Science & Technology Innovation Program

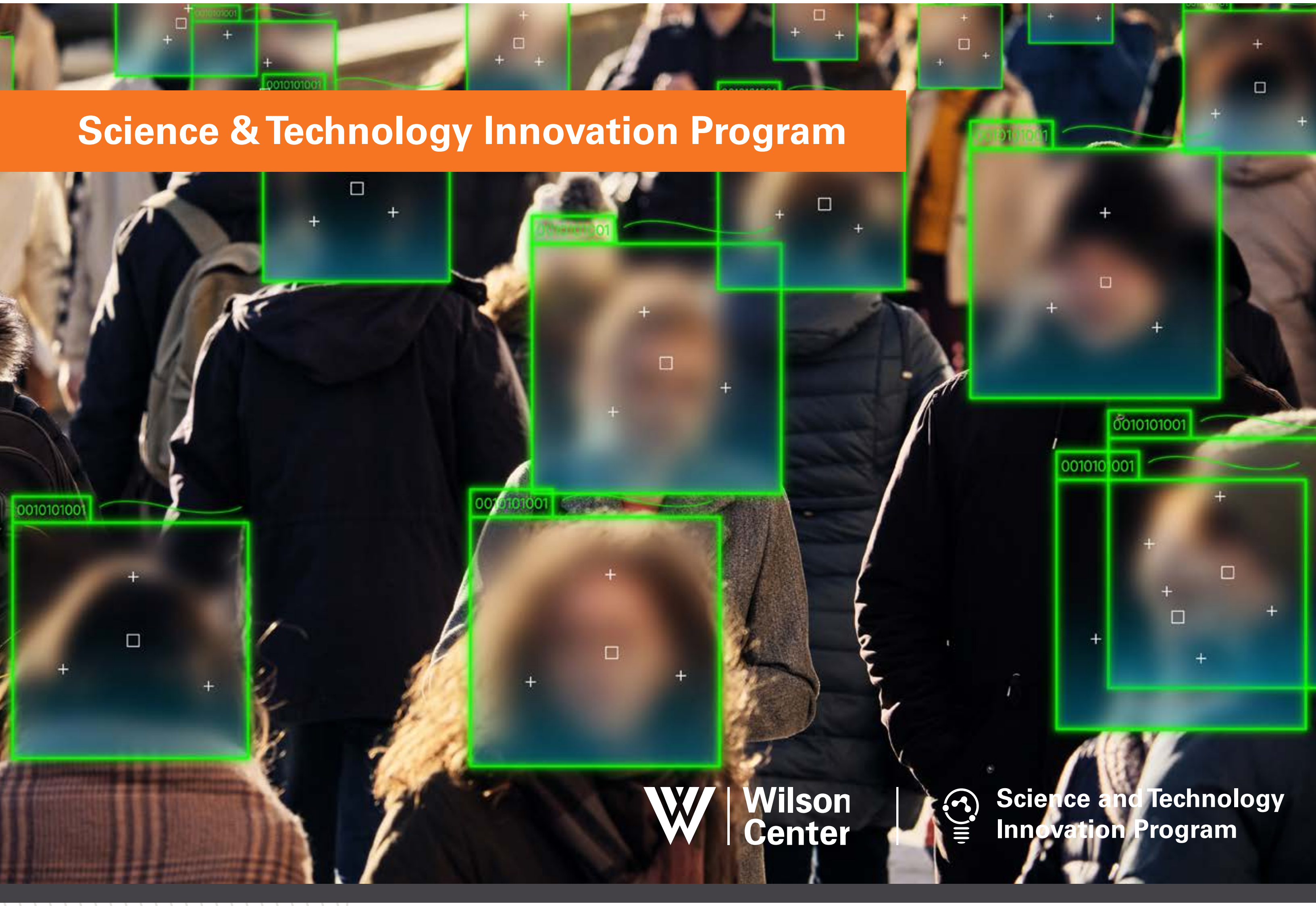


Image source: By ilikeyellow / Shutterstock.com

Authors

Anne Bowser
Meg King
Alie Fordyce
Andrew Carmona

Broadening the Conversation around Facial Recognition: Lessons from the Consumer Perspective

November 2021

Advances in artificial intelligence (AI) promise to improve our quality of life, including applications that offer increased safety and security. This promise extends to biometric technologies, including facial recognition technologies (FRT)—but academia, the private sector, civil society, and government are all flagging ethical concerns.

Much of today's public and policy discussions focus on mitigating harms of FRT in contexts like criminal justice. Analyzing such cases can help us understand where more attention is needed and, potentially, what pitfalls to avoid. However, focusing exclusively on harms cannot shed light on whether, or how, to leverage FRT to benefit society.

