DIVISION OF RESPONSIBILITIES FOR THE SECURITY OF VOICEPRINT BIOMETRIC COLLECTION AND USE

A Research Paper submitted to the Department of Engineering and Society
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Electrical Engineering

By

Laura Gustad

March 25, 2021

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

ADVISOR
Catherine D. Baritaud, Department of Engineering and Society

As biometric authentication becomes more efficient and convenient for users, the security of our personal identifiers, such as voiceprint, must also be taken into consideration. The *Gale Encyclopedia of E-Commerce* defines voiceprint as a behavioral biometric consisting of the unique pattern of pitch, dynamics, and underlying characteristics that can be obtained from a voice recording ("Biometrics", 2002). Devices that are continually listening in or that use linguistics recognition to identify voice commands have access to the owner's unique, trackable voiceprint identifier, which increases the risk of identity theft in a data breach. As COVID-19 restricted the ability for music groups to practice together, the vocal harmonizer proposes an accessible solution. The creation of this device can assist musicians in times of distance learning by applying hardware and software to make one voice sound like many. On the other end of the spectrum, the Science, Technology, and Society (STS) focus explores the data privacy legislation that regulates the collection and use of an individual's voiceprint in technologies ranging from banking to voice assistants. The technical project is tightly coupled with the STS topic because the vocal harmonizer uses, but does not save, a portion of the user's voiceprint identifiable data.

As the use of voiceprint biometrics becomes more abundant, the risks to consumers are more concerning, and the ethical considerations need to be better incorporated in the regulation and design of voiceprint technologies. The STS research portion of this thesis project will be focused on the interactions between user perceptions of biometric privacy, business use of biometric data, and the regulation of data privacy, specifically exploring the views on inclusion of the private right of action. The Actor Network theory approach will be used in the analysis to assist in the exploration of the interactions between social groups in hopes of answering the question: How can we best protect consumer biometric information and data privacy, while allowing for innovation and development of efficient, useful technology? The STS discussion

will explore the existing regulations and current practices of voice data collection, consent, and use practices and consider possibilities for improving the ethical development and regulation of devices capable of collecting, storing, and using voice and biometric data.

INCREASED ADOPTION OF BIOMETRIC AUTHENTICATION CREATES PRIVACY CONCERNS

With the rise of the Internet of Things (IoT) and remote access, voice control and biometric authentication have become the buzz terms for new home technologies and devices. In fact, a Visa (2020) survey about consumer opinion on biometric authentication revealed 72% interest in fingerprint identification, 45% in eye scan, and 32% in voice recognition (slide 4). The graphic in Figure 1 illustrates the different types of biometric identifiers currently being used in identification technology.

Figure 1: Types of Biometric Identifiers. The main biometric identifiers, from facial recognition to vein patterns, are divided into physiological or behavioral categories, based on their consistency level (Thales Group, 2020, Biometrics: trends section).

As voiceprint is the only biometric authentication method that can be used via a remote channel, it is the preferred method of confirming identity in security for applications from remote banking to connected devices for the home (Wood, 2020, para. 2). With the increased use of

connected devices and biometric authentication, a privacy concern arises with the simultaneous and increased access to biometric identifiers. In 2017, research by Fairhurst, Li, and Da Costa-Abreu illuminated the increased vulnerability of identifiable information due to the sale of consumer data to allow companies to predict a potential customer's gender, age, and more based on biometrics (pp. 369-371). As the collection of voice data grows, the risk also increases that consumers' identifiable, biometric data could be vulnerable to unlawful collection, use, storage, or leak.

