreach" responsibility. Without a clearer indication that the Office's public outreach efforts are directed toward or will somehow result in a reduction in AI and FRT bias, the assumption that coordinating public outreach with diverse stakeholders will sufficiently address bias in commercial FRT use remains unfounded. Hence, reducing bias that arises from the commercial use of FRTs is not an articulated central focus, nor an explicitly intended effect, of the Office's enabling statute.

Close analysis of the statutory language establishing the National AI Initiative and the Office reveals that the Office will likely operate more like a governmental think-tank to ensure coordinated AI innovation than a regulatory body with enforcement power. Such a scheme is inadequate to properly address the existing issues of bias shown in today's commercial FRTs since the AI Initiative will likely promulgate industry standards that stem from and reflect the market itself, including its apparent biases.

The few FTC regulatory decisions that have been handed down concerning existing commercial FRT applications are products of the FTC's recent actions to regulate private AI use (Facebook, Inc., n.d.). Based on its latest posts and statements, the FTC anticipates broadening its regulation of private AI and FRT use to not only focus on user consent, but also biased algorithms (Jillson, 2021). However, the FTC has limited enforcement power to sufficiently address the wide-ranging applications of FRTs and reduce the perpetuation of bias. The Federal Trade Commission Act empowers the FTC to, among other things:

> (a) prevent unfair methods of competition and unfair or deceptive acts or practices in or affecting commerce;
> (b) seek monetary redress and other relief for conduct injurious to consumers; and
> (c) prescribe rules defining with specificity acts or practices that are unfair or deceptive, and establishing requirements designed to prevent such acts or practices (15 U.S.C. §§ 41-58).

As stated in its enabling statute, the FTC's enforcement power is limited to "unfair or deceptive acts or practices in or affecting commerce" (15 U.S.C. §§ 41-58) The FTC asserts its authority over certain issues or subject areas by deeming a certain commercial practice unfair or deceptive, which is exactly what the FTC did when it released its recent AI blog post categorizing the use or sell of "biased algorithms" as an unfair and deceptive practice. Yet, the FTC will likely run into future enforcement issues in trying to prevent the use and sale of biased algorithms because they lack the willingness to enforce orders and expertise in AI training and development. Despite the FTC's recent blog post indicating its intention to bring enforcement actions related to biased algorithms, FTC Commissioner Rohit Chopra provided a statement to the Senate noting that "Congress and the Commission must implement major changes when it comes to stopping repeat offenders" and that "since the Commission has shown it often lacks the will to enforce agency orders, Congress should allow victims and state attorneys general to seek injunctive relief in court to halt violations of FTC orders (Federal Trade Commission [FTC], 2021).

In support of his first suggestion concerning the issue of repeat offenders, Commissioner Chopra emphasized that, “[w]hile the FTC is quick to bring down the hammer on small businesses, companies like Google know that the FTC simply is not serious about holding them accountable” (Federal Trade Commission [FTC], 2021). If the FTC is currently struggling to “turn the page on [their] perceived powerlessness” (Federal Trade Commission [FTC], 2021), then it follows that it is most likely ill-suited to successfully take on emerging global leaders in commercial AI technology. Furthermore, the Commissioner’s plea for Congress to allow victims and state attorneys general to access the courts for injunctive relief underscores the FTC’s inability and unwillingness to enforce its orders. Shifting the burden onto consumers and judges to regulate the exploding commercial use of FRTs and reduce bias is less than ideal as the courts lack the expertise and resources to adequately address bias in commercial FRT use. Also, courts are bound by justiciability principles, which limits their ability to regulate and reduce bias. Therefore, Congress should create a new agency that is solely authorized to address issues of bias in commercial FRT use, has power to regulate, and teeth to go after private parties who violate its regulations.

C. Congress must establish a new federal agency specifically, but not solely, authorized to regulate and eliminate issues of bias that arise from commercial FRT applications.

In order to effectively address the pervasiveness of bias in private FRT use, Congress must establish a new regulatory agency specifically, but not solely, authorized to regulate and eliminate issues of bias that arise from commercial FRT applications. The new agency should be created according to the following enabling statute to ensure its appropriate scope and capacity:

> The [agency] is empowered, among other things, to:
>
> (a) prevent private entities’ development, use, or sale of FRTs in circumstances that perpetuate bias based on ethnic, racial, gender, and other human characteristics recognizable by computer systems;
> (b) seek monetary redress or other relief for injuries resulting from the presence of bias in FRTs;
> (c) prescribe rules and regulations defining with specificity circumstances known or reasonably foreseeable to perpetuate bias that is prejudicial to established human and legal rights, and establishing standards designed to prevent such circumstances;
> (d) gather and compile data and conduct investigations related to private entities’ development, testing, and application of FRTs; and
> (e) make reports and legislative recommendations to Congress and the public.[8]

Part (a) of the new agency’s enabling statute delineates the scope of the agency’s enforcement power to specifically regulate private entities’ development, use or sale of FRTs in settings that perpetuate bias. The phrase “development, use, or sale” is designed to extend the agency’s regulatory scope to include the development or creation of FRTs in recognition of the fact that biases can originate from either the algorithm or the training dataset. Including the development stage within the agency’s regulatory authority will allow the agency to effectively regulate the sources of bias—the algorithm, training datasets, and photo quality. Additionally, the inclusion of all three stages—development, use, and sale—enable the agency to have the conceptual framework and authority to regulate any future sources of bias that are yet to be discovered[9] (Learned-Miller et al., 2020).

---

8 The new agency’s enabling statute is modeled after the Federal Trade Commission Act because the Act succinctly embodies the power of a narrowly focused agency. The FTC Act is primarily focused on “unfair and deceptive acts or practices affecting commerce,” which has contributed to the FTC’s broad authority. The new agency will need a similar breadth in their jurisdictional scope since researchers have only begun to scratch the surface of bias in FRT applications.

9 Specifically defining the concepts used to describe the creation and management of FRTs is of utmost importance to delineating the scope of not only the AI attendant technology, but also the breadth of an agency’s regulatory framework. Researchers have recently endeavored to provide specific definitions for the creation of a federal regulatory scheme for FRTs that will likely be a necessary addition to the enabling statutory language proposed here.

Published by University of Memphis Digital Commons, 2023

Part (b) confers the agency the power to impose sanctions in the form of monetary penalties or other appropriate type of relief for injuries caused by a private party's violation of the agency's regulations. Part (b) is of utmost importance since it will give the agency power to bring down the hammer on violating entities and shirk the perception of "powerlessness." The agency will compel compliance from large companies by bringing timely actions against violating parties, requiring violating parties to make material changes to their algorithms that eliminate or significantly reduce bias, and maintaining a reputation for rigorously holding companies accountable for their algorithms.

Part (c) functions hand-in-hand with Part (b) in that the agency's promulgation of rules and regulations creates the legal claims through which the agency can seek monetary redress or other forms of relief from violating companies. Requiring the agency to prescribe rules and regulations that specifically define circumstances known or reasonably foreseeable to perpetuate bias will require significant technical expertise. The agency should employ and regularly consult with preeminent AI and FRT scholars and researchers so that it can stay abreast of industry standards, norms, and developments. The agency must also develop rigorous testing standards to identify and address algorithms' rates of bias, which will require it to compile large datasets that are phenotypically and demographically representative.

