Reframing Algorithmic Bias with Cultural Clothes in Public Facial Recognition Technology: Improving Public Goods through a Private Sector Lens

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelors of Science, School of Engineering

Joshua Matthew De Vera
Fall 2023

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.

Advisor

Kathryn A. Neeley, Associate Professor of STS, Department of Engineering and Society

STS Research Paper

Introduction

> "Above all, we will need more public goods and more positive externalities. Star Trek teaches us that humanity's wondrous inventions do not fully realize their potential until they are freely shared."
>
> - Manu Saadia, Trekonomics: The Economics of Star Trek

In the 21st century, biometric technology has become the new standard of security. The rise of biometric technology through facial recognition can be attributed to the need for stronger security measures and its ease of use due to the individuality of everyone's faces. As a result, facial recognition has been integrated into most smartphones, tablets, and security systems. This is warranted due to its 95% accuracy in product testing across many large companies such as Apple, Google, and Amazon (Leslie 2020, pg. 2). In these commercial products, the accuracy is advertised as very precise and causes its users to place a lot of trust in the reliability of facial recognition.

Although the accuracy of facial recognition has been proven by many private companies throughout the start of the 21st century, when this software is used in the public spheres of education and police systems, minority groups, commonly those of color, have lower accuracy levels around 70-80% (Perkowitz 2021, pg. 5). Furthermore, cultural clothes such as hijabs and turbans cause the accuracy of facial recognition to decrease even more. The technical cost and the social implications of misidentifying these individuals will grow alongside facial recognition's utilization which emphasizes the need to address it before the cost to reverse it becomes too high. Despite the marketed high accuracy of facial recognition technology, algorithmic bias and decreased accuracy occur when it is developed as a product made for the public sector. With this technology becoming an integral part of security systems, the misclassification of minority groups who make up a majority of its users increases the computing cost of recognition failures (Singh et al. 4, 2020). As facial recognition is utilized more in developing countries and within the United States, it is imperative that it is able to account for all types of users and properly identify all individuals.

In recent years, researchers have recognized this fault of facial recognition and created bias reducers on top of algorithms to increase accuracy. However, the accuracy is still significantly lower for affected populations, it leads to "16% increase in [computing] cost", and it works around the problem of bias adjusters (Kalinga 2021, pg. 7). There are companies that are able to maintain high accuracy without the use of bias adjusters which brings attention to the reason why facial recognition developed in the public sector has significantly less accuracy compared to those developed in the private sector. Developers of facial recognition in both sectors have the same skills and technology available to them, but underlying organizational and cultural differences lead to a decrease in software quality and accuracy. In analysis of the development of facial recognition as a public versus private good, the differences between the two sectors provide insight on how to improve facial recognition in the public sector and technology that is integrated into society as a public good. A novel approach to facial recognition development as a public good can promote quality by increasing accuracy and decreasing bias through a business model perspective.

Problem Definition: Despite software engineers with the same level of expertise developing facial recognition technology in the public and private sector, there is a large gap in accuracy with public goods having significantly less accuracy than its private counterparts.

Facial Recognition Interdependence in Society

In 1965, the first facial recognition algorithm for computers was created by Woody Bledsoe, Helen Chan Wolf, and Charles Bisson to recognize the human face based on a few basic facial features or landmarks as they're called in the industry. Over half a century later, their work has evolved beyond the large computer rooms that would produce results after a few hours to embed facial recognition software that exists in almost every personal device. As researcher David Leslie states, "facial recognition technology has grown at an unprecedented rate in the 21st century and widespread in the industrial and commercial sectors." (Leslie 2020, pg. 1) Facial recognition was originally developed as a form of security as its unique biometrics are specific to one person, but software developers have quickly expanded the technology's use beyond that scope in the private and public sectors. As a result, facial recognition has become a large factor for security in society. In recent years, it has been integrated in the educational and judicial/police systems to "track students in classrooms and potential criminals", respectively (Wehrli 2022, pg. 3). Facial recognition technology has become a part of all personal devices and transcended from commercial goods to services being utilized for the public. Facial recognition technology plays a large role in society and therefore needs to be reliable for all users and the systems that depend on it.