Consumer technologies offer the opposite side of the coin. Studying consumer applications can help elucidate the broad benefits of facial recognition, and specific use cases for how FRT can improve quality of life. Research on the consumer perspective can also help us understand what leads consumers to consider these technologies ethical. Words like "transparency" and "disclosure" are, according to the Organization for Economic Cooperation and Development, high-level principles of open AI. Yet what approaches do consumer FRT applications leverage to ensure that users have the information they need to trust an application? Understanding how to design FRT systems to ensure transparency, disclosure, and trust is necessary for any policy action on ethical FRT, in the consumer context, and more broadly.

Introduction to Facial Recognition

Facial recognition is a specific type of biometric technology that shares some similarities with, but also differs from, facial detection and facial analysis machine learning (ML) technologies. Facial detection determines whether a digital image contains a human face. It is a type of object detection, and the first step in more complex facial recognition models.

Once a face is detected, non-biometric "facial analysis" technologies can assess characteristics such as sex, age, or emotion. Generally, the term facial recognition technology (FRT) refers to an automated or semi-automated biometric identification process where a probe image of a face is analyzed and compared to one or more templates, or reference images that mathematically represent a human face. In one-to-one matching, a template from a probe image of an individual is compared to a template of a single individual. These systems are used for verification, such as unlocking a smartphone or passport authentication. In one-to-many matching, a probe image is compared to many templates compiled in a database or gallery to explore whether a potential match exists.

These technologies can make two types of errors (Grother, Ngan, Hanaoka, 2019). In Type I errors, or "false positives," one individual is incorrectly associated with another. In Type II errors, or "false negatives," an FRT system will fail to identify a match when one exists. Many ethical concerns revolve around fairness, including when biased algorithms lead to disproportionate Type I errors for specific demographic groups. In response to these concerns, numerous federal and state-level bills aiming to address bias and fairness concerns by imposing restrictions and requirements have been proposed. The Biometric Information Privacy Act (BIPA) has been codified into law in the State of Illinois. In addition, some companies—including Microsoft, Amazon, and IBM—have announced voluntary moratoria or longer-term bans around the development and use of facial recognition technologies by police.

Bias is not the only ethical concern. The Perpetual Lineup, a project led by Georgetown Law's Center on Privacy and Technology, focuses on privacy violations, including when the data used by FRT applications are pulled from sources like Department of Motor Vehicles (DMV) records. Security is also an important consideration.

This paper focuses on understanding and analyzing a range of consumer facial recognition technologies. We scoped an initial working definition of "consumer technology" to include FRT applications where:

1. Reasonable steps are taken to make consumers aware that they are participating in a facial recognition system; and,

2. Consumers can meaningfully "opt out" of participation, and have access to reasonable alternatives if they choose to do so.

Following interviews with 16 experts at nine institutions, primarily private companies, we showcase a range of consumer FRT applications to understand high-level application domains and more specific use cases. This is important for painting a comprehensive picture of FRT, and identifying opportunities in addition to risks.

Our conversations revealed both potential vulnerabilities and a range of risk management and mitigation strategies. While not all strategies will be appropriate for all purposes, documenting these experiences can help private and public sector stakeholders think critically about the ethical use of facial recognition technologies in and beyond the consumer sector.

We conclude with a list of future technology and policy opportunities for elevating the safe and ethical use of FRT, offered from the perspective of our interviewees.

Setting the Scene: An Overview of Consumer FRT in 2021

In many cases, FRT is simply the next step in a historically analogue process. An early biometric technology, fingerprint, was first used as evidence during a landmark 1910 murder case. Within the U.S. government, the National Institute of Standards and Technology (NIST) has been working on biometrics including fingerprint, face, iris, voice, DNA, and multimodal technologies for over 60 years. NIST has collaborated with other agencies, including the Federal Bureau of Investigation (FBI), since the 1960s.

Many of the companies we interviewed have a lengthy history with these technologies, working in FRT and related biometrics for decades. They believe that the demand for accuracy and speed is increasing. They have demonstrated that significant performance gains have recently been made, a finding also documented by NIST, who reported a twenty-fold performance increase between 2014 and 2018.

One important change is the scale of activities that new technologies enable. For example, FRT can support real-time observation at speeds and scales that are not possible with historically analogue technologies like fingerprint or standard CCTV cameras that do not aim to identify individuals in the video footage they capture. This potential elevates a range of ethical concerns, many of which also touch on broader AI questions, including ethical data collection and use.

While some companies focus solely on consumer technology, others additionally (or primarily) develop applications for government use. In addition, while many companies focus solely on biometric FRT, a few also develop broader facial analysis applications. Companies sell to U.S. consumers, and internationally. A wide range of use cases and diverse applications exist:

- **Enterprise safety and security**. Facial recognition technology provides a secure authentication process that allows trusted parties in while keeping unauthorized actors out. Biometric authentication provides access to physical locations or virtual spaces, including during remote onboarding processes. For each resource accessed, authentication can happen once, or on an ongoing basis.