Part (d) significantly empowers the agency to continually request information from private FRT developers so that it can promulgate rules and standards that can effectively address the identified sources of bias in commercial FRT applications. Without the power to gather and compile data, the agency's regulations and standards would run the risk of becoming obsolete or irrelevant to the FRT industry, which would hinder its ability to reduce bias. Similarly, the power to conduct investigations related to FRT development, testing, and applications is crucial to the agency's regulatory authority so that the agency can actively ensure companies' compliance without needing to wait on injured parties, who often lack AI expertise or access to representation, to bring claims. Based on its investigations, the agency can further ensure the sustained reduction in FRT bias by making reports and recommendations to Congress and the public.

Part (e) can be best realized by the agency because of its broad authority to regulate every aspect of FRT development and application. Thus, the agency sits at a critical juncture between FRT developers, legislators, and the public. Consequently, the agency can emphasize legislative reform as needed to effectively reduce bias and contribute to a nascent body of knowledge that the public has only begun to understand.

A new federal agency, empowered to investigate and regulate FRT development, testing, and application can reduce the presence of bias more effectively than the current regulatory scheme because of its broad authority and enforcement power. The FTC is limited in its authority to regulate bias, and its regulatory power has repeatedly bowed to the will of large corporations. Furthermore, it is not clear whether the Office has the authority to even promulgate rules or standards. Yet, FRT technology is a growing market, and researchers have only scratched the surface of how FRTs perpetuate bias. To this end, the Association of Computing Machinery's U.S. Technology Policy Committee observed that industry and government have adopted FRTs "ahead of the development of principles and regulations to reliably assure their consistently appropriate and non-prejudicial use" (Association for Computing Machinery [ACM], 2020). A new agency, specifically targeting the development, training, and application of FRTs can have the necessary breadth and expertise to reduce existing sources of bias and discover unknown sources of bias. Furthermore, the agency's narrowly tailored focus on FRTs can help to lay a foundation for its future expanded regulatory authority over additional AI attendant technologies, which are likely more complex systems. Since large corporations have not dealt with the new agency yet, the agency will be able to set itself apart from agencies with waning respect from corporations by strictly enforcing its regulations, erring on the side of caution, and crafting settlement agreements with provisions that require violators to make material changes to reduce FRT bias. Though the existence of completely unbiased FRTs is sure to be difficult to realize, the new agency will deploy all of its authority and resources to reducing FRT bias to the point of elimination.

Recommendations

The inundation of commercial facial recognition technology coupled with a lagging federal regulatory framework to govern commercial FRT development and use has led to a precarious environment where individuals bear the undue burden of redressing unprecedented harms. The following policy recommendations, while ambitious, aim to support a national regulatory scheme that would reduce the frequency and severity of FRT bias and discrimination:

- Establish a federal agency with the explicit authority to regulate commercial AI and its attendant technologies, like FRTs, in accordance with the following enabling statute:

  The [agency] is empowered, among other things, to:

  > (a) prevent private entities' development, use, or sale of FRTs in circumstances that perpetuate bias based on ethnic, racial, gender, and other human characteristics recognizable by computer systems;
  > (b) seek monetary redress or other relief for injuries resulting from the presence of bias in FRTs;
  > (c) prescribe rules and regulations defining with specificity circumstances known or reasonably foreseeable to perpetuate bias that is prejudicial to established human and legal rights, and establishing standards designed to prevent such circumstances;
  > (d) gather and compile data and conduct investigations related to private entities' development, testing, and application of FRTs; and
  > (e) make reports and legislative recommendations to Congress and the public.

- Create uniform guidelines for states' regulation of the commercial collection and use of biometric data;
- Develop and encourage the increased implementation of phenotypically and demographically diverse face datasets in commercial FRT development, training, and evaluation.

REFERENCES

11 Tex. Bus. & Com. Code § 503.001 (2019).

11 Tex. Bus. & Com. Code § 503.001(D).

15 U.S.C. §§ 41-58.

740 Ill. Comp. Stat. 14/1 (2008).

740 Ill. Comp. Stat. 14/20 (2008).

Amazon Web Services [AWS]. (2021). *The Facts on Facial Recognition with Artificial Intelligence*. Retrieved from https://aws.amazon.com/rekognition/the-facts-on-facial-recognition-with-artificial-intelligence/

Association for Computing Machinery [ACM]. (2020, June 30). *U.S. Technology Policy Committee: Statement on Principles and Prerequisites for the Development, Evaluation and Use of Unbiased Facial Recognition Technologies*. Retrieved from https://www.acm.org/binaries/content/assets/public-policy/ustpc-facial-recognition-tech-statement.pdf

Brownlee, J. (2019, July 5). A gentle introduction to deep learning for face recognition. *Machine Learning Mastery*. Retrieved from https://machinelearningmastery.com/introduction-to-deep-learning-for-face-recognition/

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research 81*, 1-15.

Cal. Civ. Code § 1798.100 – 1798.199.100 (2018).

Cal. Civ. Code § 1798.150, § 1798.199.90 (2018).

Calvello, M. (2019, Oct 23). 22 eye-opening facial recognition statistics for 2020. *Learn Hub*. Retrieved from https://learn.g2.com/facial-recognition-statistics

Chellappa, R., Phillips, P.J., Rosenfeld, A., & Zhao, W. (2003). Face recognition: A literary survey. *ACM Journals, 35*(4). https://dl.acm.org/doi/10.1145/954339.954342

Chen, J. (2022, Oct. 3). Race to the bottom. *Investopedia*. Retrieved from https://www.investopedia.com/terms/r/race-bottom.asp#:~:text=The%20race%20to%20the%20bottom%20refers%20to%20a%20competitive%20situation,can%20also%20occur%20among%20regions

Crumpler, W. (2020, May 1). The problem of bias in facial recognition. *Center for Strategic and International Studies Blog Post*. Retrieved from https://www.csis.org/blogs/technology-policy-blog/problem-bias-facial-recognition

Everalbum, Inc., (n.d.) FTC Complaint No. 1923172. Retrieved from https://www.ftc.gov/system/files/documents/cases/everalbum_complaint.pdf

Facebook, Inc., FTC Stipulated Order No. 19-cv-2184. Retrieved from https://www.ftc.gov/system/files/documents/cases/182_3109_facebook_order_filed_7-24-19.pdf

Federal Trade Commission [FTC]. (2019, July 24). FTC imposes $5 billion penalty and sweeping new privacy restrictions on Facebook. Retrieved from https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions

Federal Trade Commission [FTC]. (2021, April 20). Prepared opening statement of Commissioner Rohit Chopra, U.S. Senate Committee on Commerce, Science, and Transportation hearing on "Strengthening the FederalTrade Commission's authority to protect consumers". Retrieved from https://www.ftc.gov/system/files/documents/public_statements/1589172/final_chopra_opening_statement_for_senate_commerce_committee_20210420.pdf

Federal Trade Commission [FTC]. (2021, Feb. 10). Protecting Consumer privacy in a time of crisis, remarks of acting Chairwoman Rebecca Kelly Slaughter, future of privacy forum. Retrieved from https://www.ftc.gov/system/files/documents/public_statements/1587283/fpf_opening_remarks_210_.pdf

Federal Trade Commission [FTC]. (2021, Jan. 8). Statement of Commissioner Rohit Chopra, in the matter of Everalbum and Paravision, Commission File No. 1923172. Retrieved from https://www.ftc.gov/system/files/documents/public_statements/1585858/updated_final_chopra_statement_on_everalbum_for_circulation.pdf