CONFUSION IN PATCHWORK OF STATE BIOMETRIC PRIVACY LAWS

At a glance, the outdatedness of biometric data protections in the United States is on display as federal law specifically protects only data related to healthcare information, children under the age of thirteen, and financial or credit information (Tschider, 2018, pp. 122-126). Lack of consumer protection on a national scale has sparked the creation of more comprehensive data privacy provisions by individual state legislatures to protect biometric privacy and better enforce consent standards for data collection. As outlined in Prescott's (2020) article in *The National Law Review,* Illinois enacted the Biometric Information Privacy Act (BIPA) in 2008, closely followed by Texas' Capture or Use of Biometric Identifier (CUBI) Act in 2009 (paras. 4-10). Prescott (2020) also describes the more recent legislation passed, including Washington's HB 1493 in 2017, New York's Stop Hacks and Improve Electronic Data Security (SHIELD) Act in 2019, California's Consumer Privacy Act (CCPA) in 2020, and Arkansas' Code §4-110-103(7) in 2020 (paras. 11-14).

However, each state privacy act provides a different definition of personal information, outlines a different standard for consent, and proposes different penalties for non-compliant companies. In terms of enforcement, Illinois' BIPA provides for private right of action, while

Texas, Washington, and California authorize the attorney general to enforce the policy (Prescott, 2020, paras. 9-11). The private right of action allows individuals to file a civil suit when their biometric privacy has been violated, when an entity is collecting biometric data without written consent or is storing that data for a nondisclosed amount of time. Prescott (2020) also emphasizes the disparity between the "$1,000 penalty for each negligent violation, or a $5,000 penalty for each willful violation" imposed by BIPA and the $25,000 civil penalty per violation imposed by CUBI (para. 7). The patchwork of different legislation leads to complications in identifying which businesses certain restrictions apply to, how penalties will be decided, and what the exact guidelines for the boundaries of the requirements are.

CLASS ACTION AND THE ACTOR NETWORK ANALYSIS OF BIOMETRICS

A crucial step in finding solutions in voiceprint biometric technology is understanding the interactions of major stakeholders affecting the collection, storage, and regulation of voiceprints. The Actor-Network Theory framework, first implemented by Callon (1984) and Latour (2005), is often used for analyses seeking to find a balance of power between actors surrounding a technology. As Callon (1984) explains, the actor-network framework allows for analysis of not only each actor's interaction with the technology, but also the "interessement," identities imposed by actors on each other, which will help to bring into focus the unique web of social forces between the actors in the network of voiceprint biometrics technology (p.207). Under existing regulations, the scarcity of private rights of action limits opportunities for citizens to be involved in pushing for regulation or amendment to existing legislation, which creates a missing interaction the Actor-Network theory model of biometric identification, as shown in Figure 2 on page 5.

