Improper use of facial recognition technology due to the problem of many hands: The case of Robert Williams, Detroit Police, and DataWorks Plus

STS Research Paper
Presented to the Faculty of the
School of Engineering and Applied Science
University of Virginia

By

Joshua Barnard

May 14, 2021

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: 

Approved: _____ Date _____
Benjamin J. Laugelli, Assistant Professor, Department of Engineering and Society

Introduction

In a world of rapidly evolving, easily accessible technology, law enforcement agencies are turning to private companies to aid in investigations. Facial recognition technology (FRT), such as DataWorks Plus's Face Plus (FP), has been adopted by at least one quarter of US police departments and with its use has come definitive controversy from the tool's imperfections (Garvie, 2020). In the case of Robert Williams, this imperfection in FP, and lack of federal guidance led to an unnecessary arrest by Detroit P.D. that, while eventually overturned, showed that the flaws in the system could lead to a failure in law enforcement's use of FRT. There has been a large call to have federal regulation of FRT in all spaces, not only in government use but also in the private sector, and scholars have wrestled with the ethical considerations of the use of FRT—with heavy importance being placed on privacy rights and bias mitigation. However, much of the responsibility of the technology being pushed off to either the companies developing the technology or to the government exclusively, with no regard to the collective responsibility of both. By failing to consider the moral responsibility that the DataWorks Plus, the Detroit police department have above their own ethical obligations, the moral responsibility concerning biased, privacy-invading systems will not be fully realized, and similar technologies will fail.

I will explore the failure to meet the collective responsibility shouldered by DataWorks Plus and the Detroit police department through the ethical dilemma of the problem of many hands, aided by the systems, technology, and society framework (STS) of Actor Network Theory (ANT). Using ANT, I will develop and observe the moral responsibility of both the Detroit Police Department and DataWorks Plus. I will show that a moral failing still occurred despite the individual responsibilities being met—through the Detroit P.D. manual and DataWorks Plus's validation of their software, and a larger joint responsibility has to be recognized.

Facial Recognition Technology and Its Use in Law Enforcement

FRT is a tool that, given a given image, attempts to match facial features of a given image to other available images in order to verify or identify the individual in the original image (About Face, 2020). This technology is widely used in law enforcement across the United States, with over 4500 law enforcement agencies having access to an FRT tool (Garvie, 2020). The way FRT functionality generally works, is when given an image, the system will return matches with a respective confidence of match. FRT has been adopted by many law enforcement agencies across the United States and beyond, accompanied by significant controversy, in part from improper arrests of individuals incorrectly identified using the technology (Harwell, 2019).

Literature Review

On the topic of FRT, there is a large body of scholarly work analyzing the ethical ramifications of using the new technology, the need for regulation, and the responsibility of the developers of novel digital technology, however there is a definite lack of work looking at the interplay between responsibilities of the companies developing FRT and the government agencies using them. Although not in the field of law enforcement, Martinez-Martin (2019) demonstrates the numerous ethical considerations that users of FRT face in health care which parallel many of the same dilemmas in law enforcement. In each facet of the issue, there are numerous possible failure points, precipitated by both the end user and the developer. Touching on a concern in the space of biases and privacy, Martinez-Martin lays out evidence for both the importance and existence of the issue, then provides thoughts on ways to combat these challenges. However, the analysis is incomplete as no consideration is given in regards to the assignment of responsibility if the challenges are not overcome, and the system fails.

There is nearly an overwhelming amount of analysis exploring the legal ramifications of FRT, discussion ranging from its constitutionality and possible discriminatory effects to how to best regulate this relatively novel technology. In her work, Rowe (2020) argues that the "absence of federal regulation […] has created much uncertainty for companies and consumers alike", going on to cite the controversy created by the FRT company Clearview AI. Rowe cites numerous cases where the activities of Clearview and agencies utilizing their software have generated problems—including perceived privacy violations in several US states (including some lawsuits), and discomfort with the technology in the European Union. After providing these cases, Rowe furthers her argument, "Clearview's story introduces the kinds of consequences (both intended and unintended) that can arise from the absence of federal oversight," proposing that the failures of FRT are precipitated by the lack of effective regulations. Rowe falls short in her analysis of the issue, failing to acknowledge the obligation of companies developing and agencies using FRT to act morally outside of legal requirements.

