CYBERLAW CLINIC
HARVARD LAW SCHOOL | BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY

MEMORANDUM

To: Representative Stephanie Jerome

From: Harvard Cyberlaw Clinic

Date: January 5, 2024

Re: Biometrics and Data Privacy

1. OVERVIEW

Biometric data is usually defined as data arising from measurements of the human body, often used to personally identify people.[1] The theory of why biometric data is particularly high risk is that it is uniquely identifiable. If someone steals a person's credit card number or social security number, they can get new ones. But as the saying goes, "you can't replace your face." Biometric data's special identifiability means that collecting and using biometric data can implicate privacy, free expression, autonomy and dignity in significant ways. In recognition of the highly sensitive nature of biometric data, data privacy legislation usually contains a special section for biometric data.

This section describes how state, federal, and foreign entities have defined biometric data. Drawing on these examples, it then provides recommendations for how Vermont could incorporate biometric data regulations into a general data privacy law.

2. DEFINITIONS OF BIOMETRIC DATA

As of November 2023, thirteen US states have passed comprehensive data privacy laws, many of which address biometric data. Additionally, Illinois has a specific biometric privacy law.[2] Almost a dozen more have introduced data privacy bills in their respective state legislatures.[3] Each state defines biometric data slightly differently. At issue in selecting a definition is how to ensure we capture just enough data, and how to ensure that it won't become out of date. The below selection is meant to highlight distinct ways to define biometric data in statute.

2.1. Vermont 2023 Bill

The draft Vermont data privacy legislation that was introduced in 2023 contains the following definition of biometric data:

> "Biometric data" means data generated by automatic measurements of an individual's biological characteristics, such

[1] Computer Security Resource Center, *Biometrics*, National Institute of Standards and Technology, https://csrc.nist.gov/glossary/term/biometrics
[2] F. Paul Pittman, US Data Privacy Guide, White & Case (Dec. 26, 2023) https://www.whitecase.com/insight-our-thinking/us-data-privacy-guide; *Biometric Information Privacy Act (BIPA)*, ACLU Illinois, https://www.aclu-il.org/en/campaigns/biometric-information-privacy-act-bipa.
[3] *Which States Have Consumer Data Privacy Laws?*, Bloomberg Law (Nov. 27, 2023), https://pro.bloomberglaw.com/brief/state-privacy-legislation-tracker.

> as a fingerprint, a voiceprint, eye retinas, irises or other unique biological patterns or characteristics that are used to identify a specific individual. (B) "Biometric data" does not include: (i) a digital or physical photograph; (ii) an audio or video recording; or (iii) any data generated from a digital or physical photograph, or an audio or video recording, unless the data is generated to identify a specific individual.[4]

This definition identifies specific forms of biometric data, but also includes a residual clause for "other unique biological patterns or characteristics" that could encompass other forms of data. The definition of biometric data is limited to information that is "used to identify a specific individual." Finally, the definition excludes photographs and audio or video recordings, although (B)(iii) leaves open the possibility of data extracted from such sources to be biometric information, depending on how it is used.

2.2. Illinois

The Biometric Information Privacy Act (BIPA), an Illinois statute specifically addressing biometric information, defines biometric information as "any information . . . based on an individual's biometric identifier used to identify an individual."[5] A biometric identifier is in turn defined as follows:

> "Biometric identifier" means a retina or iris scan, fingerprint, voiceprint, or scan of hand or face geometry. Biometric identifiers do not include writing samples, written signatures, photographs, human biological samples used for valid scientific testing or screening, demographic data, tattoo descriptions, or physical descriptions such as height, weight, hair color, or eye color.

The definition goes on to list several exclusions, including donated organs, tissue, or blood; genetic information; medical data used for health care treatment, payment, or operations under HIPAA; and medical imaging such as X-rays or MRIs.

The BIPA definition lists six specific biometric identifiers and does not include a residual clause that could be read to encompass other forms of identification. The definition of biometric information is further limited to

[4] https://legislature.vermont.gov/Documents/2024/WorkGroups/House Commerce/Bills/H.121/Drafts, Amendments, and Legal Documents/H.121~David Hall~Draft 2.1, 5-22-2023~6-20-2023.pdf

[5] https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57

the use of a biometric identifier "to identify an individual." Combined with numerous exceptions, BIPA provides an example of a very narrow definition.