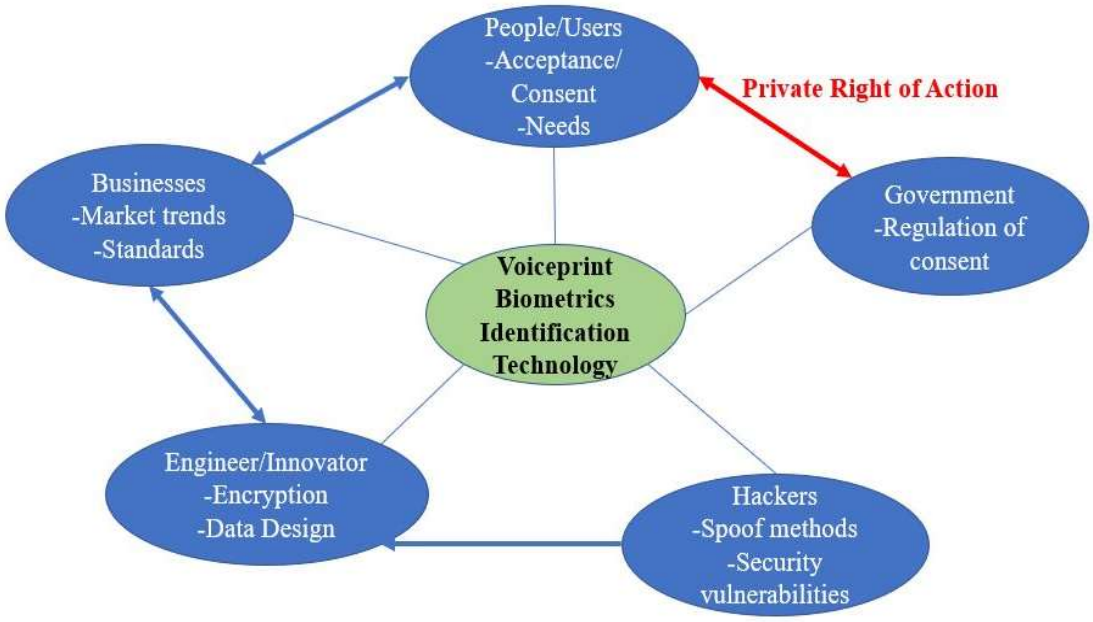

Figure 2: Actor-Network of Voiceprint Biometric Identification Technologies. The actors in the network of biometric identification are shown to interact outside of their main connection to biometrics; the red arrow highlights the often-missing connection between user citizens and protective legislation development (Adapted by Gustad, 2020, from Carlson, 2007).

When the Actor-Network framework is applied to voiceprint biometrics identification, the main actors can be defined as businesses, engineers, hackers, users, and government and their interactions can be mapped as shown in Figure 2. The hackers find new spoofing methods, which forces the engineers to advance encryption defense and data storage techniques. Engineers provide the biometric identification technology and algorithms for businesses to create new connected devices, either through industry-independent research or company-funded design. Businesses use predictive algorithms to target consumers. The consumer response to products, specifically, which product lines receive more consumer attention, influences the direction of market trends. Finally, government regulations protect consumer data and require businesses to notify consumers of a breach of their information. But in this Actor-Network model, the user impact on government regulation requires further investigation.

Narrowing Focus to Businesses, People/Users, and Government

The major public players in the voiceprint biometrics network are the Businesses, People/Users, and Government. The motivation behind voiceprint collection for businesses is not just to allow for efficient user authentication, but also to use it to collect other personal information for market research and targeted ads. Fairhurst, Li, and Da Costa-Abreu (2017) explain the growing ability to collect "soft" biometrics like age or gender from voiceprint, and explore research in identifying emotional state, which currently has an "accuracy of 68%" using voiceprint (p. 372). Increased market research in biometrics has led to the creation of larger shared databases and heightened incentives to sell consumer data. Yallen (2020) explains that due to the shocking number of data breaches between January 2013 and July 2018, "each American, on average, was a victim of data theft nineteen times" (p. 787). These statistics represent real threats to the privacy and safety of consumers' biometric privacy that need to be understood by individual users for informed consent and to be considered by businesses and engineers in the design of products.

In order to standardize and monitor the level of care taken by businesses in attaining informed consent from consumers and the precautions taken when collecting and storing biometric data, the interaction of businesses and consumers must be extended to include the government. Government involvement in the voiceprint biometrics network has thus far been limited to state legislation with two concrete methods for consumer intervention in data privacy requirements for businesses and engineers. First, in some states, consumers can request that a company send a copy of their data, modify or correct their data, or delete their personal data (Alexander, 2020, para. 8). Second, in other states, consumers can directly enforce their rights to

privacy by bringing legal actions, or the attorney general can impose penalties on companies who violate them (Alexander, 2020, paras. 4-7).

Debate Over Enforcement Methods and User Influence on Regulation

A survey of the existing state privacy laws identifies a divide in enforcement method between private right of action and enforcement by the attorney general, which posits the question: Is private right of action the most efficient method of user involvement? On the side of private right of action, individual participation in government regulation of biometric identification technology has already led to important precedent in business standards in Illinois. For example, the case of *Rosenbach v. Six Flags Entertainment Corp.* held that "plaintiffs can pursue Illinois Biometric Information Privacy Act claims even in the absence of an actual harm" (Rosenthal & Oberly, 2020, para. 8). The Rosenbach decision allows consumers to address violations and protect their personal data before errors cause irreversible consequences. Further support for private right of action can be seen in proposals for new state and national biometric privacy legislation. The California Privacy Rights Act (CPRA), passed in November 2020, builds on the California Consumer Privacy Act, including the addition of private right of action after a data breach (Privacy Rights Clearinghouse, 2020, para. 7). On the national scale, the proposed National Biometric Information Privacy Act provides for the private right of action in addition to the consent standards enumerated in state legislation (Alexander, 2020, para. 5).