Facial recognition is used in over 50% of federal police systems and has a growing presence in the education system as an anti-cheating mechanism (Wehrli 2022, pg. 2). These two systems have influence over a large portion of society as a form of security and law enforcement for all citizens and a foundation for future generations. As facial recognition technology is integrated more into these important societal structures that interact with multiple communities, people assume that it should be accurate enough to support these diverse communities. This is emphasized by Sidney Perkowitz, a scientist at Emory University, who argues that "[facial recognition technology] should be able to support all members of society" (Perkowitz 3-4, 2021) in a fair way under the algorithm. With the scale at which it is deployed in society through private products and newly added public services, facial recognition software should be able to reproduce the same results with users and identify everyone as unique individuals. In personal devices such as the iPhone and Google Pixel, their facial recognition software has an accuracy around 97% (Kolla 2023, pg. 3). Even through the high accuracy of personal devices, it is important to acknowledge the technical limitations of how facial recognition technology is deployed. Cavazos notes that the software is "deployed in small, portable devices with limited technological capabilities which result in computing power tradeoffs." (Cavazos et al. 8) Facial recognition software is embedded in small personal devices in private devices and cameras in police systems. Despite these limitations, facial recognition in personal devices demonstrates that it can still yield highly accurate results. However, when facial recognition is deployed in the public, there is a clear bias in accuracy towards light skin-toned males compared to darker skin-toned females and those with cultural clothes.

Facial Recognition's Failure as a Public Good

Facial recognition exists in both the private and public sector as goods used to identify a diverse set of users. A public good is a commodity or service that is provided without profit to all members of society. An important point to emphasize is that it is designed for all members of society and thus must be able to support the diverse nature of any population. In the case of facial recognition technology, the public sector, compared to private products, has significantly decreased accuracy of facial recognition in minority and underrepresented populations and emphasizes the need for a new approach to facial recognition technology. Khalil argues that the software available to the commercial sector fails in comparison to those used in the field due to the size and computing power of commercial devices and the reports given to the public fail to emphasize the scale of inaccuracy among minority groups (Khalil et al. 2020, pg. 7). In Figure 1, the accuracy of different facial recognition technology is mapped across different skin tones and genders. In all of these, there is a clear bias in accuracy towards light skin-toned males compared to darker skin-toned females.

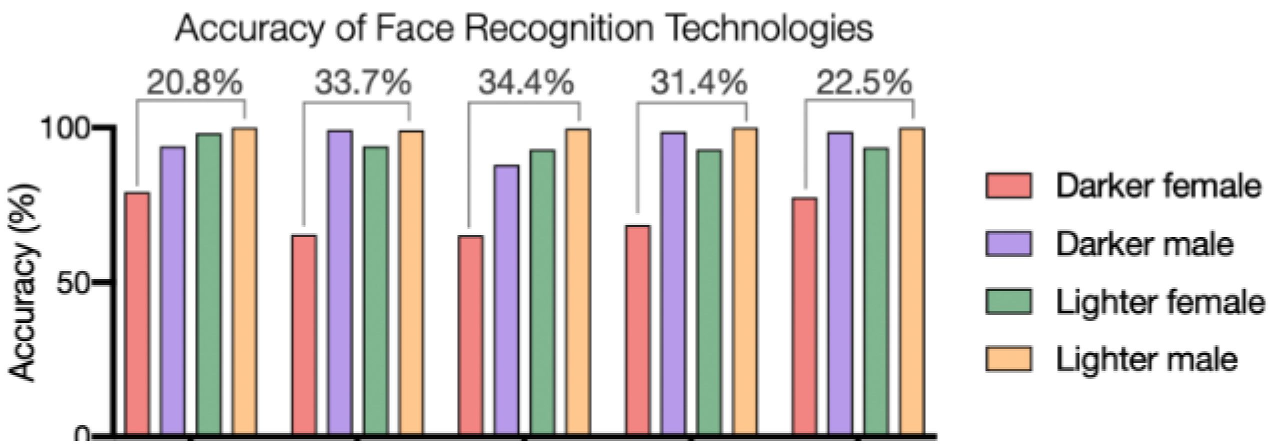

Figure 1. Accuracy of five different facial recognition softwares used as public goods testing against gender and skin tone (Khalil et al. 2020, pg. 5)