Source: By Blue Planet Studio / Shutterstock.com

- **Personal safety and security**. Devices ranging from doorbells, to personal computers, to mobile phones all support biometric authentication through facial recognition. This capability prevents unauthorized access to sensitive information on a device that is lost or misplaced and facilitates easier and safe access to homes for caregivers to ill or elderly patients.

Source: By Zephyr_p / Shutterstock.com

- **Transportation.** FRT is used at airports by government authorities like Customs and Border Patrol (CBP) and by private companies to verify identity, in some cases without the need for physical credential like a driver’s license. Automobiles, including both traditional and autonomous vehicles (AVs), are using FRT for safety, security, and personalization. Some “Smart Cities” are beginning to use biometric identification systems including FRT on public transportation to replace traditional fare cards.

Editorial credit: By 1000 Words / Shutterstock.com

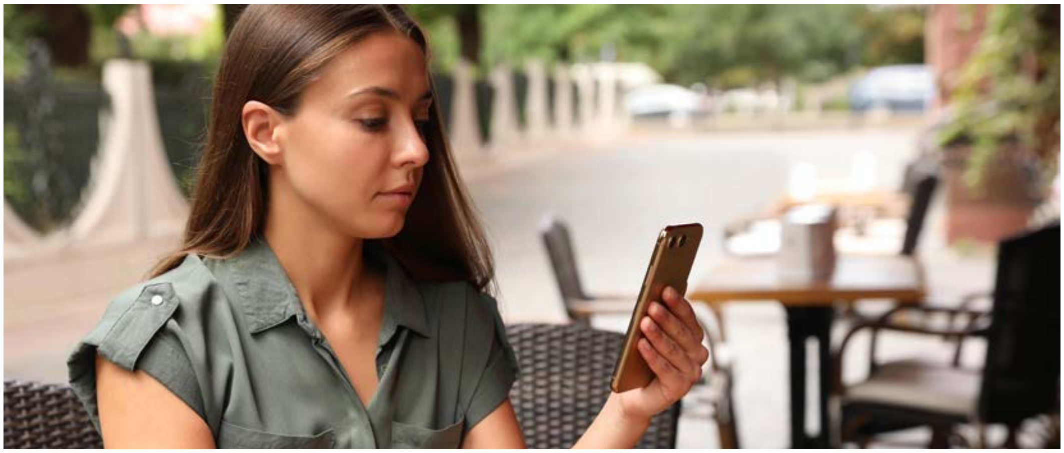
- **Health and medicine.** Biometric facial recognition can help ensure that the right patient receives the right care, including by ensuring the appropriate delivery of medical supplies. Non-biometric facial analysis applications, along with related computer vision technologies, can also be helpful for early detection of injury or illness (including stroke).

Editorial credit: By Iokanan VFX Studios / Shutterstock.com

- **Customer experience, hospitality, and retail.** Facial recognition can support more convenient and more personalized services. For example, some companies offer face matching on smart home devices so the device can recognize specific users and provide services on its screen such as personalized calendars and reminders. Similarly, resorts can reduce the need for customers to carry credit cards or keys while providing a personalized experience based on established preferences, particularly for VIP guests, while offering the added benefit of contactless hotel stays critical during the COVID-19 pandemic.

Editorial credit: By 1aslysun / Shutterstock.com

- **Creative and playful experiences.** A number of digital products, such as applications for artistic expression, social media apps, and video games, use a mix of biometric and non-biometric facial recognition and analysis techniques. These help support different social functions, like the recognition and tagging of friends, or simply make experiences more fun and engaging.



Editorial credit: By New Africa / Shutterstock.com

Companies offer a range of technical FRT solutions, including 2D and 3D recognition, and tools for still images and video footage. They also use a number of strategies to develop and sell facial recognition algorithms and, in some cases, access to databases for matching.

- **Software development kits (SDKs).** A software development kit is a set of tools that helps software developers create custom applications for their systems. Investing in these building block components allows companies to focus on their "core" offering—typically an algorithm. While customers who buy or license SDKs appreciate the ability to build custom applications to meet unique business needs, companies who produce SDKs have little control over how their algorithms are used, including over processes for selecting training data.
- **Direct access to existing systems or algorithms.** Many companies provide direct subscription access to facial recognition software. As with other strategies, most direct access models require an end user to "enroll" in a facial recognition program by first verifying their identity, and then providing one or more images of their face. This model is used by general purpose providers, such as Microsoft Azure's Cognitive Services Face API. Other companies provide access to software systems, but also to databases.
- **Access through channel partners.** Some companies develop facial recognition algorithms specifically for hardware providers, including manufacturers of personal computers or Internet of Things (IoT) devices. Working directly with a channel partner helps FRT companies customize their applications to the unique capabilities of a hardware device, thus ensuring quality control. Some companies are moving from selling through channel partners to direct access models, in part due to ethical concerns around who the ultimate end user might be.
- **Custom development or "co-design" of new systems or algorithms.** A number of companies build customized facial recognition systems to meet unique customer needs. Customization allows companies to consider and design for important aspects of system deployment, such as the specific hardware devices used. This approach also allows companies to select training data that matches environmental conditions, target demographics, and other controls. Companies can also customize user interfaces to ensure that workflows meet established good practices, and monitor completed systems to ensure compliance and provide technical support.