Gibson, Dunn & Crutcher LLP. (2021, April 23). Artificial intelligence and automated systems legal update (1Q21). Retrieved from https://www.gibsondunn.com/artificial-intelligence-and-automated-systems-legal-update-1q21/#_ftn3

Greenberg, P. (2020, Sept. 18). facial recognition gaining measured acceptance. *State Legislatures Magazine*. Retrieved from https://www.ncsl.org/research/telecommunications-and-information-technology/facial-recognition-gaining-measured-acceptance-magazine2020.aspx

Grisales, C. (May 18, 2023). *Schumer meets with bipartisan group of senators to build a coalition for AI law,* NPR.org. https://www.npr.org/2023/05/18/1176894731/schumer-meets-with-bipartisan-group-of-senators-to-build-a-coalition-for-ai-law

Grother, P., Ngan, M., & Hanaoka, K. (20190. Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects. Natl. Inst. of Stand. & Technol., NISTIR 8280. Retrieved from https://doi.org/10.6028/NIST.IR.8280

The White House. (2020, Nov. 17). Guidance for regulation of artificial intelligence applications, H.R. 6395, § 5102. 2, OMB Op. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf

Heater, B. (2020, Nov. 4). Portland, Maine passes referendum banning facial surveillance, *Techcrunch.com*. Retrieved from https://techcrunch.com/2020/11/04/portland-maine-passes-referendum-banning-facial-surveillance/

Hunton Andrews Kurth LLP. (2020, Sept. 10). Portland, Oregon first to ban private-sector use of facial recognition technology. *Privacy & Information Security Law Blog*. Retrieved from https://www.huntonprivacyblog.com/2020/09/10/portland-oregon-becomes-first-jurisdiction-in-u-s-to-ban-the-commercial-use-of-facial-recognition-technology/

IBM. (2020, June 3). What is artificial intelligence (AI)? *IBM Cloud Education*. Retrieved from https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

Jillson, E. (2021, April 19). Aiming for truth, fairness, and equity in your company's use of AI. Federal Trade Commission [FTC]. Retrieved from https://www.ftc.gov/news-events/blogs/business-blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai

Kang, C. OpenAI's Sam Altman Urges A.I. Regulation in Senate Hearing, NYTimes.com, (May 16, 2023), https://www.nytimes.com/2023/05/16/technology/openai-altman-artificial-intelligence-regulation.html

Krogstad, J.M. (2020, July 10). Hispanics have accounted for more than half of total U.S. population growth since 2010. Pew Research. Retrieved from https://www.pewresearch.org/fact-tank/2020/07/10/hispanics-have-accounted-for-more-than-half-of-total-u-s-population-growth-since-2010/#:~:text=Some%20of%20the%20nation's%20largest,of%20the%20nation's%20Hispanic%20population

Learned-Miller, E., Ordóñez, V., Morgenstern, J., & Buolamwini,J. (2020, May 29). *Facial recognition technologies in the wild: A call for a federal office.* Center for Integrative Research in Computing and Learning Sciences. Retrieved from https://people.cs.umass.edu/~elm/papers/FRTintheWild.pdf

MarketsandMarkets, (2020, Dec. 2). Facial recognition market worth $8.5 billion by 2025. Retrieved from https://www.marketsandmarkets.com/PressReleases/facial-recognition.asp

Metz, R. (2020, Sept. 9). Portland passes broadest facial recognition ban in the US. *CNN.com*, Retrieved from https://www.cnn.com/2020/09/09/tech/portland-facial-recognition-ban/index.html

Najibi, A. (2020, Oct. 4). Racial discrimination in face recognition technology. *Harvard University, Science in the News*. Retrieved from https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/#

National Artificial Intelligence Initiative Act. U.S. Public Law 116-283 (2020).

Natl. Inst. of Stand. & Technol. [NIST]. (2018, Nov. 30). NIST evaluation shows advance in face recognition software's capabilities. Retrieved from https://www.nist.gov/news-events/news/2018/11/nist-evaluation-shows-advance-face-recognition-softwares-capabilities

Sakin, N. (2021, Feb. 11). Will there be federal facial recognition in the U.S.? International Association of Privacy Professionals (IAPP). Retrieved from https://iapp.org/news/a/u-s-facial-recognition-roundup/

Samuel, S. (2019, March 6). A new study finds a potential risk with self-driving cars: Failure to detect dark-skinned pedestrians. *Vox.com*. Retrieved from https://www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin

Published by University of Memphis Digital Commons, 2023

Solender, A. & Gold, A. (April 13, 2023). Scoop: Schumer lays groundwork for Congress to regulate AI, Axios.com, https://www.axios.com/2023/04/13/congress-regulate-ai-tech/

Thales Group. (2021, March 30). Facial recognition: Top 7 trends (tech, vendors, markets, use cases & latest news). Retrieved from https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/biometrics/facial-recognition

The White House Office of Science and Tech. Policy [OSTP]. (2021, Jan. 12). The White House launches the National Artificial Intelligence Initiative Office. Retrieved from https://trumpwhitehouse.archives.gov/briefings-statements/white-house-launches-national-artificial-intelligence-initiative-office/

The White House Office of Science and Tech. Policy [OSTP]. (2022a). Blueprint for an AI bill of rights. Retrieved from https://www.whitehouse.gov/ostp/ai-bill-of-rights/

The White House Office of Science and Tech. Policy [OSTP]. (2022b). Biden-Harris administration announces key actions to advance tech accountability and protect the rights of the American public. Retrieved from https://www.whitehouse.gov/ostp/news-updates/2022/10/04/fact-sheet-biden-harris-administration-announces-key-actions-to-advance-tech-accountability-and-protect-the-rights-of-the-american-public/

Towards Data Science. (Sept. 14, 2018). Clearing the confusion: AI vs machine learning vs deep learning differences. Retrieved from https://towardsdatascience.com/clearing-the-confusion-ai-vs-machine-learning-vs-deep-learning-differences-fce69b21d5eb

Wang, M., Deng, W., Hu, J., Xunqiang, T., & Huang, Y. (2019). Racial faces in-the-wild: Reducing bias by information maximization adaption network, *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South). pp. 692-702, doi: 10.1109/ICCV.2019.00078 https://arxiv.org/pdf/1812.00194.pdf

Wash. Rev. Code § 19.375.020 (2017).

Wash. Rev. Code §19.375.030 (2017).

Yeung, D., Balebako, R., Gutierrez, C.I., & Chaykowsky, M. (2020). *Face recognition technologies: Designing systems that protect privacy and prevent bias*. Santa Monica, CA: Rand Corporation.