2.3. California

Under the California Consumer Privacy Act (CCPA), "biometric information" means:

> an individual's physiological, biological, or behavioral characteristics, including information pertaining to an individual's deoxyribonucleic acid (DNA), that is used or is intended to be used singly or in combination with each other or with other identifying data, to establish individual identity.[6]

The statute lists examples of biometric information, including iris, finger, voice, and face prints, but specifically notes that the list is not comprehensive.

2.4. Washington

Under Washington law, a "biometric identifier" means

> data generated by automatic measurements of an individual's biological characteristics, such as a fingerprint, voiceprint, eye retinas, irises, or other unique biological patterns or characteristics that is used to identify a specific individual.[7]

"Biometric identifier" specifically excludes photo, video, and audio recordings as well as health care information covered under HIPAA.

2.5. Connecticut

Connecticut law defines "biometric data" as

> data generated by automatic measurements of an individual's biological characteristics, such as a fingerprint, a voiceprint, eye retinas, irises or other unique biological patterns or characteristics that are used to identify a specific individual.[8]

---

[6] https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5#:~:text=(c)%C2%A0%E2%80%9CBiometric%20information,contain%20identifying%20information.

[7] http://app.leg.wa.gov/RCW/default.aspx?cite=19.375.010

[8] https://www.cga.ct.gov/2022/ACT/PA/PDF/2022PA-00015-R00SB-00006-PA.PDF

Like Washington, Connecticut specifically excludes data generated from a photograph, video, or audio recording, but adds a residual clause "unless such data is generated to identify a specific individual."

2.6. Federal Trade Commission's Policy Statement

While the Federal Trade Commission (FTC) has not released binding regulations regarding biometric information, it has released a policy statement.[9] In that statement, the FTC defines "biometric information" as

> data that depict or describe physical, biological, or behavioral traits, characteristics, or measurements of or relating to an identified or identifiable person's body

Under the FTC definition, Biometric information includes, but is not limited to, face, iris, fingerprint, voice, genetics, and gait or movement. The FTC definition expressly includes "data derived from [] depictions, images, descriptions, or recordings, to the extent that it would be reasonably possible to identify the person from whose information the data had been derived."

2.7. European Union Laws

Under the European Union's General Data Protection Regulation (GDPR), "biometric data" means

> personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic [fingerprint] data.[10]

A more robust definition can be found in the recent EU AI Act, which provides the following definitions:

> (33)'biometric data' means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data;

[9] https://www.ftc.gov/system/files/ftc_gov/pdf/p225402biometricpolicystatement.pdf
[10] https://www.legislation.gov.uk/eur/2016/679/article/4

(34)'emotion recognition system' means an AI system for the purpose of identifying or inferring emotions or intentions of natural persons on the basis of their biometric data;

(35)'biometric categorisation system' means an AI system for the purpose of assigning natural persons to specific categories, such as sex, age, hair colour, eye colour, tattoos, ethnic origin or sexual or political orientation, on the basis of their biometric data;

(36)'remote biometric identification system' means an AI system for the purpose of identifying natural persons at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database, and without prior knowledge of the user of the AI system whether the person will be present and can be identified ;

(37)''real-time' remote biometric identification system' means a remote biometric identification system whereby the capturing of biometric data, the comparison and the identification all occur without a significant delay. This comprises not only instant identification, but also limited short delays in order to avoid circumvention.

(38)''post' remote biometric identification system' means a remote biometric identification system other than a 'real-time' remote biometric identification system;

(39)'publicly accessible space' means any physical place accessible to the public, regardless of whether certain conditions for access may apply.

This definition is notable because it distinguishes between biometric identification systems, which are designed to identify a specific person, and biometric categorization systems, which place people into categories without necessarily identifying them.