There is work that shows the need for corporations to have a defined responsibility structure in this rapidly evolving digital age. In defining their concept of corporate digital responsibility (CDR), Lobschat, et al. (2019) introduce the idea of possible abuse of highly configurable AI-tools for unwanted functions, and that designers and engineers have a responsibility to be aware and attempt to anticipate these manipulations. One of the main ideas expressed, "digital technologies that assist in human decision making […] need to be subject to moral norms and ethical considerations similar to those that apply to humans," (Lobschat, et al., 2019), aims to put these technologies on the same moral ground that humans should also be held to from an outside perspective. They achieve this by establishing a set of "shared values and norms", the corporate digital responsibility of the company. From a company's CDR, there is a

viewpoint in which responsibility is placed not only on the company as a whole, but everyone involved in the production and distribution of the technology. The network they set out considers to develop a CDR, considers a wide variety of actors and stakeholders, including government institutions, legal actors, other technologies, and non-users, all of which interact to inform the creation of a company's CDR. However, Lobschat, et al. (2019) fail in their analysis of responsibility, by placing all responsibility on the developers of the technology—not acknowledging the importance of the responsibility that needs to be placed on the users as well.

Using the problem of many hands and actor network theory I will build on the ideas and concepts laid out by Martinez-Martin, Rowe, and Lobschat, et al., to provide a fully developed analysis of the concept of the responsibilities of developers and users of FRT.

The Problem of Many Hands and Actor Network Theory

In any project involving a large number of people or organizations, due to overlapping roles and ownerships of parts of the endeavor, there is often large difficulty in determining who should be held responsible if something goes awry. This predicament often has two parts, first a practical difficulty, where in large projects it's nearly impossible to find those, who through action or inaction, may contributed to a failure. The second aspect of this dilemma, is the moral failure that arises, as many people, especially those affected by the project, are often dissatisfied with the lack of ability to place blame on individuals. This has led to a development of the idea of collective responsibility, "the notion […] that there is more to responsibility in complex cases than just the sum of the responsibilities of the individuals." (Van De Poel & Royakkers). The problem of many hands can be described as "the occurrence of a situation in which the collective can reasonably be held morally responsible for an outcome, while none of the individuals can be

reasonably held responsible for that outcome." (Van De Poel & Royakkers). Analysis using the problem of many hands is aided significantly by the development of the network within which to evaluate responsibilities. The criteria for shouldering moral responsibility for a failure of technology are laid out by Van De Poel & Royakkers (2011) with four distinct parts. The first is **wrong-doing**, an action is performed or a norm violated directly precipitates a failure. The second is **causal contribution**; where-in inaction to stop wrong-doing places blame on the actor. Next is the condition of **foreseeability**, wherein an actor must reasonably be able to predict the consequences from the faulty action or inaction. Lastly an actor must have **freedom of action**, otherwise the blame cannot be placed directly upon them.

This network will be developed using the science, technology and society (STS) framework of Actor Network Theory (ANT), because it allows for the analysis of the contrasting perspectives of the primary agents. ANT views and understands the system from the mindset of the network builder(s), the primary actors who are followed and through whose eyes the network is interpreted. This process of mapping the connections that exists between all possible actors in a network—both human and non-human—can allow for powerful analysis of why a problem occurred or may occur. In the course of developing a network, complex groups of actors are become punctualizations, entire networks that because of their self-obviousness can be converted into a single actor in another network (Cressman, 2009). I will use ANT to build the system from the viewpoints of the Detroit Police Department, the analyze the moral responsibility of the primary actors—the Detroit Police and DataWorks Plus. After showing the blame cannot be placed on either, I will then leverage the problem of many hands and show how failure to consider collective responsibility by both primary actors precipitated the moral failing created in the use of FRT in a law enforcement setting.