As facial recognition's impact on society grows, there is a need for software bias to be mitigated across different cultural communities especially when deployed on a public level. This data is further emphasized in Perkowitz's paper where they state African-Americans and Hispanic populations have decreased accuracies in facial recognition compared to their White counterparts in both field and commercial testing (Perkowitz 2021, pg. 6). All these groups have different facial landmarks that are more or less prominent depending on race which demonstrates the need for software that can differentiate individuals beyond these landmarks.

Despite these decreased accuracies when deployed in public devices, facial recognition technology has high accuracy rates for all races in private products. In Figure 2, Apple released an ad with the iPhone X showing off their Face ID algorithm which is able to detect faces despite changing appearances.

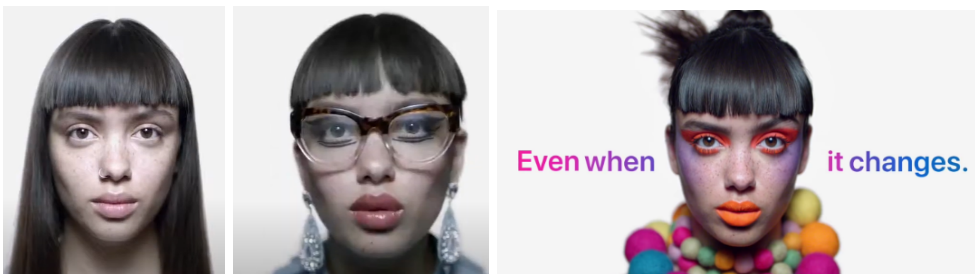

Figure 2. Apple iPhone X Ad, Introducing Face ID, demonstrates Face ID recognition algorithm that is recognize faces even when they change (Apple 2017)

This ad demonstrates that when facial recognition is applied as a private product, it can yield high accuracy over a range of faces even when facial landmarks change. Data also backs up Apple's Face ID algorithm with accuracy around 97% across all races (Kolla 2023, pg. 3). Even though the same software engineers develop facial recognition algorithms for the public and private sectors, there is still algorithmic bias when applied in public goods. This emphasized the question on what leads to facial recognition technology developed in the private sector having higher accuracies across multiple races compared to the public sector.

In the late 2010s, facial recognition technology has introduced bias reduction algorithms in order to reduce bias in devices used in public systems, but the recognition software contains inherent bias through technological and organizational factors. Bias adjusters acknowledge that there is algorithmic bias and make an attempt to reduce bias in facial recognition, but this solution leads to the cyclic problem of whether bias adjusters have bias as well. Private companies like Apple are able to achieve high performance without the bias adjusters commonly found in public facial recognition despite the same resources being available to software developers in both sectors. As a result, Khalil and many software engineers have called for a new way to approach facial recognition technology (Khalil et al. 2020, pg. 6). Although not developed yet, it would remove the additional computing cost to reduce bias and allow smaller systems with limited power to yield high accuracy without technical limitations to facial recognition. This demonstrates the need for a novel approach that is emphasized by current software developers for facial recognition as a public good in order to develop it with accuracy that is impartial to differences in diverse user groups.

In Mesthene's Book, *Technological Change: Its Impact on Man and Society*, Chapter 3: Economic and Political Organization describes the dynamic and differences between private and public goods and users. Furthermore, Mesthene provides analysis on how the two are developed differently. Mesthene observes trends in history that negatively impact public goods, recognizes the technical needs in order to approach a resolution, and emphasizes a changing social outlook in order to innovate technology's development in the public sector. In a similar methodology, facial recognition's development can be analyzed to determine the source of algorithmic bias and divergence in accuracy between the two sectors.

