Title
AI Technology Ethics in Cognitive Science: A Review of the Literature

Permalink
https://escholarship.org/uc/item/5fw6j3nr

Journal
UC Merced Undergraduate Research Journal, 13(2)

Author
Montiel, Ma Angela Edith

Publication Date
2021

DOI
10.5070/M413255536

Peer reviewed|Undergraduate

AI Technology Ethics in Cognitive Science

A Review of the Literature

Ma Angela Edith Montiel

Cognitive Science Department of University of California, Merced

WRI 100: Advanced Writing

Dr. Mary Soltis

August 12, 2021

Introduction

I will explore issues in the field of cognitive science, specifically ethical issues surrounding cognitive science research that has malicious purposes and intent. To be more precise, how does police funding into artificial intelligence (AI) research fuel and reinforce stereotypes and attitudes towards communities of color?

The purpose is to follow intentions for AI cognitive science research in police surveillance and crime prevention. Bias against communities of color is maintained in AI systems which damages the lives of people of color. It is a significant ethical issue because AI is a relatively new technology previously adopted by police departments with negative implications such as implicit bias from harmful stereotypes. (Maerowitz 2020)

From there, we can analyze the intentions behind the financial resources and the uses of the research. This technology can change the very foundations of society, for good or bad. A review of the literature will help us understand the behind-the-scenes areas of AI research. This literature review will clarify misconceptions of AI and provide transparency for communities of color. It can also create awareness within cognitive scientists for the impacts of the research.

Most people already feel the effects of AI policing. AI has changed policing by introducing predictive policing, a process in which "algorithms" predict crime "based on previous data." (Hawkins 2020) The implications are that a person will commit a crime due to the color of their skin, as there are high rates of incarceration in people of color. Facial recognition technology used by Amazon displays bias for "lighter-skinned" people and fails to recognize "darker-skinned" people. (Lifshitz 2021) The definition of AI has drastically changed over the years. For this literature review, AI is defined as systems that "think" and "act" like humans. (IBM Cloud Education 2020)

This issue is important to consumers because AI has been used for mass policing efforts and targets communities of color. (Bailey et al. 2020) Mass policing and incarceration are harmful and are the biggest causes of death for black men. (Barry 2021) Historically, black communities were victims of destructive research, which has created the urgency for regulations to prevent the exploitation of vulnerable populations. Incidents such as the Tuskegee study come to mind. (CDC 2021) Among other ethical issues, the researchers of the study lacked the informed consent of the participants. After the discovery of the Syphilis treatment, the participants were not offered treatment. (CDC 2021)

More recent incidents involve corporate giants such as Amazon and Microsoft regarding police AI surveillance. (Maerowitz 2020) During the rise of BLM protests in the summer of 2020, Amazon came under scrutiny after several articles were released about their AI technology deployed for police use.

In the wake of new technology, researchers need many resources such as participants and funding. Additionally, law enforcement is looking for ways to aid in policing. A deeper understanding of the major players behind the scenes of a constantly changing field helps communities understand their rights and fight against systematic inequality.

The issue of police research funding is established in the theory of cognitive capitalism. This theory was initially developed by Antonella Corsani, and its primary use in the study of capitalistic value on knowledge and research. (2012) The theory of cognitive capitalism indicates that the "accumulation" of knowledge becomes value and commodity. (Corsani 2012, p. 2) This theory correlates with my research on funding by police departments. The pursuit of knowledge in research is dictated by monetary value, but evidently, many corporations are acquiring data. Data obtained by research became valuable.

Literature Review

Before we look at the impacts on society by AI, we need to understand the definitions of AI. According to IBM Cloud Education, AI is a field of the science of "intelligent computer programs," which make "predictions" based on algorithms. (2020) An algorithm is a set of rules and instructions programmed for a machine to do work.