Analysis of the Problem of Many Hands in the FRT-Police Network

Robert Williams was arrested in January of 2020, after being identified as a potential match for a shoplifting crime committed in October of 2018. This identification came from the use of DataWorks Plus's Face Plus (FP), an FRT tool. This identification, along with a positive identification from a security consultant was enough to arrest Robert Williams in his own front yard and hold him for 30 hours in jail. The security consultant who provided a secondary identification was an outside contractor for the store that had been robbed, and the consultant was not present at the time of the crime. The police department did not check Williams's alibi, and provided little to no explanation for the arrest. The case was eventually dismissed in court, but not before light was shed on the fact that the false identification had occurred due to the use of FRT (Hill, 2020). This case incited a vast amount of backlash on the Detroit Police department (DPD) and companies developing FRT, as well as a call for federal regulation of the technology. While blame has been placed on both the Detroit police department and DataWorks Plus, by failing to consider the moral responsibility each have above their own ethical obligations, the moral responsibility concerning biased, privacy-invading systems will not be fully realized, and similar technologies will suffer the same fate as FRT.

Defining Moral Responsibilities

In order to show how the moral responsibilities of the both DPD and DataWorks are not sufficient to preventing a moral failing, I first need to clarify who has been failed, and define the moral responsibilities of each actor. Both organizations are responsible to the same entity in the network, the citizens of Detroit. However, there is nuance in the obligation to the citizens—the DPD has a charge to not only serve and protect the citizens, but to do so fairly, impartially, and equally. The burden of DataWorks is different in that the technology they develop needs to affect

all people without biases—or to provide clear, interpretable results that allow a population to be equally treated under the technology.

Developing the Network

To understand the interplay of responsibilities, I will map the network to can observe how and where each failure of responsibility occurred. To construct the network, I have determined the key actors in the system. I have found the critical human and non-human actors in the system as the following: (i) *Detroit Police Examiners* those in the DPD who are able to query the FRT system; (ii) *Detroit Police Investigators* those in the DPD who are in charge of running investigations, are able to ask examiners to run queries; (iii) *Detroit Police Decision Makers* those in the DPD responsible for setting rules are regulations related to police activity; (iv) *DataWorks Plus In-House Developers* the company responsible for putting our the FRT system that police departments interact with; (v) *Outside Contractors* companies such as NEC and RankOne, whose FRT algorithms are incorporated into Face Plus's functionality; (vi) *Detroit Citizens* anyone living in or around the city of Detroit and who live under the enforcement of Michigan law by the DPD. Additionally, there are non-human or technological actors, such as (vii) *Federal FRT Regulation* any legislation passed by the US government regarding the use of FRT by law enforcement; and (viii) *Federal FRT Reports* any reports generated by the federal government agencies, such as the National Institute of Standards and Technology (NIST), on FRT technologies.

It is also important to mention the punctualizations that can occur and how this shifts the interactions from the perspective of primary actors (specifically from the DPD and DataWorks viewpoints). From the perspective of the DPD, the actors of DataWorks Plus In-House Developers and Outside Contractors are punctualized into one technological actor, *Face Plus*.

Similarly, from the point of view of DataWorks Plus, Detroit Police Examiners, Investigators, and Decision Makers are combined into the single entity of the *DPD*. This change in perspective can be seen in Figure 1.