Zakrzewski, C., Tiku, N., Lima C. & Oremus W. (May 16, 2023). OpenAI CEO tells Senate that he fears AI's potential to manipulate views, *Washington Post.com*, https://www.washingtonpost.com/technology/2023/05/16/ai-congressional-hearing-chatgpt-sam-altman/

AUTOMATION AND ARTIFICIAL INTELLIGENCE IN THE FIELD OF SOCIAL WORK: UNDERSTANDING TECHNOLOGY IN PRACTICE

Susan Elswick EdD, LCSW[1]
School of Social Work, University of Memphis

Introduction

Technology use in the field of social work is not new, but during 2019-2021 there was a surge in the use of technology and telehealth practices due to the COVID-19 pandemic (Cristofalo, 2021). As the use of technology in applied clinical settings grew rapidly, practitioners and clinicians have had much to think about as it relates to direct practice work (Mishna et al., 2021). Historically, the field of social work has been hesitant to move towards technology-based supports (Mathiyazhagan, 2021). Slow acceptance and implementation of technology within the field has resulted from concerns about client confidentiality and access, as well as clinician knowledge of and competence with technology, not to mention concerns about automation of the profession and worries about the negative impact on the therapeutic relationship central to effective treatment (Elswick, 2017; Hill & Ferguson, 2014; Anderson-Meger, 2011; McCoyd et al, 2022). When the COVID-19 pandemic occurred, many of these reluctant professionals were thrust into the world of technology, and many felt ill equipped and ill prepared (Ashcroft, Sur, Greenblatt, and Donahue, 2021). During this time, social work practitioners were challenged with navigating therapeutic relationships with clients while also conducting training and educational programming in a virtual space their training and expertise did not prepare them to enter. This article will examine the use of technology in social work practice, discussing the types of technology and reviewing benefits and challenges with their use in practice.

Understanding the Need for Professional Guidance on the Use of Technology in Social Work Practice

While the COVID-19 pandemic amplified the use of technology in social work practice, the profession has long recognized potential ethical issues and concerns that come with the use of technology. To better prepare social workers for these ethical concerns, the profession utilizes a global set of standards known as the National Association of Social Worker (NASW) Code of Ethics (2017a). The core values embedded within the NASW Code of Ethics include service, social justice, dignity and worth of the person, importance of human relationships, integrity, and competence (National Association of Social Workers, 1999). Social work professionals must aim to follow the Code of Ethics in every aspect of practice.

1 Susan Elswick obtained her Master of Social Work at University of Tennessee in 2006 and her Doctor of Education in Instructional and Curriculum leadership with a specialty in Applied Behavior Analysis at the University of Memphis (UofM) in 2011. Dr. Elswick has over 17 years of clinical mental health experience that includes community mental health, case management, residential programming, ABA-based programming, school-based programming, parent coaching, integrated behavioral health, infant mental health, and home-based services. She is currently a full professor within the UofM's School of Social Work, focusing on leading the university in developing programs to support student workforce development, and in growing partnerships and direct services for the community. Dr. Elswick is a licensed clinical social worker in Arkansas, Mississippi, and Tennessee, and a licensed school social worker in Tennessee. Her research interests include the use of evidence-based behavioral interventions, and the use of informatics and technology in the field of social work.

With the deployment of technology in social work practice, it was evident that clinicians needed to be aware of access and equity, the ability for greater flexibility, and economic or geographic restraints (Jones, 2010). Organizations and institutions within the profession have developed several initiatives to help provide guidance to practitioners navigating these challenges. For example, the National Association of Social Workers (NASW) and Association of Social Work Boards developed an additional set of standards for technology and social work practice so the core values of social work would be the main motive for implementing technology when fulfilling the needs of individuals, groups, and/or communities (National Association of Social Workers & Boards, 2005). In addition, the field of social work has a recently revised set of guiding principles, known as the NASW Standards of Practice for Technology in Social Work (NASW, 2017b). There has also been a lot of work done within the field related to the Grand Challenges in the field of social work practice. The Grand Challenges in Social Work Practice were initiated by the American Academy of Social Work and Social Welfare and are a groundbreaking initiative to champion social progress powered by science. One of the specific Grand Challenges in the field of Social Work relates to harnessing technology for good (Blair et al, 2021).

The role of social work is evolving, and social workers need to adjust to the changes in social work practice in the technology age. The problem is technology training has not been openly available to illustrate how social workers can use technology for clinical practice and community service (Pascoe, 2021). The standards previously mentioned are available to demonstrate the importance of implementing technology from an ethical point of view; however, there is very little training, professional development, or education emphasizing how technology can be utilized from an applicability standpoint for the helping profession. Additionally, there is relatively little information on preparing social workers to be leaders in technology development. This level of leadership is needed so that our domain specialty and professional lens are reflected within these technologies from inception and not as an afterthought.

The research by Burton and van den Broek (2009), indicates that effective professional development and training on new technology is needed (Burton & van den Broek, 2009). Clinical practitioners face a growing need to be educated regarding the best practices for technology-based supports, and institutions of higher education are challenged with preparing these practitioners with the needed 21st century skills to provide services and participate in the development of future technologies.

Overview of Technology, Automation, and Artificial Intelligence in Social Work Practice

One of the first steps in preparing clinicals to utilize technology in practice is supporting their knowledge about the types of technology that exist and can be implemented in practice. Maheu et al (2004) have constructed the following categories, which can assist clinicians with organizing technologies even though there is much overlap between and among them:

- Telephone/audio counseling and video/web conferencing tools
- Web-based and computer-based therapeutic tools
- Web-based text communications
- Mobile/handheld/sensor technologies

Social work and social sciences have historically used email, text messaging, and telephone counseling when indicated and appropriate; however, as technology advanced so did the need for social workers to enhance their use of newer more advanced technology within the field (Elswick, 2017; Pascoe, 2021). More contemporary versions of technology within the field include video conferencing, tele-behavioral health, automated self-help web-based interventions, utilizing social networks, and, more recently, the use of Artificial Intelligence (AI) (Elswick

2017; Asakura et al., 2020). These more contemporary technologies are being utilized more than ever. The use of this type of technology in practice for the field of mental health, behavioral health, and social sciences is on the cusp of growth.

Some newer technologies raise concerns about the potential loss of the therapeutic relationship between the clinician and the client (McCoyd et al, 2022). That relationship is at the heart of social work practice and is indicated as one of the major influences in positive therapeutic outcomes (Ardito & Rebellino, 2011; Wampold, 2012). Of particular concern are automated technologies and artificial intelligence-based technologies (Asakura et al., 2020). As we start to investigate the impact these technologies may have on therapeutic outcomes, professionals must understand the differences between these types of technologies and their use in practice.

Automated Technologies

Automation is a term for technology applications where human input is minimized in a transaction. Typically, automation is utilized to make something easier, more efficient, and more streamlined for professionals and consumers (Goldberg, 2012). Automated technologies can be useful for behavioral and mental health education and interventions. There are types of automated technologies focused on supporting the clinicians in practice, and there are automated technologies focused on supporting the client directly. Many of these automated processes have made practitioner's assessment, data collection, and evaluation of client progress much more streamlined and productive (Pascoe, 2021). These technology advances have been valuable to the profession but should still rely on the clinician validating the automated documents and processes for accuracy.

There are many examples of client-focused automated technologies. These include interactive media, online courses, artificial intelligence-powered chatbots, voice assistants, and even video games, as well as self-help media such as books, videos, audible media (e.g., podcasts), blog and print articles, and self-contained internet sites (Pascoe, 2021; Elswick, 2017). Social media, online courses, and mass-market mobile apps also can include such automated media (Pascoe, 2021; Elswick, 2017). These technologies serve to decrease geospatial, temporal, and financial barriers to consumers accessing needed services (Carayon et al., 2019). These client-focused technologies are excellent adjuncts to direct services and are often used to enhance the direct therapeutic interventions provided by the mental health practitioner (Clough & Casey, 2011). Practitioners are also involved in the use of automation for clients. Practitioners must evaluate, assess, and identify in advance which automated technologies should be offered to clients and consumers to ensure that we are advocating for consumer rights (Pascoe, 2021).