All facial recognition developers have the same accessibility to software resources barring financial costs, but there are differences in company structure and office culture that lead to decreases in software quality because of a lack of a metric for social benefit. Mesthene's methodology can be applied in the development of facial recognition to extract the specific features that can decrease the gap in accuracy between the two sectors. Understanding these differences and connecting the two sectors can help create a plan to reduce bias and provide an approach to developing public facial recognition technology for the betterment of society.

Research Approach: Analyzing Differences in Developing Facial Recognition in the Public vs Private Sector

Developing Facial Recognition in the Public Sector as Private Good Increases Accuracy

As stated by Mesthene in his book, *Technology and Culture*, public goods usually start off as private goods that eventually transition into goods or services that become provided to all members of society as the government acquires the monetary means to support its creation. The ideation and evolution of products is often made in the private sector because of the clear economic benefits to the company. Mesthene provides multiple examples that demonstrate this trend that repeatedly occurs throughout history. The transportation system, which used to be accessible to only the elite of society, is now a form of public transportation and the education system which used to be available to rich men to free education for all children from Kindergarten to 12th grade. However, when public buses were deployed in the public sector, there were few tools available to quantify the economic or social benefits of implementing them for all users (Mesthene 1970, pg. 65). This is known as the system of feedback and reward which is very different from the perspective of public and private goods. Analyzing the differences in visibility and types of feedback and rewards from developing technology in the public vs private sector is an important factor when determining the differences in accuracy between the two.

Another point Mesthene introduces is that in these transitions, the scale at which the product is transformed and the change in societal powers in play requires “built-in efficiency that will lead to technological possibilities going unexplored for long periods of time” (Mesthene

1970, pg. 73). In the case of the public education system, the social effects of discrimination delayed the enrollment of women and other races in the United States for centuries and has led to decreased societal benefit despite being a public good. Therefore, the ideology of integrating technology into the public sector focuses on keeping what is already known to work instead of creating new approaches to a larger-scale and diverse group of people that it will affect. Since facial recognition software in private goods is already known to have significantly higher accuracy compared to public goods, analysis of the differences between the two products can lead to insight on where the differences originate from and how to develop a new approach to facial recognition in the public sector.

To understand the different influences that affect the development of facial recognition in the public sector, it is important to understand the political factors of the algorithm in two different sectors. Mesthene notes that political institutions go beyond the stigma towards the word politics in the 21st century. Rather than associating political institutions with Republican versus Democratic political parties, he defines political institutions as "all the decision making structures and procedures that have to do with the allocation and distribution of wealth and power in society." (Mesthene 63) Beyond the government institutions that affect the policy of public goods, politics also encapsulates the economic institutions that fund these goods and the large community of field experts who inform decisions made in the process. By addressing the economic, democratic, and technological systems in play, insights on where there are differences and how to develop a novel approach to developing facial recognition as a public good will rise.

As demonstrated by Mesthene, the lack of monetary incentives and a metric to define social benefit in public goods decreases quality compared to its private sector counterparts. He states that "bias in the economy generally in favor of development and satisfaction of private goods [results in] a relatively slow rate of innovation in the public sector." (Mesthene 1970, pg. 72) In the development of transportation, private interest took precedence over public usability because of the monetary incentive of its rich end users. For the public sector, the goods and services produced don't directly benefit the developers and there isn't a method to describe the benefit to users in society if they invest in higher quality. In Figure 1, Mesthene continues to emphasize that this "system of feedback and reward is speedy and visible" for private goods because there is a direct relationship between profits and technological innovation (Mesthene 1970, pg. 72). Mesthene visualizes the lack of incentives in the development of public goods which he states is the reason why the difference in quality is so drastic between the two sectors of society.