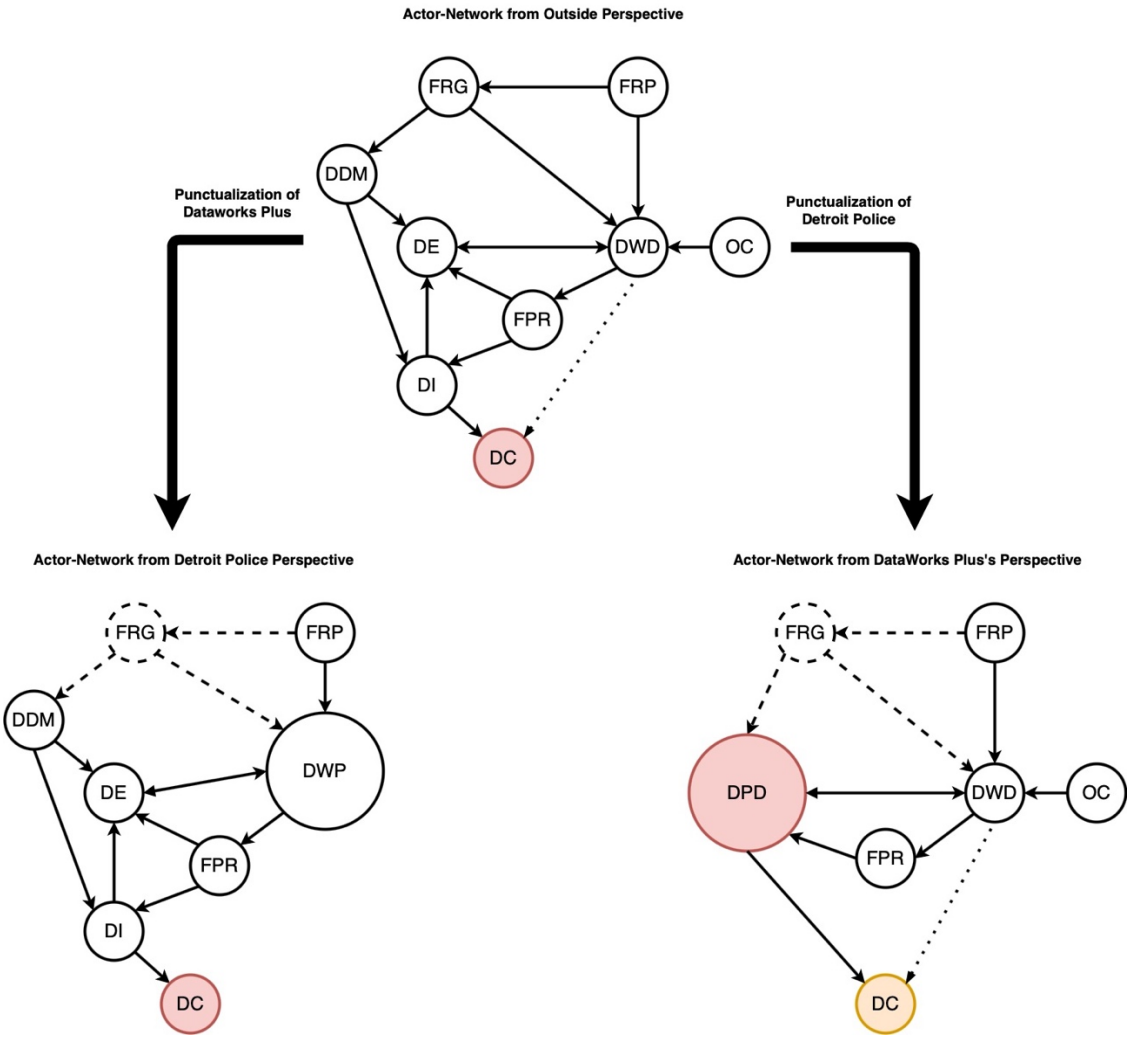

Figure 1 – The general Actor-Network's transformation to perspectives from DataWorks Plus and Detroit Police. FRG is Federal Regulations, FRP is Federal Reports, DDM is Detroit Decision Makers, DE is Detroit Examiners, DI is Detroit Investigators (DDM, DE and DI all are punctualized as DPD), DWD is DataWorks Plus Developers, OC is Outside Contractors (DWD and OC are punctualized as DWP), FPR is Face Plus Reports, and DC is Detroit Citizens. Shaded actors are those who a moral responsibility is required towards. In all cases DPD and DWP have a moral obligation to Detroit citizens (DWP has an indirect responsibility).

Using this network, I will show how both the DPD and DataWorks believed to have met their own moral responsibilities, yet a moral failing still occurred, due to a lack of consideration for the joint responsibility both had towards the citizens of Detroit. In this network diagram the larger circles represent the primary actors when punctualized, and the dashed arrows represent failed connections in the system, while dotted lines represent weak connections in the system.