The main arguments against private right of action are the strain on the court system and increased costs to businesses. Prescott (2020) connected the BIPA provision for recovery of attorney fees as the incentive for attorneys to take on BIPA cases, resulting in the 200 BIPA lawsuits filed in just 2018 and 2019 (paras. 7-8). For all the businesses that fall within the scope of the California Consumer Privacy Act, Yallen (2020) estimates total compliance costs "to

reach $55 billion in initial costs with up to an additional $16 billion to maintain compliance over ten years" (p. 818). As more states pass individual legislation, compliance becomes more complex and expensive for companies, and state courts are flooded with new cases. Instead of continuing this trend of overlapping state legislation, national legislation has been proposed as an alternative to provide a simpler source for compliance requirements and penalties for violations. Due to the increased scope of national legislative protections, enforcement by the attorney general or another individual party would not be as effective as expanding the responsibility to include consumers' use of the judicial system through private right of action.

NATIONAL BIOMETRIC PRIVACY LAW ENCOURAGING CHANGE IN DESIGN PARADIGM

The ideal actor-network for voiceprint biometrics on a national scale would include a shared understanding of risks posed by collection and storage of voiceprint biometrics, the consent being given, and the consequences of unethical design. National provisions for consent enforcement and protection of biometric privacy would encourage self-regulation in industry and consumer awareness/involvement in consent. The National Biometric Information Privacy Act proposes some direct security provisions, but also relies on entities to "maintain biometric information using the reasonable standard of care within the private entity's industry," a standard which is not always concretely defined (Alexander, 2020, para. 6). The national consent requirements would effectively change the design paradigm to consider, specifically, the length of storage, level of biometric detail needed, and the merits of the solution being provided as well as incorporating consent requirements before the sale of the product. Figure 3 shows the current design paradigm as enumerated by Martin and Schinzinger (2009), which will be used to discuss current practices in design and recommendations for better incorporation of consent regulations and the professional judgement of engineers in design tasks for biometrics devices (p. 6).