Artificial Intelligence Technologies

If automation takes the effort out of computation and repetitive tasking for clinicians and clients, artificial intelligence (AI) seeks to embody human nature (Moore et al., 2018). Although AI technology holds great promise to possibly transform mental healthcare, there are also potential pitfalls that clinicians must consider when determining if they will utilize technology in practice. The term AI was originally coined by a computer scientist, who defined it as the science of making intelligent machines (McCarthy, 1989). The concept of AI is giving human capabilities to a computer to perform a specific task. While there is evidence that AI is beginning to be leveraged in clinical settings, especially in medical processes, the field of social sciences is far from the adoption of AI in mental healthcare (Jiang et al., 2017). The reason for the slow adoption of AI into the field of mental health is related to the potential higher risks to consumers within this field of practice that use AI to facilitate simple modern-day conveniences (Miller et al., 2017; Hengstler et al., 2016). AI is being used in certain fields to support the identification of disease progression (Reddy et al., 2019), and within the field of social sciences it has been used to detect and predict various conditions, support the use of screening and assessment, and even be used for supporting clinical decision-making (Graham et al., 2019).

Although the implementation of AI in mental healthcare offers a potential solution to some of the problems with availability, attractiveness, promoting potential personalized mental health care for each client, and accessibility of mental healthcare services via technology-based apps (Bickman, 2020; Topol, 2019), many advocates for technology within practice agree that the loss of the human connection within the field could be problematic (Ramsey et al., 2014). Because of this potential loss of human connection, many professionals indicate that the use of technology should be used as an adjunct to clinical work and not necessarily a standalone intervention (Clough & Casey, 2011). According to a study done in 2015, mental health workers were noted as one of the least likely professions to be automated by machines because of the complex nature of the relationships between the client and the clinician and because of the continued need for advocacy within the field for consumers' and clients' rights (Bui, 2015). While utilizing technology to increase scalability for consumers could address the current shortage of healthcare professionals, it also could reduce the reliance on mental health professionals in the field. Such a reduction in interactions between mental health professionals and their clients is a large area of concern for the field of mental and behavioral health.

Technology, automation, and artificial intelligence are continuing to expand within the practice of social work and come with both opportunities and challenges. To better understand how to utilize technology in practice, social workers must understand the barriers and concerns that many families, specifically marginalized families, have when it comes to accessing and utilizing technologies.

Understanding the Social Injustice in Technology and the Digital Divide

Preparing community and clients for technology-supported work requires an appreciation of the historical marginalization many populations have experienced throughout the world. If the field of social work plans to utilize technology-based supports, then we must be aware of the potential barriers and issues with this mode of practice for all potential consumers. Otherwise, the technologies will simply re-create social injustices and disparities in treatment that many in the field are working to eliminate.

Oppression and racism have contributed to families of color being locked out of systems and mechanisms that could facilitate their economic mobility, growth, and full participation in society. The use of technology in practice is no different (Hamilton, 2020; Pittman, 2010). Individuals who have been provided access to and have been immersed in technology-rich environments are known as digital natives (Prensky, 2001), while those who had little to no exposure to technology-based skills are often referred to as digital immigrants (Presnky, 2001b; Hoffman & Novak, 1999). Understanding these terms helps to lay the groundwork for clinician understanding of the marginalization of people of color as it relates to technology-based supports.

In today's public education curriculum, 21st century skills are embedded into the fiber of public education (Jacobson-Londenberg, 2016; Beers, 2011); however, many families were never exposed to this type of education, training, or skills development. Historically, families of color have experienced inadequate or no education regarding 21st century skills, and this form of social injustice has contributed to poverty and poor health outcomes (Pittman, 2010; Beers, 2011; Hoffman & Novak, 1999). Access to digital resources available via the internet and other electronic platforms is increasingly important to nurture healthy child development, support family needs, and reduce risk for poor health outcomes (Sieck et al., 2021); however, for many communities of color, both hard capital and soft capital investments have been sparse or nonexistent. The term for this evident social-injustice phenomenon has been termed the “digital divide” (Cullen, 2001; Van Djik, 2020; Ramsetty & Adams, 2020; Sieck et al., 2021; Hoffman & Novak, 1999). This “digital divide” refers to the gap between individuals, households, businesses, and geographic areas at different socio-economic levels regarding both their opportunities to access information and communication technologies and their use of the Internet for a wide variety of activities (Van Djik, 2020; Sieck et al., 2021). The digital divide has always been in existence but during the COVID-19 pandemic, the digital divide for people of color and marginalized groups was exposed in the US on a national level (Sieck et al., 2021).

The children of families in communities that experience this digital divide are at a significant disadvantage in completing high school, college educational requirements, and job applications (Moore et al., 2018). The digital divide also further complicates poor health outcomes experienced by communities of color at disproportionate rates. Many families in these communities lack internet access, consistent internet speed, and the digital literacy skills necessary to be successful consumers of technology (Moore et al., 2018). These aforementioned factors contributed to the underutilization of technology-mediated services like tele-health, when they became more available to the community in response to COVID-19 lock-downs.

In addition to understanding the digital divide and the impact it has on underserved families, it is important to understand the digital divide in the context of utilizing technology in practice from a macro level perspective. To do so, we must understand the evident gaps in access to technology for consumers as well as the pros and cons of technology use in practice.

Pros of Technology, Automation, and Artificial Intelligence in Practice

As technology is more prominent, understanding the benefits of the use of technology in social work education is important. There are evident pros to the use of technology in practice from both the consumer and the clinician perspectives. Benefits include decreasing common barriers to behavioral health services for the community at large. Specifically, technology may assist with confronting problems with access to the delivery of services, lack of education, limited affordability, and stigma around requesting and accessing services (Coombs et al., 2021).

Access to effective service delivery can be a complex problem and one that many practitioners have been attempting to overcome for some time. There are times when there are more consumers in need of services than there are practitioners who can serve them (Saxena et al., 2007). With high needs and a limited number of professionals, the use of technology in practice could aid in supplementing the services for the consumer population at large.

Lack of education about mental and behavioral health needs impacts the access of appropriate and effective services. Poor understanding of mental illness impairs families’ abilities to provide adequate care for their mentally ill relatives (Saxena et al., 2007). The lack of knowledge about mental health disorders and the need for consistent treatment is a significant barrier to service delivery (Saxena et al, 2007). Increasing mental health awareness using technology could potentially positively impact the outcomes for millions of individuals who are not sure if they need service delivery.

Stigma is also a large factor in consumers accessing mental health services. Stigma is noted as one of the foremost barriers deterring people who need help from seeking it (U.S. Department of Health and Human Services, 1999). Multiple studies have shown that stigma associated with mental illness often prevents clients from seeking and eventually sustaining appropriate treatment for their specific needs (Whal, 2012); however, with the use of

Published by University of Memphis Digital Commons, 2023

technology-supported services, the stigma related to seeking services is dramatically lessened (Rodriguez-Rivas et al., 2022).

Technology-based assessments and interventions are helpful therapeutic tools that clinicians can integrate into their daily work. For example, web-based supports are extremely helpful for consumers who do not have daily contact with their care provider but may need additional daily supports to improve outcomes. These same web-based interventions also have the flexibility of being offered in the home, community, school, and beyond. The fact that web-based supports are also accessible on demand when the consumer needs them is helping to break down barriers to service delivery for many.