The Responsibilities of the Detroit Police Department

The DPD did not shoulder the moral responsibility of the failure of FRT use in the arrest of Robert Williams due to a lack of foreseeability that the software had vulnerabilities. The DPD had to consider a variety of factors in the decision-making process of whether or not to engage in the purchase and use of DataWorks FRT software. They had to consider the legality of the process, given to them by both state and federal regulations, as well as the ethical ramifications of the process. Notice in Figure 2, the dashed arrows from FRG, Federal Regulations. At the time of the DPD engaging in their contract with DataWorks Plus, no legal regulation existed governing the use of FRT in law enforcement. Additionally, DataWorks Plus believed their FRT to be an unbiased technology (Hill, 2020). With this in mind, the DPD was ready to move forward with the use of the technology, but not before setting out rules and regulations of their own, updated regularly in Manual Directive 307.5, a document to laying out the DPD's plan, research, and deployment strategy for FRT. This document is seemingly exhaustive to the use cases of the FRT technology and lays out one extremely critical piece of information that helps

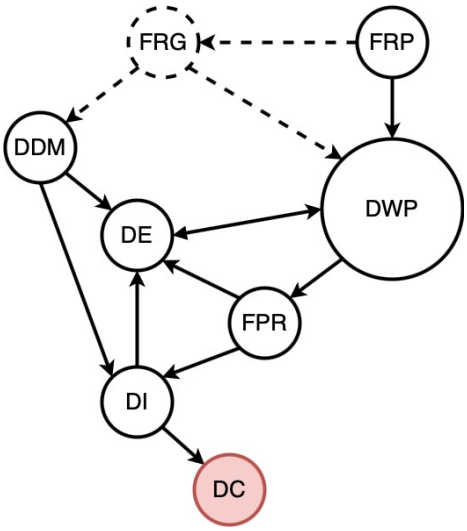

Figure 2 – The actor-network from the perspective of the DPD

keep the DPD ethically sound in their use of the technology, language at the top of every report given to Investigators from Examiners, “The result of a facial recognition search is provided by the [DPD] only as an investigative lead and IS NOT TO BE CONSIDERED POSTIVE IDENTIFICATION OF ANY SUBJECT” (Revised facial recognition directive, 2019). The DPD in spite of little or no regulation created what would appear to be a document laying out both ethical and legal use of FRT technology. In the case of Robert Williams, these rules were followed exactly and a moral failing still occurred, due to the fact that all that was needed was a secondary human identification at the time to arrest. Despite all of the language in the Manual Directive that explained the importance of the FRT report being a tool to aid in potentially building leads, to arrest a man on effectively no evidence. The protocols and regulations were inadequate at the time of Robert Williams’s arrest, and the DPD has acknowledged and made steps towards preventing future mistakes (Rahal & Hicks, 2020). To determine moral blame, we must look at the four criteria: wrong-doing and causal contribution occurred, with the system being improperly used in (either knowingly or unknowingly) that led to an incorrect arrest. Lastly the DPD had freedom of action, as they chose to any FRT technology or to not even use the technology in the first place. However, the DPD could not foresee the consequences, because with the information they had from DataWorks Plus they believed the system to work properly and without bias. Because of their inability to foresee this consequence, the DPD did not fail in their moral charge to to not only serve and protect the citizens, but to do so fairly, impartially, and equally. From both an internal and outsider’s perspective the DPD met the moral responsibility towards Detroit Citizens when using technology.