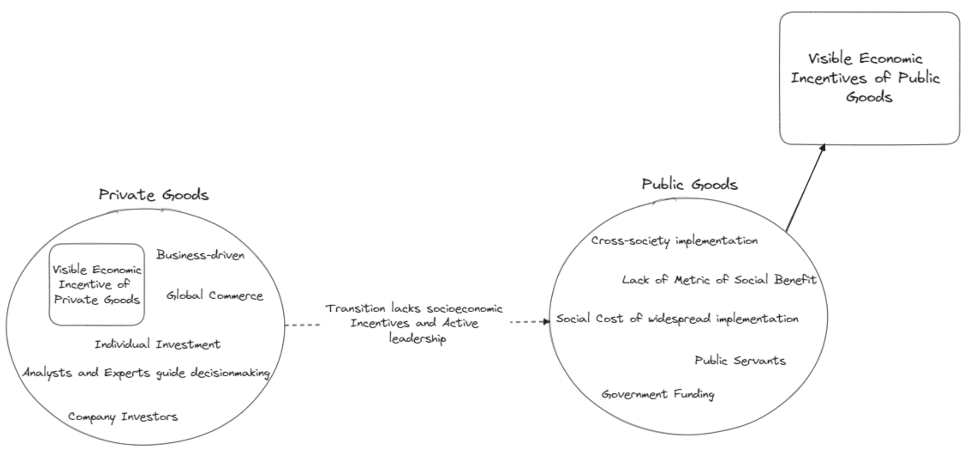

Figure 3. System of Feedback and Reward in the Private and Public Sectors. This graphic illustrates the factors that lead to the discrepancies between public and private goods and the visible socioeconomic factors in play from each domain (Created by Author)

After recognizing that there needs to be a change in approach to developing technology for the public sector, Mesthene calls for a way to measure its benefits for stakeholders. Compared to private goods, it is hard to quantify the marginal benefit to society based on societal responses alone. Therefore, in order to bridge the gap in quality goods and services between the public and private sector, there needs to be an economic incentive for all public workers to foster innovation and higher quality comparable to private firms.

Utilizing Mesthenes's approach in the system of feedback and reward to understand the underlying organizational model and company culture, the analysis of facial recognition as a public and private good can provide insight on the gap of accuracy between the two sectors. Firstly, the financial budgets of each system provides insight on technical limitations and personal incentive throughout the entire company. Secondly, the metric of benefit for creating each system for their respective end users quantifies the return on investment that the organization and members experience. Finally, an organization structure analysis breaks down development from a human perspective and explains choices made from an individual to organizational level. By analyzing these aspects of the two systems in the development of facial recognition technology, the divergence in accuracy can be identified and resolved.

Results: The differences in product development ideology coupled with significantly lower visible economic and social benefit causes facial recognition technology in the public sector to have decreased accuracy and innovation.

In the development of software, the organizational and financial factors have a large role in its development. From the CEO who runs the company, the officers who oversee different departments, the representatives who extract requirements from customers, the managers who lead tasks, and the engineers who actually build the product, each level of the organization contributes to the development of a product. Therefore at each step, an individual has the ability to change the final product. In Mesthene's perspective, it is important that there are experts at all levels who are able to make decisions based on their specialty in order to drive innovations and technological excellence. However, these technical and organizational factors are often limited by financial factors. Without the proper equipment, resources, and marginal return on investment, there is a lack of incentives and resources to innovate and develop products causing the smartest engineers to be limited in their capabilities. Utilizing Mesthene's methodology to compare the difference in system between the development of facial recognition technology in the private and public sector, this research will bring attention to the organizational and financial factors beyond the known technical limitations.

Monetary Incentives Drive Innovation and Accuracy

For any software to be developed, there is an associated financial plan that considers the cost and profit and determines the trajectory of the final output. For each new feature or increase in quality, there are increased costs to develop it and also produce it. Therefore, when considering the quality of facial recognition and innovation that leads to increased accuracy, it is important to know the financial aspects that are associated with each product. In figure 4, there is a comparison of costs in developing Face ID in Apple's iPhone X versus facial recognition technology in the US Immigration and Customs Enforcement Agency due to the Stamps Act in 2022.