The software algorithms powering face detection, facial analysis, and facial recognition can be run in two places: on a hardware device or in the cloud. Each of these approaches is associated with different strengths and weaknesses.

Companies who run algorithms via on-premise devices often consider user privacy a primary concern. If a hardware device with an embedded camera captures a photo during an enrollment process, it can immediately translate that photo into a template and potentially even discard the initial image. This template is then saved on device, and matched to templates from later probe images. This process prevents the user's image from ever leaving the capture device and minimizes stored data (because templates do not contain enough information to facilitate reverse-engineering the original face images), a safeguard designed to reassure end users that their privacy needs are respected and met.

Companies who run algorithms on the cloud do so for a number of reasons. Chief among these is ensuring the use of facial recognition services can be tracked. Monitoring helps companies ensure compliance with stated use policies, and provides information for audits on an as-needed basis. In most cases, images can be translated into templates before being sent to the cloud, mitigating some—but not all—user privacy concerns.

The question of which strategy is more secure depends on the use-case. Running algorithms on an on-premise device provides users with higher levels of control. Running algorithms on the cloud takes control away from the user and places it in the hands of a cloud provider. When best practices for security like encryption are followed, this may be an asset. However, compared to on-device storage, cloud storage may be easier for certain bad actors to find, and security breaches do occur. Furthermore, cross-border data flows can pose compliance challenges and different vulnerabilities.

Best Practices for Consumer FRT

In FRT, AI, and beyond, there is a consistent tension between the need to embrace general, high-level ethical principles on one hand, and to recognize unique, contextual challenges and requirements on the other.

Many companies have codified high-level ethics around the use of facial recognition technology. One company uses principles based on global privacy law. Another embraces "tenets" that help developers think through the values behind complex tradeoffs, such as collecting less demographic information to protect participant privacy, or collecting more demographic information to help ensure equitable performance across different groups.

Codes of conduct are another common ethics support tool. These can be used internally, and may be shared with external audiences. These codes may outline the types of FRT applications a company will develop, the types of partners a company will work with, or otherwise limit how their technology can be used. In some cases, companies believe codes of conduct are a business asset that help create trust and foster good practices: "customers do not want subjective ethical choices to be able to decide when the technology can be used." Each of these methods could be evaluated as the basis, in part, for industry best practices.

But not all ethics are universal, even among western democracies. One company headquartered in Asia generally follows domestic norms, though also takes into account key U.S. policy developments. Exact ethical requirements may also differ based on the specific use case, where "duty of care is proportional to reach and impact."

Ethics can be challenging to translate into practice, and difficult to enforce. Some companies rely on contractual requirements to ensure adherence, including through pass-through terms, which stipulate that all future parties or secondary customers must agree to similar contractual agreements. Companies also leverage approaches like time- or device-limited licenses, where a technology ceases to work if certain requirements are violated.

Voluntary Limitations on the Ethical Use of FRT

Multiple companies draw hard lines around the types of applications they will develop. Major use cases deemed "categorically unethical" include the use of non-biometric facial analysis for predicting gender or sexual orientation. For some, these uses of facial analysis applications should be considered "pseudo-science," and raise quality considerations as well as ethical concerns. Many companies also decline to develop biometric FRT technologies for use in real-time surveillance or predictive policing.

None of the companies we interviewed have an indiscriminate sales model; all vet their partners before beginning a business relationship. In addition, some decline to do business in countries that have questionable track records regarding cybersecurity or human rights. This noted, it can be difficult for companies to navigate the ethics around an American company doing business on foreign soil, for example in an airport setting.

Finally, companies may limit how flexible technological solutions are developed and deployed. Codes of conduct may stipulate that humans must remain "in the loop," for example by requiring that an authenticated expert review and approve suggested matches. Training requirements can be used to limit who can be authenticated, and therefore has access to a FRT system.

One major shared understanding transcends this diversity of applications, use cases, and business strategies. All the companies we spoke with characterize facial recognition as an "assistive technology," or "a tool, an indicator: and it should be used this way." As an assistive technology, the goal of FRT is to provide an objective recommendation that can help compensate for human limitations, including fatigue and the tendency to multi-task. Many interviewees noted that, when properly developed and implemented, automated facial recognition can also help mitigate ethical issues like human bias.

There were mixed perceptions regarding whether FRT is appropriately used as assistive technology. Some of the companies we interviewed also marketed their technology for use by government, and believe that misinformation exists on how FRT is commonly used in this domain, often asserting FRT requirements given by government clients are often more restrictive than the requirements offered by private sector partners. Others emphasized that over-reliance on AI recommendations is still a significant concern.