The Responsibilities of DataWorks Plus

DataWorks Plus's responsibility for the technology it develops to affect all people without bias—or to provide clear, interpretable results that allow a population to be equally treated under the technology, is one that they believed to have met in providing FRT technology to the DPD. The moral fault cannot be placed on DataWorks Plus due to a lack of foreseeability of consequences, because the program is given to police departments to use at their own discretion. DataWorks Plus had analyzed the algorithms they were using to provide the identification results and believed themselves to be "pseudo-experts in the technology" (Hill, 2020). Believing their internal vetting to be accurate, DataWorks Plus could reasonably expect the tool they provided to function as an unbiased tool and thus they have would have no need to provide clear, interpretable results for the user of that tool, meeting the expectations of no wrong-doing or causal contribution. This reasonable expectation is exemplified by the lack of direction on how to use Face Plus given to the DPD (Koebler, 2020). Todd Pastorini, DataWorks Plus's general manager, gave the viewpoint that Face Plus is just a tool that brings back a single candidate, but a variety of candidates. DataWorks Plus is assuming that the DPD will put in place their own rules and regulations needed to mitigate the risks associated with the use of FRT (Koebler, 2020), this removes the ability for them to foresee any consequences as a result, adopting an out of sight out of mind mentality towards their FRT software. Because DataWorks Plus punctualized all of the DPD and their regulations as a single actor in the network they cannot notice the nuance that exists between the individual actors in that system; however, by doing so they do meet their moral responsibility, even if by flawed assumptions.

Ignorance of Collective Responsibility Lead to Joint Failure

Having both sides not violating their moral responsibilities, how then can there have been a moral failing? This is where the Problem of Many Hands assists itself to my analysis of the situation and why a moral failing still occurred. To illuminate the reasons behind the improper result of DataWorks Plus's FRT use by the Detroit Police, I first have to define the collective responsibility shouldered by both DataWorks Plus and the DPD towards the citizens of Detroit who will be affected by the use of the FRT tool. This obligation is to the citizens, to implement caution and care in the use of a new technology—its use will only lead to overall benefit for citizens without enacting any undue harm to any law-abiding citizens. This means the DPD implements independent review of technologies when in the process of acquiring access to them; to ensure fairness and equity of the technology's usage on citizens. This review must be a continuous process, evolving as new reports come out. Because similar machine learning driven technologies have shown bias in the past these potential imperfections in the software should be known, and by not conducting an independent review the DPD has a causal connection to this moral responsibility and thus bear blame under the joint responsibility. If the DPD had taken on this collective responsibility they could have halted the use of Face Plus after the NIST report in December of 2019 found many FRT algorithms to have biases in "age, sex, race, and other demographic properties" (Grother et al. 2019). On the side of DataWorks, the failure to uphold this collective responsibility is evident in their lack of instruction in running the system, exemplified by Pastorini's statement "We, [DataWorks Plus], don't tell our customers how to use the system. There's already law enforcement policies" (Koebler, 2020). This inadequacy of information unfairly transfers the responsibility to find the biases and risks in using the system contained in the collective responsibility off to the DPD, risks that only DataWorks can foresee

and has the freedom to disclose and thus need to shoulder this moral responsibility of providing guidelines to the users of their technology. This moral responsibility to disclose risks and provide guidance in the use of their FRT tool is necessary for DataWorks in this newly recognized collective responsibility. If either or both sides had assumed responsibility for the joint obligation outside of their individual charges, this moral failure could have been largely prevented, as the DPD would have had a better understanding of the system and could have designed more effective protocols surrounding the use of the FRT tool.

The Role of the Federal Government

One of the most common calls from those questioning the controversial use of FRT in law enforcement has been for the federal government to step in, with an expectation for either a blanket ban or strict regulation on use of the technology, with a bill being introduced in June of 2020 (Solon, 2020). This point, while quickly solving the immediate issue of FRT in the context of law enforcement, fails to understand the roles and responsibilities face individually and jointly by DataWorks Plus and the Detroit police. Regulation and bans would remove any biased algorithms from the market, and provide clear direction for how law enforcement should use this technology. However, with the growing speed of development in technology spurred by explosion of Artificial Intelligence it is almost certain that controversial technologies like FRT appear more frequently in law enforcement and outside settings. Without recognizing the deeper issues of the problem of many hands, these new technologies could fall into the same pitfalls as FRT, doomed from the start by not having awareness of the collective responsibility to the citizens of any city, state, or country. Placing a ban or intensely strict regulation only pushes the problem down the road, when the next innovation creates similar dilemmas, and by

understanding the problem of many hands, law enforcement and developers can combat their occurrence by jointly accepting the collective responsibility.