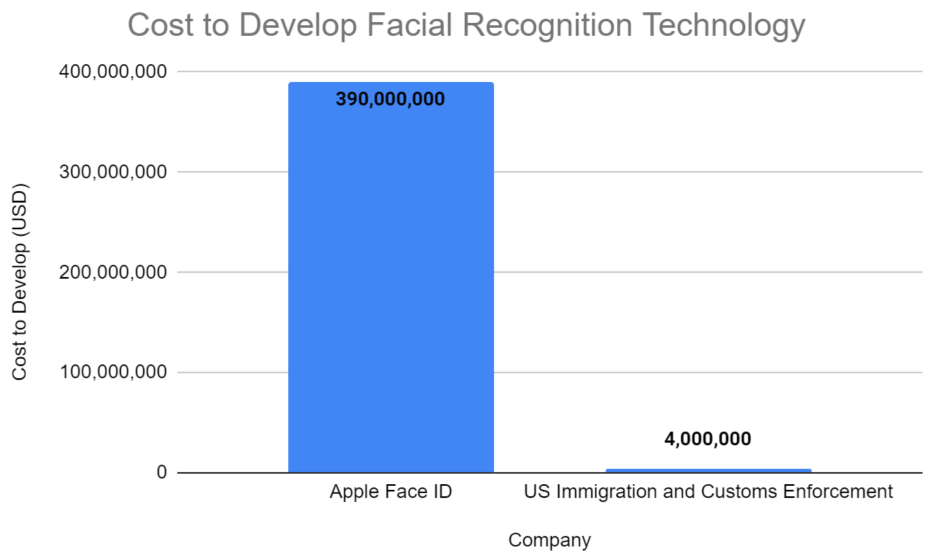

Figure 4. Cost to Develop Apple's Face ID for the iPhone X (2017) and Cost to Develop Facial Recognition Technology for US Immigration and Customs Enforcement (2022)

In the graph, there is a clear difference in funding when facial recognition is developed as a private good rather than for use in the public sector. The lack of financial limitations allows the software engineers who develop the algorithm to concentrate on technology quality and increased accuracy.

Likewise, the money can be used to innovate beyond the existing model of facial recognition technology. Figure 2 previously introduced in the paper demonstrates that Face ID for the iPhone X has 97% accuracy even when the user changes their facial features including haircuts, piercings, hair color, makeup, and skin color (Kolla 2023, pg. 3). With the large budget allocated towards Face ID, engineers at Apple and other private companies went beyond the average facial recognition technology accuracy at the time and utilized resources to create a product with the highest accuracy at the time. Compared to the software used by Apple, the US Immigration and Customs Enforcement has about 1% of the budget and consequently causes their software to have accuracy around 80% with gaps between dark skinned females and light skinned males as high as 34.4% (Khalil et al. 2020, pg. 5). This financial limitation leads to decreased accuracy and no room for innovation. In developing facial recognition technology for the public, this relationship demonstrates that there needs to be more monetary resources allocated to its development.

Beyond the cost of improving facial recognition technology accuracy, the marginal profit from selling it incentivizes software quality. For the iPhone X, the quarterly revenue after the release of the iPhone X jumped to $88.3 billion USD (Apple 2018). The profit margin from releasing the iPhone X with quality facial recognition technology resulted in a significant increase in revenue that incentivizes Apple to create quality products. Compared to the profits in Apple, the US Immigration and Customs enforcement doesn't directly profit from using the facial recognition technology and there currently isn't a metric to determine the societal benefit or profit makes some of the rewards invisible to the organizations tasked with making it. Therefore, in order to design facial recognition for public use, a profit system, whether monetary or socially, must be created in order to uncover the true benefit of having quality facial recognition technology in the public sector. Once there is a metric to determine a true estimate to the benefit to society including the intangible features such as reduced errors, decreased computing cost, technological innovation, and societal inclusion of all populations, algorithmic bias in facial recognition technology can decrease and improve its utilization for public good.