One application of an AI system is predictive policing. According to Hawkins, predictive policing is an algorithmic process in which the system "attempts" to predict crime, victims, and offenders based on data. (2020). Hawkins explains that predictive policing may create or "worsen" issues due to "bad data, institutional biases in law enforcement, and a lack of transparency and public input." (2020)

As artificial intelligence (AI) research advances, many potential avenues for "AI-assisted crime" are discussed in a workshop by professionals from diverse backgrounds such as academia, industry, and government. (Caldwell et al. 2020, p. 1) The article "AI-enabled future crime" summarizes the findings of the "AI & Future Crime" workshop hosted by Caldwell and others (2020) and introduces possible areas of research in AI crime prevention. The fields of focus are academia, industry, and government, including the public and private sectors: Security and Crime Science, Computer Science, Public Policy, finance, retail, defense, police, and government. (Caldwell et al. 2020) Objectives are to classify threats from possible AI crime and the severity of the said crime in the "short to medium term." (Caldwell et al. 2020, p. 2) Caldwell and others (2020) conducted this research to see potential ways to prepare for AI crime prevention as uses for AI by bad actors come forward.

A key point in the conclusion of the article discusses a pattern in which the delegates determined AI crime severity through "scalable threats." (Caldwell et al. 2020, p. 12) Crimes that

affect individuals are lesser than scalable crimes that affect "society at large." (Caldwell et al.

2020, p. 12) These delegates are professionals with experience in their fields for years. This

ranking pattern is an indication of the delegates' bias. The workshop delved into the topic of bias

reflected in AI from the way real-world data is "affected by systemic, social, economic, and

historical bias." (Caldwell et al. 2020, p. 5) However, this article does provide insight from the

expertise and knowledge of field professionals. If expansion on this research occurs, researchers

must conduct this exercise with a random sampling of non-professionals. This additional

participant demographic can help determine societal views of AI crime and AI crime prevention.

It also provides transparency to the non-professional community as AI gradually becomes a way

of life.

A specific example of AI crime prevention is discussed in "IoT-Guard: Event-Driven

Fog-Based Video Surveillance System for Real-Time Security Management." (Sultana and

Wahid 2019) An issue that Sultana and Wahid attempt to solve is the reactive approach of

analyzing CCTV footage or witness reports by police enforcement after the crime. This approach

has been proven ineffective, and property crime continues to rise.

The article reports the findings of an "event-driven and fog-based smart-surveillance

system" called the "IoT-Guard" for a "smart-home environment (SHE)." (Sultana and Wahid

2019, p. 134881) Sultana and Wahid define SHE as a dwelling with different computing

applications connected to the groundwork. (2019) The fields of focus are artificial intelligence

(AI), computer science, crime, and police. A smart surveillance framework was designed by

utilizing past security and safety techniques with developing AI technologies to address rising

property crime. The researchers provide background knowledge of smart surveillance and

security and safety techniques in "smart-city" environments such as: "transportation, healthcare,

industry, and residencies." (Sultana and Wahid 2019, p. 134882) There is a description of the

proposed surveillance framework, performance evaluations in a laboratory setting, comparisons

with existing surveillance systems, and future research possibilities of the proposed framework.

Sultana and Wahid designed and developed a prototype model of IoT-Guard. (2019) The

researchers assessed various parts of the IoT-Guard system: motion detection algorithm, data

transmission, prediction accuracy, crime data transmission, and SMS alert. (Sultana and Wahid

2019) The motion detection algorithm successfully identified movement and immediately

"captured and transmitted images." (Sultana and Wahid 2019, p. 134886) Sultana and Wahid

trained the "convolutional neural network (CNN)" model on recognition through "1,000 knives"

and "800 guns" images; this model was "equally efficient" as "real-world CCTV images" in

prediction accuracy. (2019, p. 134886) The crime data transmission increased crime image

resolution with the name of the "highest probability of the crime object" in a "Short Message

Service (SMS)" messaging notification to "protective services in real-time." (Sultana and Wahid

2019, p. 134886) IoT-Guard was compared to State-of-the-Art (SoA) architecture and video

compression techniques for further research on "resource-efficiency." (Sultana and Wahid 2019,

p. 134887)

The research attempts to provide a solution for security and safety concerns from rising

property crime. The IoT-Guard surpassed other systems in energy, bandwidth, and CPU usage.

Additionally, the IoT-Guard architecture outperformed SoA architecture in "ability, scalability,

energy, and CPU and memory usage." (Sultana and Wahid 2019, p. 134893) However, the

experiments are in a controlled laboratory setting; results may differ in a real-world application.