3. RECOMMENDATIONS

3.1. Address both identification and categorization

The definition of biometric data should include biometric systems that identify *and* categorize people based on protected characteristics. Biometric systems have traditionally been defined as systems that uniquely identify a

person based on body-related measurements. But biometric systems are now designed not only to pick out a particular person from a crowd, but to sort a group of people into various categories. For example, systems may purport to categorize people by gender based on their face print,[11] or to determine whether someone is "criminal" based on their walking style ("gait"),[12] or interpret someone's emotional state or suitability for a job based on facial movements.[13]

This shift from identifying a specific person to categorizing groups of people is important because with identification, people generally agree that the task at hand is possible (is this person who they say they are? Who is this person?). But with categorization, the baseline task (what is this person's emotional state? What is their race? Are they "criminal"?) is in some cases not possible, or is highly contested. Because the task at hand is more ambiguous, there is a risk that cultural stereotypes will inform which categories people are sorted into, and will inform the process of categorizing.

Accounting for this risk, the FTC definition of biometric information notes, "In some contexts, the terms 'biometrics' or 'biometric technologies' have been used to refer specifically to technologies that are used to identify individuals. We use the term 'biometric information technologies' to refer to the broader category of all technologies that use or purport to use biometric information for any purpose."

This clarification in the definition of biometric data also serves a purpose of closing a loophole whereby vendors might avoid liability by collecting and using biometric data in ways that would pose the same risks as only to not personally identify people.

3.2. List examples rather than requirements

Each biometric privacy bill (or biometrics sections of general data privacy bills) define the scope of what biometrics are a little differently. Some are high-level descriptions, whereas others list examples. The most recent versions include both high-level descriptions and a non-exclusive list of examples.

---

[11] Os Keyes, *The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition*, Proceedings of the ACM on Human-Computer Interaction (2018).
[12] Coalition for Critical Technology, *Abolish the #TechToPrisonPipeline*, Medium (June 23, 2020), https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16
[13] Lisa Feldman Barrett et al, *Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements*, Psychological Science in the Public Interest (2019), available at https://pubmed.ncbi.nlm.nih.gov/31313636/; Ifeoma Ajunwa, *Automated Video Interviewing as the New Phrenology*, Berkeley Technology Law Journal (2021), available at https://papers.ssrn.com/abstract=3889454

Vermont should follow this example in an effort to be as clear as possible about what fits within the scope, while also ensuring that the list is non-exclusive and does not become out of date.

3.3. Consider whether photographs, video, and audio count as biometric data

Most state legislation specifically exclude photographs, video, and audio recordings from the definition of biometric data. But the recent FTC definition considers recordings (e.g. of a person's face or other stand-alone biometric data) as a form of biometric data. Conceptually, this makes sense because a photo a person's face could be used to create a face-print, just as a live-captured rendering a person's face would. The inclusion of photographs in the definition of biometric data might be desirable for instance, in the case of Clearview AI, which scraped photos of people's faces from Facebook in order to create a huge database of faces for its facial recognition tool.[14] Considering photographs biometric data would allow individuals to consent to this kind of data gathering, and would potentially prevent this sort of broad data collection in the first instance. Including photographs or other recordings also closes a logical loophole that might otherwise enable a bad actor to skirt regulation. If photographs are not included, a bad actor could simply take static images of people as they enter a public space, and store those images rather than record live images and capture face prints.

In practice though this would place compliance burdens on a lot more entities that hold photos of people's faces. Many entities collect photographs of people for non-privacy invasive purposes, ranging from an ID photo at a local gym, or a cafe that has polaroid pictures of customers on a bulletin board. To ensure that these types of data uses are not affected and that small entities are not burdened, the definition of an affected entity should be limited to entities beyond a certain customer size, or in instances where data is re-shared or sold to another party.

3.4. Include specific and narrow exemptions

One of the key questions we've discussed over the past couple of months is the difference between private and public sector, and whether certain public sector uses should be exempt from regulation. The line between public and private is not always crystal clear when it comes to biometric technologies. For example, a government actor may use a tool developed by the private sector,

[14] *See* Kashmir Hill, YOUR FACE BELONGS TO US (2023).

such as law enforcement's use of Clearview AI.[15] And a private actor, such as an individual homeowner or company operating a doorbell camera on their property may also capture video footage of the public sidewalk and may share this information with the police.[16]