Ethics Across FRT Design and Development Processes

A number of steps to ensure the ethical use of FRT can be taken during various stages of design and development processes, including planning, development, deployment, and testing and validation. These lifecycle specific opportunities are complemented by crosscutting considerations, including around privacy and security.

Planning

Many of the companies we spoke with critiqued FRT providers who make their technology available "to anyone with a credit card," or "let anyone use it, [without] background." Generic technologies can easily be misused, and are associated with performance limitations. For some companies focusing on a niche market, the use case around a new product may be similar enough to existing systems that models can simply be tweaked or tuned. In other cases, "to get 90% [or higher] accuracy, you have to build things from scratch."

Customizing an existing algorithm or creating a new solution requires developing a deep understanding of a particular use case and broader application domain. For example, tolerance for different types of levels and errors is context specific: in access control, Type I errors (false positives) are "a disaster," while false negatives are an "inconvenience." There are also variations in the level of acceptable error rate, for example between a FRT application used to enable better customer service in a retail setting, and an application that airlines use to facilitate secure, efficient check-in and boarding for an international flight.

Planning requires documenting technical requirements, including hardware-software compatibility. The quality of a probe image in relation to clarity, resolution, contrast, and lighting is partially determined by the configuration of hardware including capture cameras. In many settings, infrared cameras are preferable to RGB devices, 3D cameras have higher accuracy than 2D versions, and active illumination can improve quality, particularly for darker toned faces. Attention to camera placement can also mitigate quality concerns. From one company's perspective, a best-case scenario is one where "[Customers] give us the hardware, we start developing the FRT so our algorithms get the maximum amount of data under the intended use cases." When this is not possible, iterative development processes can evaluate the links between hardware requirements, algorithmic parameters, and training data.

Across all phases of the development process, but particularly during planning, consulting with relevant stakeholders through "user-centered" or "participatory" approaches is an ethical ideal. Some companies recognize that not enough consultation takes place, especially when development processes follow the "typical engineering style: we just did it." Others are using approaches from human factors research, including "deep ethnographic methods." Increasingly, companies use user-centered and participatory processes to understand not just the immediate users of an FRT, but also "people downstream," including "the ones at risk to harm."

Development

The development of new facial recognition algorithms—like many other AI applications—is an iterative, cyclical process where key parameters are tuned, data are collected, and models are trained. Across these development processes, access to comprehensive, high-quality data is an important limiting factor that can lead to bias and other quality concerns. Data that are easily available may generally over-represent faces associated with a particular sex or Fitzpatrick skin type, and additional gaps may exist regarding data from a particular target population.

While incorporating additional training data can help mitigate these concerns, all training data must be procured "legally and through appropriate means." However, many outstanding questions remain about what constitutes legal and appropriate means. One interpretation of "legally and through appropriate means" means, when possible, using data from users who consented, or "opted in," to having their data used for training purposes. During enrollment, some companies encourage users of a specific FRT applications to explicitly opt in to general training databases, while almost all enable participants to opt out to inclusion in general training databases. Beyond information collected through enrollment, data may be publicly accessible but still ethically fraught. For example, Tweets are "consented" in the sense that a user is actively contributing content to the Twitter platform that may include a face. But, Twitter users do not explicitly consent to data sharing for purposes of FRT system training, and—in many cases of large-scale data scraping—was not given a meaningful opportunity to opt out of inclusion. As a related consideration, it can be difficult or impossible to meaningfully withdraw consent once an individual's face is used to train a FRT system.

The challenge of whether to prioritize reducing bias or protecting privacy can cause an ethical or even legal conflict. IBM was subject to a class action lawsuit for BIPA violation when images in a 2019 dataset, "Diversity in Faces," were traced to the social media platform Flickr. Two of the companies we spoke with cited this lawsuit as important in their decision to err on the side of user privacy over collecting additional data to help support bias mitigation techniques.

Beyond training data, development processes also offer opportunities to design systems so humans remain "in the loop." Good user-experience (UX) design is critical to meeting this goal. The careful design of user interfaces can clarify FRT's appropriate role of an assistive technology, including by portraying an automated recognizer as "an assistant," or requiring human verification as an important step in any identification workflows. User interfaces can also help enforce certain requirements. For example, some interfaces might only display matches above a certain threshold to prevent users from acting on dubious recommendations. Other interfaces may refuse to perform a search if the probe image quality is too low. Other companies offer broad-brush assessments such as "high confidence" or "no confidence" to avoid creating a false certainty and reinforce the value of human verification. Still others simply provide a rank candidate list without percentage or confidence assessments.

UX can also act as a "tripwire," showing where uses might require additional policies or guidelines that are required to ensure the ethical use of FRT systems. Beyond support for ethical interfaces, additional checks and balances can be put into place during development processes. Many companies create systems that document, for audit purposes, who is using a system and how, including through a record of user logins or searchers.