The results show promising applications of IoT-Guard and ways of improvement for further

research. The proposed IoT-Guard, particularly its ability to upgrade without changes to the

system configuration. (Sultana and Wahid 2019) Hacking attacks commonly found in centralized internet frameworks can be prevented by "fog computing" as a "decentralized computing infrastructure." (Sultana and Wahid 2019, p. 134882)
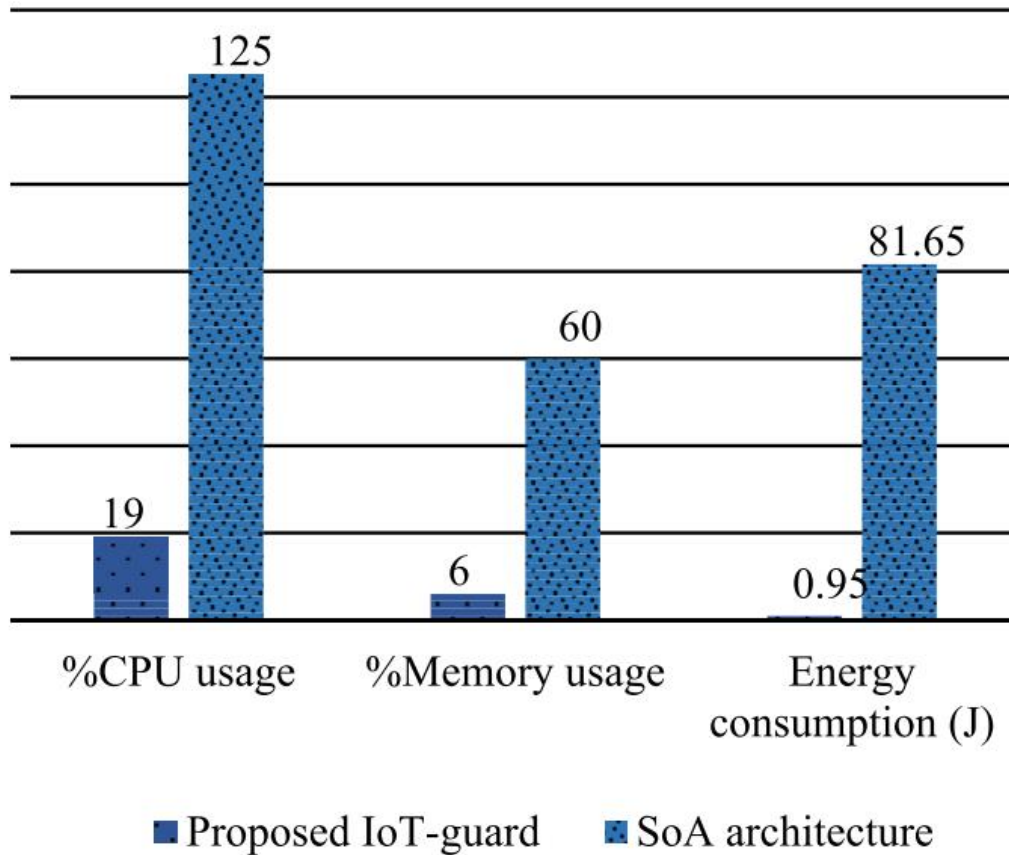
FIGURE 18. **Comparison of CPU and memory usage, and energy with SoA at edge.**

The presentation of the data in the bar graphs makes it difficult to discern the bars for comparisons, which is misleading to the reader. Specifically, Figure 18 is incomplete; it does not show the parameters for the x- and y-axis. It also combines two different parameters like percentage and joules, which is misleading. The proposed IoT-Guard works on collecting patterns, acting on its own for the best course of action, and reporting the collected data

"constantly" to police enforcement. The researchers fail to account for the infrastructure is in place to handle this system.