Leadership Affects the Direction of Facial Recognition Development

The governing powers in private firms contain experts in facial recognition development to inform decisions while the policy makers of the federal scale often lack the domain expertise in the development of facial recognition technology. In private firms, the CEO contains some domain knowledge and has a staff of individuals who know the product better than them to lead innovation within the company. Furthermore, CTOs, directors, product managers, and engineers all influence the development of facial recognition technology in the private sector and can make design decisions that can yield the best results. In the public sector, the democratic systems that define how facial recognition software is developed lack the same amount of domain expertise in the decision making process. When analyzing the 535 members of Congress in 2022, only nine members have engineering backgrounds or domain expertise in software development (Congressional Research Service 2022 pg. 4). Therefore, only 1.7% of Congress has an educated understanding of software development and yet these individuals vote and determine how facial recognition software and other public goods are created through government contracts. In public good contracts where requirements are made by these policy makers, the lack of experts in facial recognition technology in the process limits developer innovation and closes off possible technical capabilities to upper management. Despite the software developers having the ability to create facial recognition software comparable to private goods, the democratic systems in place for developing private goods limits the resources and requirements available to them when developing new software and decreases innovation.

The differences in incentives, company culture, and organization structure between public and private development of facial recognition technology provide insight on the non-technical sources of algorithmic bias. The differences in personal benefit for facial recognition developers and ability to innovate solutions at every step should be transitioned into the development of facial recognition as a public good. These features not present in the current development ideology allow for technical limitations to take precedence over innovation and lead to algorithmic bias. Therefore, creating a clearer system of feedback and rewards for the benefit of quality facial recognition in the public and developer innovation without organizational limitations can enable increased accuracy and benefit for civilians affected by its implementation.

Conclusion

Facial recognition technology has the ability to affect millions of users as it is integrated more into society as a form of security in the United States. Despite the ease of use of facial recognition that takes seconds to identify individuals, there is a complex algorithm made by software engineers that takes a lot of computational power in order to create accurate results. As this technology is developed in the private and public sectors, the accuracy between the two sectors diverged as facial recognition as a public good has significantly decreased accuracies for minority groups with darker complexion. Utilizing Mesthene's methodology, analysis on the differences between the two developments of the same facial recognition technology provides insight on how to bridge the gap between the two.

After analyzing facial recognition that is developed in the private and public sector, the most prominent differences are the financial incentives and visibility of the system of feedback and rewards, company culture that promotes quality, and organizational structure that enables innovation. Most evidently, private companies have an obvious system of feedback and rewards from their customers with profits that directly benefit the company. Likewise, the organizational goals of private companies such as Apple, Google, and Amazon are aligned from all levels from the CEO to the software developers which enable facial recognition quality and accuracy. These two main factors influence the technical aspect of facial recognition development which emphasizes the importance of software engineering ideologies when approaching this problem. The same software developers create facial recognition software in the private and public sector which emphasizes the need to address the non-technical differences that decrease accuracy.

Therefore, a private sector system of visible rewards and space for software developer innovation should be employed into the public sector in order to bridge the gap of facial recognition quality and accuracy. This environment will enable developers to produce the best algorithm and potentially eliminate algorithmic bias due to the scale at which facial recognition will be deployed in the public.

In analyzing the differences in the development of facial recognition technology, the analysis can provide insight into the development of all public goods in the future. As seen in the case of facial recognition, the organizational discrepancies in the federal system and the invisible system of rewards and public benefit metric causes a decrease in facial recognition quality and accuracy. Utilizing elements for development seen in its private counterparts can yield improved results. Although some engineers argue against the capitalist approach of public goods, this ideology and environment can enable more innovation due to a larger user base with diverse backgrounds. The results found in this analysis can be applied to the development of all goods and services by fostering an ideology for engineers to produce quality products through a more visible metric of societal good. By implementing and testing the validity of private company practices in developing public facial recognition software, this new approach can be further expanded into engineering principles that affect the development of all goods and services implemented in the public.