Deployment

Around deployment, at least two user groups are considered: the end users responsible for using a FRT application to submit a probe image and/or evaluate a potential match, and the subjects of FRT systems, or people whose biometric data are collected and evaluated.

Most of the companies we spoke with consider dedicated support for end user training a necessary practice, and include training requirements in their contractual agreements. At least two types of training are offered: support for effectively conducting a face recognition search, and general guidance on the appropriate use of a FRT system. Training is helpful for users in consumer contexts, and critical for those working on government use cases like criminal justice. Good training can help the users of a facial recognition application understand when and how much to trust a match result or a similarity score between probe image templates and gallery image templates, and—equally important—when not to trust a match result or similarity score.

Many companies have also thought carefully about the ethics surrounding consent. Beyond consent that takes place during enrollment, such as the decision to make a probe image broadly available as training data, users must explicitly opt in to any participation in FRT systems. For many software systems, a subject typically consents to certain conditions offered through a user agreement, but this process may not actually secure informed consent, especially if "no one reads terms and conditions." Some companies believe that posting a sign or making an announcement may be more effective at actually informing consumers about facial recognition practices, especially in public spaces such as airports and shopping centers.

Testing and Validation

In general, training and development cycles will continue until acceptable accuracy results are achieved. Additional alpha and beta testing can help identify bugs and refine systems before a product is released. After release, real-world use identifies more opportunities for improvement through additional development processes, and subsequent versions are released to the public through periodic updates or new product roll-outs.

Third party evaluation and assessment processes are also used. NIST's well-regarded facial recognition vendor testing (FRVT) program, measures various aspects of performance. These include performance across demographic groups, an assessment important to identify potential sources of bias, and—more recently—performance for photos of masked individuals, an important quality consideration that arose due to COVID-19. While NIST's testing program is voluntary (and could use a significant boost in funding and personnel to increase capacity), many of the companies we interviewed routinely submit their algorithms to NIST for testing, and cite their high performance in marketing conversations.

Testing is typically conducted to evaluate potential sources of error and opportunity for improvement. While often led by companies, additional evaluations can happen when voluntary watchdogs assess performance, especially on facial detection and facial analysis algorithms that are not submitted to NIST testing. Perhaps most notably, research led by the MIT Media Lab previously demonstrated the poor performance of different commercial facial analysis algorithms in regard to different interrelated aspects of an individual's identity, such as sex and skin tone. In addition to work that evaluates algorithms based on performance, the practice of making training data or source code available as open access resources offers additional opportunities for peer assessment.

Cross-Cutting Considerations

Privacy and security are two important issues that transcend development processes.

To protect user privacy, many companies only store the templates derived from images captured during enrollment, quickly discarding the images themselves. This good practice can happen regardless of whether templates are stored on a device or in the cloud. Offering transparent policies around data retention for training data as well as probe images can also help mitigate privacy concerns. Good policies should not only clarify collection and retention of images used in FRT systems, but also how data is (or could be) linked to other personal and personally identifying information (PII).

Many privacy considerations are linked to security concerns, particularly if sensitive data are vulnerable to breaches. As with other technologies, encrypting data transmitted by FRT systems is an important best practice to follow. Some companies augment encryption with tools such as two-factor authentication, including through other biometrics like fingerprint and iris, which is often used for primary identity verification. Some companies also believe that "uniqueness is a security feature," including when algorithms are customized to a particular, proprietary template. While this practice may not make a single facial recognition application more secure, it does mean that neither the template nor the algorithm could be used in other FRT systems, limiting the implications of a breach from crossing vendors or products.

In addition to security and privacy, numerous ethical issues emerge across multiple stages of development and deployment. To summarize, these include:

- **Bias**. In line with the philosophy of "garbage in, garbage out," securing access to representative training data is an important first step in bias identification and mitigation. Bias can also be evaluated during testing and evaluation processes, including through NIST's FRVT program. In addition to technical bias, considering contextual bias is also useful. Even if an FRT system does not consider or make assumptions about race, for example, it could be used to prevent physical access to a private organization thereby perpetuating bias and discrimination.
- **Consent**. Training data must be secured through ethical processes, and many companies ask subjects to opt in to broader data sharing practices during enrollment in biometric systems. Slightly different issues related to consent arise when systems are deployed, particularly in public spaces, where subjects might not explicitly opt in through behaviors like purchasing a personal device.
- **Transparency**. Providing notice and, where possible, securing informed consent is a necessary first step towards transparency. Our interviewees also advocated for broader public transparency around data collection and sharing policies, along with information on how and for what purposes different facial recognition systems are being used. Testing is an opportunity for companies to be transparent around performance, particularly as some believe that "the code doesn't need to be visible, the process of creation and use does."
- **Accountability**. Ensuring the ethical use of facial recognition requires being able to identify the source of a problem when something goes wrong. Some companies track the use and potential misuse of FRT applications for verification and audit purposes. While these practices can ensure accountability for specific applications, many of the people we spoke with believe that additional, policy-based measures are required.