As ethical issues arise from AI usage, policing concerns for vulnerable communities of color caused a pushback against police institutions. One specific example of mass policing of Muslim communities is the "racialized state surveillance" of US Muslims independent of "actual criminal activity." (Alimahomed-Wilson 2019, p. 871) The article, "When the FBI Knocks: Racialized State Surveillance of Muslims," focuses on the fields of police surveillance, culture, religion, government, and defense. (Alimahomed-Wilson 2019) The objectives of the research are to analyze FBI questioning patterns of Muslim individuals due to the release of "FBI training materials." (Alimahomed-Wilson 2019, p. 871) The researcher collected data based on three components: 1) "demographic patterns," 2) "stated FBI motivations" for contact, and 3) "questioning patterns." (Alimahomed-Wilson 2019, p. 872) According to the researcher's findings, the FBI's surveillance of US Muslims is based on unfounded and racial grounds such as the "suspicion[s] of who is engaged" in common behaviors rather than criminal behaviors. (Alimahomed-Wilson 2019, p. 873)

The data collected in the research enforced the conclusion of the article while revealing new information about the demographics of the study. Police departments and the FBI are often intertwined when it comes to state surveillance. The political and historical commentary, along with the recent views of state surveillance, of the research expands the understanding of the issues Muslim communities face.

The war on terror has lasting effects on Muslim communities as they navigate the US political landscape as "perpetual target[s]" of police surveillance. (Ali 2016, p. 93) The article, "Citizens under Suspicion: Responsive Research with Community under Surveillance," highlight

the concerns and fears of Muslim citizens where "anything could be known by the state, " leading to Muslim youth to reconstruct their "social relationships." (Ali 2016, p. 93) The fields of focus are politics, police surveillance, culture, history, and community. The model of the study is the "participatory action ethnography of youth activism and organizing." (Ali 2016, p. 83). The principles of this research model are that the research and researchers "respond to" and "be willing to serve" the communities they study. (Ali 2016, p. 83) The researcher involves themselves as an observer and participant in the communities under study. (Ali 2016) The research objective is to collect data about the "citywide concerns of police monitoring" from two community service programs, the Community Activist Literacy and the New York Muslim Student Coalition. (Ali 2016, p. 83) The researcher is attempting to analyze critical discourse practices from how the students "construct meaning, interpret, and engage" in political discourse in response to their relationships with "friends, community members, community leaders." (Ali 2016, p. 84)

The article provides political commentary to the effects of police state surveillance expands on the fears and safety concerns from the systems meant to protect their citizens. The researcher inserts themselves as a participant in the study, which makes it biased. If the researcher remained an observer, it may affect the findings because the researcher grew rapport with the community as an active participant. Rapport is essential for communities that are closed off to outsiders due to past abuses and distrust. It's difficult to see the big picture of the article from the method of data analysis by the dissection of historical events and commentary.

One specific example is the Baltimore Black communities who have suffered from "state-sanctioned expropriation." (Barry 2021) Barry states that "police killings are a leading cause of death for young Black men." (2021) Particularly, Baltimore police attempted to use the "sky

plane program" to "target and track people's movements." (Barry 2021) The collection of data from Black communities only worsens police institutional bias due to the high crime rates from these communities.

Police forces often enter agreements with corporations in the widespread use of AI in mass policing. (Maerowitz 2020) Corporate giants were faced with public scrutiny as news came to light about AI technology and facial recognition software contracts with police forces. IBM has come forward with a statement of apology and a firm stance against facial recognition technology for uses of "mass surveillance, racial profiling, violations of basic human rights and freedoms." (Maerowitz 2020) Although Amazon has also stated that it will stop "facial recognition contracts," it still has contracts with "1,300" US police forces. (Maerowitz 2020) Microsoft released a statement, "facial recognition software" would not be sold to "police departments until the federal government issues clear regulation around use of the technology." (Maerowitz 2020)

Researchers train AI on data they collect. If the researcher has a bias towards a specific demographic, AI will reflect that bias in the algorithms that the researcher also programs. (Lifshitz 2021) Research on Black communities can be malicious in the infringement of their rights. The Tuskegee study "record[s] natural history of syphilis" on Black men. (CDC 2021) Informed consent of the participants was "not collected." (CDC 2021) After the development of penicillin as a treatment for syphilis, the participants were "not offered treatment." (CDC 2021)

Due to past historical research abuses to communities of color, regulations ensured fair and safe treatment. The article, "Using Human Rights Law to Inform States' Decisions to Deploy AI," identifies the steps of AI deployment to limit illegal institutional operations of AI. (Murray 2020) The fields of focus are politics, government, police, AI technology, and law. The objective

is to discuss the AI deployment for police use in compliance with current human rights laws. (Murray 2020) It attempts to aid in state decisions to determine when AI deployment is required to prevent unlawful and unethical uses of AI technology. The researcher is presenting a process in which the scale of potential harm from AI deployment is to be limited in specific "parameters" and "frameworks." (Murray 2020, p. 158) Additionally, the article only presents itself as an introduction to the conversation of AI deployment; it does not discuss the "assessment of potential harm" of AI deployment. (Murray 2020, p. 160)

The exploration of AI technology according to human rights laws views AI technology as a tool for policing and surveillance. With the introduction of mainstream AI technology, old and new regulations may be proposed or changed to reflect the changing political and technological landscape. AI technology is new to the discussion of state surveillance; the significance of AI technology needs to be addressed in light of the injustices many communities have faced from the state.