References

A. Khalil, S. G. Ahmed, A. M. Khattak and N. Al-Qirim. (2020). "Investigating Bias in Facial Analysis Systems: A Systematic Review," in Institute of Electrical and Electronics Engineer Access, vol. 8, pp. 130751-130761, 2020, doi: 10.1109/ACCESS.2020.3006051.

Congressional Research Service. (2022). *Membership of the 117th Congress: A Profile*. https://doi.org/https://crsreports.congress.gov/product/pdf/R/R46705

Gwyn, T., & Roy, K. (2022). Examining Gender Bias of Convolutional Neural Networks via Facial Recognition. *Future Internet*, *14*(12), 375. MDPI AG. Retrieved from http://dx.doi.org/10.3390/fi14120375

J. G. Cavazos, P. J. Phillips, C. D. Castillo and A. J. O'Toole, "Accuracy Comparison Across Face Recognition Algorithms: Where Are We on Measuring Race Bias?," in *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 1, pp. 101-111, Jan. 2021, doi: 10.1109/TBIOM.2020.3027269.

Kalinga, A. T. N. (2021). *Towards An Ethnic-Bias Free Approach For Facial Recognition* (Doctoral dissertation).

Kolla, M., & Savadamuthu, A. (2023). The Impact of Racial Distribution in Training Data on Face Recognition Bias: A Closer Look. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 313-322).

Leslie, D. (2020). Understanding bias in facial recognition technologies. *arXiv preprint arXiv:2010.07023*.

Perkowitz, S. (2021). The Bias in the Machine: Facial Recognition Technology and Racial Disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing*, *Winter 2021*. https://doi.org/10.21428/2c646de5.62272586

Salvador, T., Cairns, S., Voleti, V., Marshall, N., & Oberman, A. (2021). Bias Mitigation of Face Recognition Models Through Calibration. *arXiv preprint arXiv:2106.03761*.

Singh, R., Agarwal, A., Singh, M., Nagpal, S., & Vatsa, M. (2020). On the Robustness of Face Recognition Algorithms Against Attacks and Bias. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(09), 13583-13589. https://doi.org/10.1609/aaai.v34i09.7085

Singh, R., Majumdar, P., Mittal, S., & Vatsa, M. (2022). Anatomizing Bias in Facial Analysis. *Proceedings of the AAAI Conference on Artificial Intelligence*, *36*(11), 12351-12358. https://doi.org/10.1609/aaai.v36i11.21500

Vangara, K., King, M. C., Albiero, V., & Bowyer, K. (2019). Characterizing the variability in face recognition accuracy relative to race. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0).

Waelen, R.A. The struggle for recognition in the age of facial recognition technology. *AI Ethics* 3, 215–222 (2023). https://doi.org/10.1007/s43681-022-00146-8

Wang, A., Liu, A., Zhang, R. *et al.* REVISE: A Tool for Measuring and Mitigating Bias in Visual Datasets. *Int J Comput Vis* 130, 1790–1810 (2022). https://doi.org/10.1007/s11263-022-01625-5

Wehrli, S., Hertweck, C., Amirian, M. *et al.* Bias, awareness, and ignorance in deep-learning-based face recognition. *AI Ethics* 2, 509–522 (2022). https://doi.org/10.1007/s43681-021-00108-6

Xu, T., White, J., Kalkan, S., Gunes, H. (2020). Investigating Bias and Fairness in Facial Expression Recognition. In: Bartoli, A., Fusiello, A. (eds) Computer Vision – ECCV 2020 Workshops. ECCV 2020. Lecture Notes in Computer Science(), vol 12540. Springer, Cham. https://doi.org/10.1007/978-3-030-65414-6_35