Not only are actors sometimes a mix of public and private, but places or venues themselves can blur the line between a public versus private. For instance, a pharmacy operating a real time biometric surveillance system serves the public but is private property. The concept of a "place of public accommodation," as coined in civil rights legislation and the Amerians with Disabilities Act may be a useful way to categorize this type of venue.[17] Alternatively, the EU AI Act's prohibition of real-time remote biometric identification systems (i.e. real time facial recognition) applies in "publicly accessible spaces."[18]

Fully accounting for both the public-private blend of actors who use biometric data, and the public-private nature of places where biometric technologies are used means not limiting moratoria on certain biometric technologies to law enforcement, and on the flipside, not exempting public use of biometric data on the assumption that public use of biometric data is somehow less risky. Many biometric privacy laws include exemptions from prohibition or stringent compliance processes where biometric technologies are used to prevent imminent harm like terror attacks and to find victims of crimes, including missing children.[19] In designing exemptions, Vermont may want to consider specific narrow instances where the benefit of biometric technologies may outweigh the costs, rather than exempting large categories of actors or contexts in which biometric technologies are used.

3.5. Implement a tiered approach to regulating biometric data

Before attempting to design safeguards for various biometric technologies, it is worth distinguishing them based broadly on their function, possible issues,

---

[15] James Clayton & Ben Derico, *Clearview AI Used Nearly 1m Times by US Police, It Tells the BBC*, BBC (March 27, 2023), https://www.bbc.com/news/technology-65057011
[16] Daniel Wroclawski, *What to Do If the Police Ask for Your Security Camera or Video Doorbell Recordings*, Consumer Reports (May 2, 2023), https://www.consumerreports.org/legal-rights/police-ask-for-video-doorbell-recordings-what-to-do-faq-a8950763605/
[17] Department of Justice, Civil Rights Division, *Title II Of The Civil Rights Act (Public Accommodations)*, available at https://www.justice.gov/crt/title-ii-civil-rights-act-public-accommodations
[18] *EU Moving Closer to an AI Act?*, Sidley Austin LLP (Nov. 17, 2023), https://datamatters.sidley.com/2023/11/17/eu-moving-closer-to-an-ai-act/.
[19] *See* Nora Santalu, Biometrics under the EU AI Act, IAPP (Oct. 18, 2023), https://iapp.org/news/a/biometrics-under-the-eu-ai-act/.

and possible benefits. The following use cases provide different benefits and raise different concerns:

- To verify identity in a one-on-one capacity (i.e., is this person indeed who they are claiming to be?)
    - **Example**: person presents their ID at the airport, and a face-scanner matches their face-print to the ID photo
    - **Assumed Benefit**: speedier verification
    - **Issues**: inaccuracy, data breach, wrongful denial of vital functions
- To identify individuals in a group capacity (i.e. who are these people?)
    - **Example**: each person's face at a gathering or protest is scanned and matched against a database of faces, essentially de-anonymizing people in public
    - **Assumed benefit**: easier to find an individual in a large group (as compared to manually scanning faces)
    - **Issues**: inaccuracy, data breach, chilling of political expression, differential use based on the topic of the gathering or protest
- To categorize individuals (i.e., what are this person's traits?)
    - **Example**: a video hiring tool claims to use facial geometry as an indicator of job success
    - **Assumed benefit**: greater standardization of hiring procedure (as compared to current practices)
    - **Issues**: inaccuracy, data breach, harmful stereotypes or even phrenological pseudo-science, high barrier to use by people not familiar with technology or unable to use technology for various reasons including disability

As illustrated in the examples above, each broad goal carries unique potential for risk and possible benefit. Thus, a tiered approach based on the goal and function of the technology may be best suited, as it can ratchet up or down in terms of stringency and oversight. For example, technologies that use biometric data to categorize or identify groups in real time might be strictly regulated and, in some cases, prohibited. Technologies that verify identity in a one-to-one capacity, where opt-out consent is reasonably possible, might fall under a more permissive scheme.

While this method of regulation provides for a great deal of flexibility, a key challenge would be the ongoing task of classifying new technologies among these tiers. Vermont might consider periodically updating its classifications through rulemaking, or through enabling the Attorney General to review the classifications and update them on a set regular basis.