Conclusion

AI technology has great potential to cause further harm to communities of color. It's a new field that researchers may not fully understand the implications of their research without understanding the history of liability perpetrated by police institutions. Research that benefits the communities it studies is one step in the direction of ending systematic inequality. Additionally, educating communities through transparency of the research process allows them to make informed decisions about AI technology as it gradually becomes a part of their daily lives.

Cited References

Ali AI. 2016. Citizens under Suspicion: Responsive Research with Community under

Surveillance. *Anthropology & education quarterly*. [accessed 2021 Jul 30]; 47(1):78–95.

doi:10.1111/aeq.12136.

Alimahomed-Wilson S. 2019. When the FBI Knocks: Racialized State Surveillance of Muslims.

Critical sociology. [accessed 2021 Jul 30]; 45(6):871–887.

doi:10.1177/0896920517750742.

Barry KC. 2021 Jun 29. Policing Is Not "Public Safety." *The Appeal*. [accessed 2021 Aug 11].

https://theappeal.org/the-lab/explainers/policing-is-not-public-safety/.

Bailey J, Burkell J, Steeves V. 2020 Aug 24. AI technologies — like police facial recognition —

discriminate against people of colour. *The Conversation*. [accessed 2021 Aug 11].

https://theconversation.com/ai-technologies-like-police-facial-recognition-discriminate-
against-people-of-colour-143227.

Caldwell M, Andrews JT., Tanay T, Griffin L. 2020. AI-enabled future crime. *Crime science*.

[accessed 2021 Jul 30]; 9(1):1–13. doi:10.1186/s40163-020-00123-8.

Centers for Disease Control and Prevention (CDC). 2021 Apr 22. The Tuskegee Timeline. *The

U.S. Public Health Service Syphilis Study at Tuskegee*. [accessed 2021 Aug 11].

https://www.cdc.gov/tuskegee/timeline.htm.

Corsani A. 2012. The "wage/labor relations" in neoliberal cognitive capitalism. In: *Economic

Culture of Modern Capitalism*. St. Petersburg. https://halshs.archives-ouvertes.fr/halshs-
00684845.

Hawkins J. 2020 Aug 15. Problems With Predictive Policing. *Criminal Legal News*. [accessed 2021 Aug 11]. https://www.criminallegalnews.org/news/2020/aug/15/problems-predictive-policing/.

IBM Cloud Education. 2020 Aug 24. Artificial Intelligence (AI). *IBM Cloud Learn Hub*. [accessed 2021 Aug 11]. https://www.ibm.com/cloud/learn/what-is-artificial-intelligence.

Lifshitz B. 2021 May 6. Racism is Systemic in Artificial Intelligence Systems, Too. *Georgetown Security Studies Review*; [accessed 2021 Aug 11]. https://georgetownsecuritystudiesreview.org/2021/05/06/racism-is-systemic-in-artificial-intelligence-systems-too/.

Maerowitz S. 2020 Jun 11. Face, Off: IBM, Amazon and Microsoft Limit Police Use of Surveillance AI. *PR News*. [accessed 2021 Aug 11]. https://www.prnewsonline.com/IBM-Amazon-Microsoft-surveillance-PR.

Murray D. 2020. Using Human Rights Law to Inform States' Decisions to Deploy AI. *AJIL unbound*. [accessed 2021 Jul 30]; 114:158–162. doi:10.1017/aju.2020.30.

Sultana T, Wahid KA. 2019. IoT-Guard: Event-Driven Fog-Based Video Surveillance System for Real-Time Security Management. *IEEE access* [accessed 2021 Jul 30]; 7:134881–134894. doi:10.1109/ACCESS.2019.2941978.