This level of automation has benefited the client populations as well as the clinician. Automating processes such as scoring screenings and assessments once completed, managing case notes and client documentation has greatly improved accuracy and the time often spent by social work practitioners in the field completing documentation, case notes, screening, and outcome measures (Mois & Fortuna, 2020). This type of automation provides supports to the client and clinician, but still heavily relies on the client and clinician connection in practice.
Artificial Intelligence has benefited the field by supporting the identification of possible behavioral health needs through the use of predictive analytics based on things such as blood pressure monitoring, cortisol level markers, and even possible diagnostic supports based on specific patterns identified in large data sets that may often be missed or unseen by the human clinician (Graham et al., 2019; Garcia-Ceja, et al., 2018). Utilizing additional data such as the ones mentioned above, not often provided to the clinician during direct service processes, can be used to enhance treatment outcomes for the individual client.

Cons of Technology, Automation, and Artificial Intelligence in Practice

Although there is value in the use of technology in practice, effective social work practitioners must also weigh the cons of every potential service option to make the best decisions in serving families and the community. Some of the more commonly discussed cons in the use of technology in practice include the following: data management and storage; confidentiality and privacy; informed consent; professional boundaries; safety and anonymity; licensing and regulations; practitioner competence; and, as previously mentioned, the digital divide (Van Djik, 2020; Ramsetty & Adams, 2020; Sieck et al., 2021).

Automated processes in the field of social sciences can streamline services, but without the oversight of the direct clinicians, errors can occur which could have long term effects on the clients being served. For example, within the field of automation and artificial intelligence, computerized decision making is perceived to be objective in making clinical decisions and diagnosis; however, the individuals who developed these automated tools may have bias unknowingly built into the system (Eubanks, 2018; Keddell, 2019). The use of big data in the context of automation in the field of social sciences also raises concerns around the client's ability to participate in truly informed consent if service users are not made aware of the use of their data or are unable to remove their data from the system (Keddell, 2019).

Additionally, one major concern is the loss of the human connection within these automated systems. These automated and artificial intelligent system utilize predictive algorithms to make clinician decisions in practice, which in turn threatens a client's right to be treated as an individual (Eubanks, 2018). The automated system will utilize calculations to determine treatment regiments and outcomes, but not consider the need to collaborate with the client about his/her wants in treatment planning (Keddell, 2019). Collaborative decision making is at the heart of evidence-based behavioral health practices for clients (Drisko, 2017). Finally, despite assumed objectivity, predictive algorithms are not value free and an automation of decisions in the field of social sciences could potentially reinforce inequalities and threaten service users' rights if not fully monitored for such inaccuracies. As the implementation of automated and artificial intelligence in the field of social work increases, social workers must advocate for client rights to informed consent, participation in collaborative decision making, and the ability to be treated fairly as individuals and in an unbiased manner. The need for human exchange, connection, and monitoring in the field of social sciences is evident, and not likely capable of being completed by computers alone; however, by creating leaders in the field of technology and social sciences, we can develop the future leaders of this movement to ensure that technology in practice is utilized for the benefit of the client.

Implications for Practice

The pros and cons of automation in the field of social work practice are evident from the above paragraphs, and there are some evident implications for direct practice. The field of social work must prepare future clinicians by increasing access to effective training and education in the use of technology in practice, focus on leadership skill development within the field, and train practitioners on advocacy-based best practices to ensure social justice in addressing the need of our clients and the community.

Ensuring that there are avenues for professional development, education, and training on newly adopted technologies and more novel approaches to technology in treatment is widely needed within the field. Institutions of higher education must be on the forefront of this need and must develop opportunities for professionals to acquire technological skills and to work within interdisciplinary teams to grow their leadership skills in a collaborative space. Additionally, producing future clinicians who are strong advocates for social justice in technology will be needed to ensure that we are always harnessing technology for the common good.

REFERENCES

Anderson-Meger, J. (2011). Critical thinking and e-learning in social work education. *International Journal of Business, Humanities and Technology 1*(2), 17-27.

Ardito R.B., & Rabellino D. (2011). Therapeutic alliance and outcome of psychotherapy: historical excursus, measurements, and prospects for research. *Front Psychol.* 18, :270. doi: 10.3389/fpsyg.2011.00270.

Asakura, K., Occhiuto, K., Todd, S., Leithead, C., & Clapperton, R. (2020). A call to action on artificial intelligence and social work education: Lessons learned from a simulation project using natural language processing. *Journal of Teaching in Social Work*, 40, 501-518. 10.1080/08841233.2020.1813234.

Ashcroft, R., Sur D, Greenblatt A, Donahue P. The Impact of the COVID-19 Pandemic on Social Workers at the Frontline: A Survey of Canadian Social Workers. Br J Soc Work. 2021 Jul 27:bcab158. doi: 10.1093/bjsw/bcab158. PMCID: PMC8406887.

Association of Social Work Boards. (2015). Model Regulatory Standards for Technology and Social Work Practice: ASWB International Technology Task Force, 2013-2014.

Beers, S. (2011). Teaching 21st century skills: An ASCD action tool. ASCD.

Blair G.S., Bassett, R., Bastin, L., Beevers, L., Borrajo, M.I., Brown, M., Dance, S.L., Dionescu, A., Edwards, L., Ferrario, M.A., Fraser, R., Fraser, H., Gardner, S., Henrys, P., Hey, T., Homann, S., Huijbers, C., Hutchison, J., Jonathan, P., Lamb, R. . . .& Watkins, J. (2021) The role of digital technologies in responding to the grand challenges of the natural environment: The Windermere Accord. *Patterns* (N Y), 2(1):100156. doi: 10.1016/j.patter.2020.100156.

Bickman, L. (2020). Improving mental health services: A 50-year journey from randomized experiments to artificial intelligence and precision mental health. *Administration and Policy in Mental Health*, 47, 795–843. https://doi.org/10.1007/s10488-020-01065-8

Bui, Q. (2015). Will your job be done by a machine? *NPR*. Retrieved from https://www.npr.org/sections/money/2015/05/21/408234543/will-your-job-be-done-by-a-machine

Burton, J., & van den Broek, D. (2009). Accountable and countable: Information management systems and the bureaucratization of social work. *British Journal of Social Work*, 39, 1326-1342.

Carayon, P, Hundt, AS, Hoonakker, P. (2019). Technology barriers and strategies in coordinating care for chronically ill patients. *Appl Ergon*,78, 240-247.

Clough, B.A. & Casey, L.M. (2011). Technological adjuncts to increase adherence to therapy: a review. *Clin Psychol Rev. 31*, 697-710.

Coombs, N.C., Meriwether, W.E., Caringi, J., & Newcomer, S.R. (2021). Barriers to healthcare access among U.S. adults with mental health challenges: A population-based study. *SSM Popul Health, 15*, 15:100847. doi: 10.1016/j.ssmph.2021.100847.

Cristofalo, M.A. (2021). Telehealth, friend and foe for health care social work. *Qualitative Social Work, 20*, 399-403.

Cullen, R. (2001). Addressing the digital divide. *Online Information Review, 25*, 311-320

Drisko, J. (2017). Active collaboration with clients: An underemphasized but vital part of evidence-based practice. *Social Work, 62*(2), 114–121.