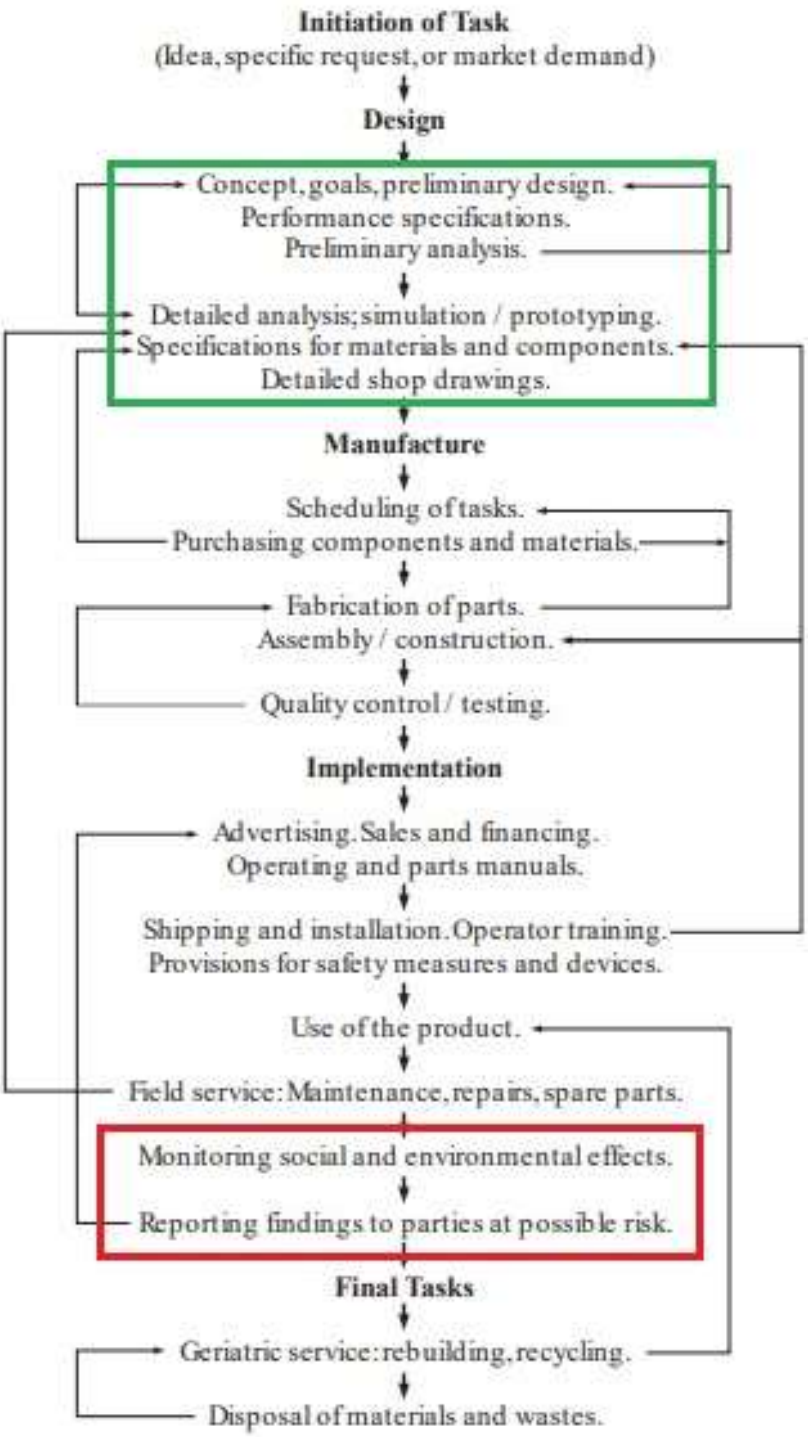

Figure 3: Current Progression of Engineering Tasks. Proposed national data privacy legislation could be implemented and further enforced through changes to the progression of engineering design and integration tasks. The tasks in red demonstrate the late consideration of social effects, while the tasks highlighted in green demonstrate a potential point much earlier in the process for incorporation of data privacy considerations to comply with proposed legislation (Adapted by Gustad, 2021, from Martin & Schinzinger, 2009b).

As demonstrated in Figure 3 on page 9, in the current design paradigm, consideration of the social effects of adopting a technology and communication of risk to the public are limited to vaguely defined tasks post implementation and sale of the product. As highlighted in red in Figure 3, Martin and Schinzinger (2009b) define tasks for the "monitoring of social and environmental effects" and "reporting of findings to parties at possible risk" that often occur only after the sale and use of the product (p. 6). Proposed national legislation with private right of action would split the responsibility for consumer biometric privacy considerations between the government, businesses and engineers, and users. The monitoring tasks post implementation and sale of the product would include individual consumer requests to businesses for termination of data storage, as well as any private right of action cases filed against businesses for alleged violations, with the expectation of correction by businesses as demonstrated in Figure 4 on page 11. However, professional ethics also call for action on the part of engineers in the design analysis and specifications stage of the progression highlighted in green in Figure 3 on page 9.