Conclusion

I have now analyzed the case of Detroit police incorrectly arresting Robert Williams due to use of flawed FRT technology. I used ANT to develop a network in which to observe the interplay of the responsibilities of both the DPD and DataWorks Plus. Through this analysis I showed that both had met an individual moral obligation, however the technology still created an ethical breakdown to the citizens of Detroit. This was caused by an ignorance towards the collective responsibility both parties had in this network, and this lack of awareness is what brought about this misdeed. By showing that in situations where the moral responsibility is shared and amplified, I have shown there is a need for all parties with any individual duty need to accept a role in the joint obligation. The reader will be better equipped to think about how to assign and distribute blame in the failure of the DPD's use of Face Plus, and can further revise this application of the problem of many hands, to better understand and mitigate deficiencies in the future use of emerging controversial technologies by law enforcement agencies.

Word Count: 4167

References

About Face*:* Examining the department of Homeland Security's use of facial recognition and other biometric technologies, part II: Hearing before the Committee on Homeland Security (H. R. hrg) 116th Cong. (2020) (Testimony of Charles H. Romine)

Cressman, D. (2009). A brief overview of actor-network theory: Punctualization, heterogeneous engineering & translation. *CPROST Catalogue Number: 09-01*, 1-12.

Garvie, C. (2020, June 24). *The untold number of people implicated in crimes they didn't commit because of face recognition. American Civil Liberties Union.* https://www.aclu.org/news/privacy-technology/the-untold-number-of-people-implicated-in-crimes-they-didnt-commit-because-of-face-recognition/

Grother, P., Ngan, M., & Hanaoka, K. (2019). *Face recognition vendor test part 3: Demographic effects* (NIST IR 8280; p. NIST IR 8280). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8280

Harwell, D. (2019, April 30). Oregon became a testing ground for Amazon's facial-recognition policing. But what if Rekognition gets it wrong? *Washington Post*. Retrieved March 16, 2021, from https://www.washingtonpost.com/technology/2019/04/30/amazons-facial-recognition-technology-is-supercharging-local-police/

Hill, K. (2020, August 3). *Wrongfully accused by an algorithm*. The New York Times. Retrieved from https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html

Koebler, J. (2020, June 29). *Detroit police chief: Facial recognition software misidentifies 96% of the time.* VICE. Retrieved from https://www.vice.com/en/article/dyzykz/detroit-police-chief-facial-recognition-software-misidentifies-96-of-the-time

Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2021). Corporate digital responsibility. *Journal of Business Research*, *122*, 875–888. https://doi.org/10.1016/j.jbusres.2019.10.006

Martinez-Martin, N. (2019). What are important ethical implications of using facial recognition technology in health care? *AMA Journal of Ethics*, *21*(2), E180–E187. https://doi.org/10.1001/amajethics.2019.180

Rahal, S., & Hicks, M. (2020, June 26). *Detroit police work to clear record of man wrongfully accused by facial software*. Detroit News. Retrieved from https://www.detroitnews.com/story/news/local/detroit-city/2020/06/26/detroit-police-clear-record-man-wrongfully-accused-facial-recognition-software/3259651001/

Revised facial recognition directive. (2019, September 19). *Detroit Police Department manual*. Retrieved from https://detroitmi.gov/sites/detroitmi.localhost/files/2019-09/Revised%20facial%20recognition%20directive%20transmitted%20to%20Board%209-12-2019.pdf

Rowe, E. A. (2020). Regulating facial recognition technology in the private sector. *Stanford Technology Law Review*, *24* (1).

Solon, O. (2020, June 25). *Facial recognition bill would ban use by federal law enforcement*. NBC News. Retrieved from https://www.nbcnews.com/tech/security/2-democratic-senators-propose-ban-use-facial-recognition-federal-law-n1232128

van de Poel, I., & Royakkers, L. (2011). *Ethics, technology and engineering: An Introduction.* Blackwell Publishing