Next Steps: Technology

Our interviewees see some opportunities for new technologies to create better FRT systems.

Innovations in hardware, such as the development of higher-quality cameras at lower price points, can increase quality and accuracy. Many techniques, like the use of LiDar for 3D imaging, have already established value and may become increasingly common. Such hardware innovations are particularly promising for cameras embedded in devices like mobile phones and IoT applications. A second important innovation associated with IoT—the development of increasingly smaller microprocessors, or chips, to power a wider and more diverse range of devices—will also enable the commercial use of facial recognition to proliferate.

Many companies are investigating how to integrate FRT with other biometric technologies, including iris recognition. On one hand, integrated biometrics can provide an additional layer of security and verification. On the other, increased complexity and data collection may be associated with elevated privacy and security concerns.

While these and other technology developments can be helpful, interview participants stress that much more work is needed on the policy front.

Next Steps: Public Policy

For the companies we spoke with, there is widespread agreement that enhanced government and policy attention to FRT, including and beyond consumer systems, is both necessary and desirable.

What Policy Guidance is Required?

Interviewees are looking for enhanced guidance specific to FRT systems, and generally applicable to AI.

In FRT, there is a widespread belief that self-regulation is a risky game," particularly around "the appropriate extent to use the technology... we look for those playbooks constantly." Recognizing that approaches such as bans can be harmful and unrealistic, general, high-level guidance should address how FRT technology should be ethically used, including by whom and for what purposes. Such guidance may establish different requirements for public and private spaces, and should address key, cross-cutting ethical questions including "what constitutes legal and appropriate data collection and use?" More specific information would be helpful regarding "how to implement different use cases," including requirements for keeping humans in the loop during different processes and workflows.

More general guidance on ethical AI could include consensus around "a precise definition of transparency and fairness" along "holistic standards" and information on how to evaluate, measure and implement these ethical constructs, or "put them on a scale." Such guidance will be helpful to developers of FRT systems, and can be informed by FRT use cases and/or include FRT-specific recommendations or examples, but should also be applicable to non-FRT applications of ML and AI.

Guidance is also needed around data protection. This should include general guidelines for protecting the privacy of individuals. It should also touch on how and under what conditions data should be shared, including among different government entities, and between government stakeholders, private sector partners, and members of the general public. Some companies also believe it would be particularly helpful for the government to "define acceptable provisions for data re-use outside of narrowly consented processes."

More work is needed to establish ethical standards and practices, including for consumer FRT and beyond. While Congress has the important authority to establish legal guidelines, other federal stakeholders also have a role to play in clarifying appropriate use. The Department of Commerce (DOC), for example, can help companies understand where, with whom, and under what conditions to safely and ethically do business, a question linked to "complex trade considerations." The Government Accountability Office (GAO) has already done important work on consent for Customs and Border Protection (CBP), which might be a model adaptable for consumer use cases. A range of federal agencies can issue guidance on acceptable use in application areas that fall under their jurisdiction, such as the National Institutes of Health (NIH) for healthcare.

NIST can play a bigger role in setting standards and other guidelines for FRT software systems. NIST could also invest in making representative data sets that are demographically diverse, though some companies believe more information on NIST's ground truth validation processes would be helpful. Some non-sensitive data could be shared as an open public asset, and more restricted data could be offered to trusted vendors willing to adhere to appropriate protocols, including those that prevent sensitive data from leaving a particular warehouse.

Agencies including the Department of Homeland Security (DHS), which already conducts Biometric Technology Rallies, can bring additional, much-needed authority and expertise on capture devices and other hardware concerns and can begin or continue conducting operational testing of FRT systems. Private sector companies also have high levels of knowledge, influence, and therefore responsibility. They can be "actively involved" in discussions around good practices, especially those related to consumer FRT, though public sector authorities are still required to establish and restrict inappropriate use.

One hurdle to more effective regulation is education. Multiple companies believe there are "not enough people in government who understand technology." The absence of widespread knowledge intensifies other challenges, but can be mitigated through education and training programs.

A second hurdle is the increasingly fragmented landscape of global and state-specific regulations. In the absence of global alignment, many companies are tracking key policy developments in a range of countries. They recognize both "good actors to learn from," such as Singapore and Japan, and countries to be "cautious about," including China and Russia. They also recognize that some norms are emerging. For example, Germany has a law that autonomous vehicle manufacturers must prioritize harm to property over harm to people. Equally concerning is the patchwork quilt of different laws and policies emerging in various U.S. states, suggesting that national-led, globally aware action is urgently needed.

Finally, any work on ethical guidelines, standards, and good practices must be accompanied by incentives for implementation. These may include high-level requirements set by Congress, along with more limited requirements leveraging (for example) agency-specific procurement policies and processes. Other incentives might align with market forces, for example by using testing and certification processes to place companies with good performance on preferred vendor lists, or to offer other incentives to companies that help create, or comply with, newly established best practices.