Elswick, S. (2017). Informatics in social work practice: Technology within the field. Hauppauge, NY: Nova Science Publishers, Inc.

Eubanks, V. (2018). Automating inequality: How High-tech tools profile, police and punish the poor. Manhattan, NY: St. Martin's Press.

Garcia-Ceja, E., Riegler, M., Nordgreen, T., Jakobsen, P., Oedegaard, K.J., & Třrresen, J. (2018). Mental health monitoring with multimodal sensing and machine learning: A survey. *Pervasive Mobile Comput, 51*, 1–26. Goldberg, K. (2012). What is automation?. *Automation Science and Engineering, 9*, 1-2.

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports, 21*(11), 116. https://doi.org/10.1007/s11920-019-1094-0

Hamilton, A.M. (2020). A genealogy of critical race and digital studies: Past, present, and future. *Sociology of Race and Ethnicity, 6,* 292–301. https://doi.org/10.1177/2332649220922577

Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust — The case of autonomous vehicles and medical assistance devices. *Technol Forecast Soc Chang, 105*, 105–120.

Hill, K., & Ferguson, S. (2014). Web 2.0 in social work macro practice: Ethical considerations and questions. *Journal of Social Work Values & Ethics, 11*(1), 2-11.

Hoffman, D.L. & Novak, T.P. (1999). The evolution of the digital divide: Examining the relationship of race to internet access and usage over time. Retrieved from http://ecommerce.vanderbilt.edu/research/papers/pdf/manuscripts/EvolutionDigitalDivide-pdf.pdf/ accessed 7 March 2022.

Jacobson-Lundeberg, V. (2016). Pedagogical implementation of 21st century skills. *Educational Leadership and Administration, 27*, 82-100.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101McCarthy, J. (1989). Artificial intelligence, logic and formalizing common sense. In: Thomason, R.H. (eds) Philosophical Logic and Artificial Intelligence (pp. 161-190). Dordretch, NL: Springer. https://doi.org/10.1007/978-94-009-2448-2_6Jones, P. (2010). Collaboration at a distance: Using a Wiki to create a collaborative learning environment for distance education and on-campus students in a social work course. *Journal of Teaching in Social Work, 30*(2), 225-236. doi: 10.1080/08841231003705396

Keddell, E. (2019). Algorithmic Justice in child protection: Statistical fairness, social justice and the implications for practice. *Social Sciences 8*(10): 1–22.

Maheu, M.M., Pulier, M.L., Wilhelm, F.H., McMenamin, J.P., & Brown-Connolly, N.E. (2004). The mental health professional and the new technologies: A handbook for practice today. Mahwah, NJ: Lawrence Erlbaum Associates.

Mathiyazhagan, S. (2022). Field practice, emerging technologies, and human rights: the emergence of tech social workers. *J. Hum. Rights Soc. Work, 7*, 441-448 https://doi.org/10.1007/s41134-021-00190-0

McCoyd, J.L.M., Curran, L., Candelario, E. & Findley, P. (2022). "There is just a different energy": Changes in the therapeutic relationship with the telehealth transition. *Clin Soc Work J*, 50, 325–336 (2022). https://doi.org/10.1007/s10615-022-00844-0

Miller, D.D., Facp, C.M., Brown, E.W. (2018). Artificial intelligence in medical practice: The question to the answer? *Am J Med, 131*(2), 129–133.

Mois, G., & Fortuna, K.L. (2020). Visioning the Future of gerontological digital social work. *J Gerontol Soc Work, 63*, 412-427.

Moore, R., Vitale, D., & Stawingoa, N. (2018). The digital divide and educational equity. A look at students with very limited access to electronic devices at home. *Insights in Education and Work* Retrieved from https://www.act.org/content/dam/act/unsecured/documents/R1698-digital-divide-2018-08.pdf

Mishna, F., Milne, E., Bogo, M., & Pereira, L. F. (2021). Responding to COVID-19: new trends in social workers' use of information and communication technology. *Clinical Social Work Journal, 49*, 484–494.National Association of Social Workers. (2017a). Code of Ethics of the National Association of Social Workers. National Association of Social Workers. Retrieved from https://www.socialworkers. org/About/Ethics/

Code-of-Ethics/Code-of-Ethics-English.aspx

National Association of Social Workers, & Boards, A. O. S. W. (2005). NASW & ASWB standards for technology and social work practice (p. 22).

National Association of Social Workers. (1999). Code of ethics. Retrieved July 28, 2013, from http://www.socialworkers.org/pubs/Code/code.asp.

National Association of Social Workers. (2017b). NASW, ABSW, CSWE, & CSWA Standards for Technology in Social Work Practice. Washington, DC: National Association of Social Workers. Retrieved from http://www.socialworkers.org/includes/newIncludes/homepage/PRA-BRO- 33617.TechStandards_FINAL_POSTING.pdf

Pascoe, K. M. (2023). Considerations for integrating technology into social work practice: A content analysis of nine professional social work associations' Codes of Ethics. *International Social Work, 66*(2), 298–312. https://doi.org/10.1177/0020872820980833

Pittman, C. T. (2010). Race and gender oppression in the classroom: The experiences of women faculty of color with white male students. *Teaching Sociology, 38*(3), 183-196.

Prensky, M. (2001). Digital natives, digital immigrants, part 1. *On the Horizon, 9*(5), 1-6. doi10.1108/10748120110424816

Ramsey, A. T., & Montgomery, K. (2014). Technology-based interventions in social work practice: a systematic review of mental health interventions. *Social work in Health Care, 53*(9), 883–899. https://doi.org/10.1080/00981389.2014.925531

Ramsetty, A., & Adams, C. (2020). Impact of the digital divide in the age of COVID-19. *Journal of the American Medical Informatics Association : JAMIA, 27*(7), 1147–1148. https://doi.org/10.1093/jamia/ocaa078

Reddy, S., Fox, J., & Purohit, M. P. (2019). Artificial intelligence-enabled healthcare delivery. *Journal of the Royal Society of Medicine, 112*(1), 22–28. https://doi.org/10.1177/0141076818815510

Rodríguez-Rivas, M. E., Cangas, A. J., Cariola, L. A., Varela, J. J., & Valdebenito, S. (2022). Innovative Technology-Based Interventions to Reduce Stigma Toward People With Mental Illness: Systematic Review and Meta-analysis. *JMIR Serious Games, 10*(2), e35099. https://doi.org/10.2196/35099

Sieck, C. J., Sheon, A., Ancker, J. S., Castek, J., Callahan, B., & Siefer, A. (2021). Digital inclusion as a social determinant of health. *NPJ Digital Medicine, 4*(1), 52. https://doi.org/10.1038/s41746-021-00413-8

Saxena, S. Thornicroft, G., Knapp, M., Whiteford, H. (2007). Resources for mental health: scarcity, inequity, and inefficiency. *Lancet. 370*, 878-889.

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Wampold, B. E. (2012). Humanism as a common factor in psychotherapy. *Psychotherapy, 49*(4), 445–449. https://doi.org/10.1037/a0027113

Whal, O.F. (2012). Stigma as a barrier to recovery from mental illness. *Trends in Cognitive Sciences. 16*(1), 9-10.

THE IMPACT OF AUTOMATION ON OUR WORKFORCE

Meka Egwuekwe, MS[1]
Executive Director
CodeCrew

In the mid-1990s, when I was a computer science graduate student, our department at Duke University had a handful of core research areas, including artificial intelligence. Those of us who chose other focus areas often laughed at the glacial progress that AI had made over the previous 40 years since the term was first coined. We could not understand why so many government research dollars were being “wasted” on technology that seemingly would only ever work in the science fiction of books and movies.

Fast forward 25 years and what was science fiction has increasingly become fact. Artificial intelligence and robotics are driving a fourth industrial revolution that is changing the nature of the global workforce.

Automation of work is nothing new. Over centuries, mechanization, mass production, electricity, and microprocessors all ushered in waves of transformative workforce changes that resulted in some jobs becoming obsolete and eliminated, while new jobs were created. Even the new jobs of one wave, such as telephone operators, elevator attendants, or gas pumpers, were eliminated in later waves. And with each successive cycle, humans eventually reskilled and adapted.

The difference with the current wave is the breakneck pace at which modern automation, fueled by AI and robotics, has rounded the long exponential curve of productivity. Today, far fewer people are needed to generate the same amount of work. Machines are often performing tasks that humans find repetitive, boring, or unhealthy--and doing so with an increasing capacity to learn, adapt, and make decisions on their own. It is the adaptability of machine learning that is the real game changer.

Some mistakenly assume that only low wage, low education, or blue-collar jobs will be affected by this adaptability. AI and robotics are performing tasks once thought safe from automation. Machine learning is disrupting work to such a degree that high wage, high education, and white-collar jobs are increasingly affected, reducing the need for loan officers, paralegals, journalists, car and truck drivers, lawyers, bankers, tax advisors, and even radiologists. Eventually, every occupation and every task will be touched by automation--even art is being AI-generated!

---

1 Meka Egwuekwe is co-founder and executive director of CodeCrew, a Memphis-based non-profit that specializes in youth and adult computer science education. Prior to founding CodeCrew, Meka was a professional software developer for more than nineteen years, building software systems for governments and Fortune 500 companies around the world. Meka has also worked with youth and education in various capacities. He has been an Alumni Admissions Representative for Phillips Academy Andover, served as youth director at his church, and has led a spring break college tour for high school students. Meka has been a local leader in the youth coding movement since founding the Memphis Chapter of Black Girls Code in 2012. Meka is also a longtime mentor in StartCo’s business and technology accelerators, where he has mentored various companies in tech. Meka also serves on the boards of the National Civil Rights Museum, the Soulsville (Stax) Foundation, the Greater Memphis IT Council, and the Greater Memphis Workforce Development Board.

Meka has a B.S. and M.S. in Computer Science from Morehouse College and Duke University, respectively. He resides in Memphis with his wife, Pamela, and their two daughters.

What does this all mean? On one hand, AI is producing better, cheaper, more advanced products, more efficient services, safer workplaces, safer transportation, and greater access to information than ever (abuses such as algorithmic bias and disinformation aside). This inevitable and irreversible evolution of work is resulting in the widespread improvement of lives through much higher standards of living.

Except it isn't. As worker productivity has increased, income and wealth, factored for inflation, have remained flat or even declined for most Americans for decades. While technology has undoubtedly positively impacted everyone to some degree, the chief benefits of these productivity gains continue to go to those who own or finance the technology. The fast pace of technology is contrasted by an inertial plantation mindset where labor, be it human or mechanized, is in place solely to make rich owners richer--a legacy mindset that promotes inequality and continues to dominate the global economy, especially in the US.

Automation has advanced to such a degree that real questions about how we structure our society and our economy are on the table. Will there be enough jobs in the future for everyone? Will there be widespread unemployment and underemployment? Will wages further stagnate and decline? How much farther will inequality spread, and can it be mitigated? How will market economies persist if not enough consumers have sufficient income and purchasing power to buy the goods and services being produced? Will jobs maintain their role in being the chief way that income and purchasing power are distributed? Does the future hold more social unrest and violence as inequality compounds?

The fourth industrial revolution is just beginning. As modern automation accelerates, there is an opportunity to finally plan and build a society and economy that works more equitably. Risk takers should still reap rewards for the work they put into manifesting innovations, while also tangibly remembering the massive taxpayer investments in AI, robotics, and other research that made it possible to launch such innovations in the first place. As machines make it easier to produce, these gains should be serving and benefiting society in a much more balanced way.

Recommendations

Implement legislation and fund programs that support making reskilling and upskilling faster and easier. Bend the cost curve of post-secondary education to being free or near free. Establish more reskilling/upskilling programs that take less time to complete, perhaps inspired by the German vocational model and others.

Implement a universal basic income (UBI). UBI pilots around the country have shown promising results, including reduced poverty rates, improved employment prospects due to less dependence on gig and part-time work, reduced food insecurity, and overall improved physical and mental health. Intelligently coupling a UBI with achievable post-secondary educational attainment goals can raise the overall skill level of the population, raising the talent base for employers and equipping more people for entrepreneurship, while significantly reducing the aforementioned personal and social challenges.

Increase state and federal funding to support the increased number of startups and small businesses likely to manifest as the economy is further transformed by artificial intelligence. As well, increase promotion of these programs and streamline the process to receive such support.

BIBLIOGRAPHY

Bessen, J., Maarten Goos, M., Anna Salomons, A., & van den Berge, W. (2020, January). Automation: A guide for policymakers. Economic Studies at Brookings Institute. Retrieved from https://www.brookings.edu/wp-content/uploads/2020/01/Bessen-et-al_Full-report.pdf

Dawson, G. S., Desouza, K. C., & Denford, J. S. (2022, September 22). Understanding artificial intelligence spending by the U.S. federal government. Brookings Institute. Retrieved from https://www.brookings.edu/blog/techtank/2022/09/22/understanding-artificial-intelligence-spending-by-the-u-s-federal-government/

Desilver, D. (2018, August 7). For most U.S. Workers, real wages have barely budged in decades. Pew Research Center. Retrieved from https://www.pewresearch.org/fact-tank/2018/08/07/for-most-us-workers-real-wages-have-barely-budged-for-decades/

Federal Foreign Office. (1995-2023). The German vocational training system: An overview. German Missions in the United States. Retrieved from https://www.germany.info/us-en/welcome/wirtschaft/03-Wirtschaft/-/1048296

Guilford-Blake, R. (2020, February 18). Wait, will AI replace radiologists after all? Radiology Business Magazine. Retrieved from https://radiologybusiness.com/topics/artificial-intelligence/wait-will-ai-replace-radiologists-after-all

Napoletano, E. (2022, May 12). An unexpected outcome of the Covid-19 pandemic: A slew of universal income programs. Forbes. Retrieved from https://www.forbes.com/advisor/personal-finance/universal-basic-income-programs/

West, D. M., & Allen, J. R. (2018, April 24). How artificial intelligence is transforming the world. Brookings Institute. Retrieved from https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/

NOTES

NOTES

NOTES

NOTES

Published by University of Memphis Digital Commons, 2023

THE BENJAMIN L. HOOKS INSTITUTE FOR SOCIAL CHANGE

memphis.edu/benhooks

bhi@memphis.edu | 901.678.3974

@benhooksinstitute @hooksinstitute benlhooksinstitute

THE UNIVERSITY OF MEMPHIS®

ISBN 979-8-218-23365-5

90000>

9 798218 233655

The University of Memphis is an Equal Opportunity/Affirmative Action University. It is committed to the education of a non-racially identifiable student body.

Designed 2023 by PureImpressions.com