Many of the necessary puzzle pieces are already on the table. The wide range of positive use cases demonstrated by the consumer sector, along with examples of good practices that have been identified and implemented, show that it is possible to develop safe and ethical FRT applications. Action from the public policy community can help cement and build on progress to date, and ensure that facial recognition technology, in and beyond consumer devices, continues to contribute economic benefit while creating social good moving forward.

Acknowledgments

We are grateful to all the individuals who contributed to this research by sharing their experience during the interview process.

This work is part of a larger research program called Beyond Bans. The Wilson Center maintains full independence over the research process and results.

References

Buolamwini, J. (n.d.). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. MIT Media Lab. Retrieved October 12, 2021, from https://www.media.mit.edu/publications/gender-shades-intersectional-accuracy-disparities-in-commercial-gender-classification/.

Facial recognition: Microsoft Azure. Facial Recognition | Microsoft Azure. (n.d.). Retrieved October 12, 2021, from https://azure.microsoft.com/en-us/services/cognitive-services/face/.

Georgetown Law Center on Privacy and Technology. (n.d.). *The perpetual line-up. Perpetual Line Up*. Retrieved October 6, 2021, from https://www.perpetuallineup.org/.

Grother, P., Hanoaka, K., & Ngan, M. (n.d.). *Face recognition vendor test (FRVT)* - NIST. NIST Publications. Retrieved October 6, 2021, from https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf.

Hansen, J. M. (2020, April 3). *Flashback: How fingerprinting made chicago famous: New technology led to 1910 murder conviction in a first for the nation*. chicagotribune.com. Retrieved October 6, 2021, from https://www.chicagotribune.com/opinion/commentary/ct-opinion-flashback-fingerprinting-clarence-hiller-slaying-20200403-jgihdoi7xfdmra7jqs3zbq2bqq-story.html.

Illinois General Assembly. (n.d.). 740 ILCS 14/ biometric information privacy act. Retrieved October 6, 2021, from https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57.

Ngan, M., Grother, P., & Hanaoka, K. (n.d.). *NIST*. Retrieved October 12, 2021, from https://pages.nist.gov/frvt/reports/facemask/frvt_facemask_report_6b.pdf.

Patricia.flanagan@nist.gov. (2020, June 7). *Biometrics*. NIST. Retrieved October 12, 2021, from https://www.nist.gov/programs-projects/biometrics.

Patricia.flanagan@nist.gov. (2020, November 30). *Face recognition vendor test (FRVT)*. NIST. Retrieved October 12, 2021, from https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt.

Robin.materese@nist.gov. (2018, December 6). *NIST evaluation shows advance in Face Recognition Software's capabilities*. NIST. Retrieved October 12, 2021, from https://www.nist.gov/news-events/news/2018/11/nist-evaluation-shows-advance-face-recognition-softwares-capabilities.

Sarah.henderson@nist.gov. (2020, May 18). *NIST study evaluates effects of race, age, sex on face recognition software*. NIST. Retrieved October 6, 2021, from https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software.

OECD. (n.d.). *OECD principles on Artificial Intelligence - Organisation for Economic Co-operation and development*. OECD. Retrieved October 6, 2021, from https://www.oecd.org/going-digital/ai/principles/.

Rizzi, C. (2020, January 30). *Class action accuses IBM of 'flagrant violations' of Illinois biometric privacy law to develop facial recognition tech*. ClassAction.org. Retrieved October 12, 2021, from https://www.classaction.org/news/class-action-accuses-ibm-of-flagrant-violations-of-illinois-biometric-privacy-law-to-develop-facial-recognition-tech.

What is an SDK? Red Hat - We make open source technologies for the enterprise. (n.d.). Retrieved October 12, 2021, from https://www.redhat.com/en/topics/cloud-native-apps/what-is-SDK.

WOODROW WILSON INTERNATIONAL CENTER FOR SCHOLARS

The Wilson Center, chartered by Congress in 1968 as the official memorial to President Woodrow Wilson, is the nation's key non-partisan policy forum for tackling global issues through independent research and open dialogue to inform actionable ideas for the policy community.

THE SCIENCE AND TECHNOLOGY INNOVATION PROGRAM (STIP)

The Science and Technology Innovation Program (STIP) brings foresight to the frontier. Our experts explore emerging technologies through vital conversations, making science policy accessible to everyone.

Woodrow Wilson International Center for Scholars
One Woodrow Wilson Plaza
1300 Pennsylvania Avenue NW
Washington, DC 20004-3027

The Wilson Center

- www.wilsoncenter.org
- wwics@wilsoncenter.org
- facebook.com/woodrowwilsoncenter
- @thewilsoncenter
- 202.691.4000

Wilson Center

STIP

- www.wilsoncenter.org/program/science-and-technology-innovation-program
- stip@wilsoncenter.org
- @WilsonSTIP
- 202.691.4321

Science and Technology Innovation Program