Besides futile check boxes for consumers to affirm that they have read and understand the privacy terms, there is no concrete method for communicating privacy risk to customers until after a breach when breach notification becomes relevant. Therefore, the responsibility to assess biometric data privacy risk falls to the professional judgement of engineers within the design process. For the progression of engineering devices with voiceprint audio, an improved paradigm will impose a task within the design analysis and specifications stage of engineering to assess and reduce the risks of storage and sale of consumer data from the perspective of a consumer as shown in Figure 4. The procedures for risk assessment outlined by Martin and Schinzinger (2009a) would require engineers to consider the unvoluntary risks and worst-case scenarios of consequences, should biometric data be leaked or stolen, and whether the efficiency and benefits

from using voiceprint biometrics are worth the risks (pp. 109-112, 117-120). After a risk-benefit analysis of the proposed device was completed, engineers would be able to make better informed decisions on collecting biometric data, necessary length of storage to complete the desired task, and the level of encryption required to reduce the risks posed to users.

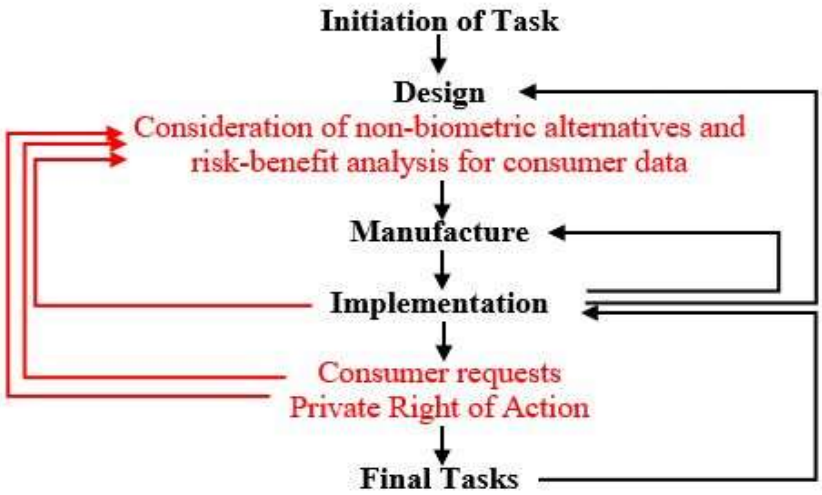

Figure 4: Proposed Design Paradigm to Address National Legislation. Proposed national biometric data privacy legislation can be implemented with the addition of the tasks in red to protect consumer data both before a product is manufactured and after its adoption (Gustad, 2021).

The full proposed design paradigm for ethical development of devices using biometric data is shown in Figure 4 with the incorporation of the professional responsibility of engineers and post-adoption enforcements by government and users. Due to the close coupling of the STS research and the technical project, the ethical responsibilities highlighted in engineering design tasks were also added to the technical project. In developing the vocal harmonizer for the technical project, the team sought to incorporate the recommended design paradigm by considering the impact of the product if it were to be mass produced and addressing those risks in the design. Privacy concerns with the collection of the user's voiceprint were alleviated by reducing the storage time and isolating data collection to only the necessary subset of the voiceprint, the fundamental frequency and harmonics. Due to the scope and timeline of the technical project, the progression of engineering tasks for the vocal harmonizer extended only to

manufacturing the product, so the considerations of changes to the tasks after implementation could not be incorporated.

UNDERSTANDING ETHICAL IMPLICATIONS OF VOICEPRINT AUTHENTICATION

In terms of the STS analysis, future research would depend on how the actor network of voiceprint biometrics is impacted by a national privacy statute. Further analysis could include suggestions for redefining the teaching of the design paradigm for voiceprint technologies and the creation of a new code of ethics. Proposals for a new code of ethics would also need to include example situations to provide guidance to engineers, which would require discussion of specific technologies. Case studies would require monitoring the social and environmental effects of products that were created according to national consent regulations and the legal proceedings for any violations, including how data risks or breaches were rectified and discussion of the ethical responsibilities involved. Depending on the adoption of a national biometric privacy policy, research could expand to consider an international privacy basis as proposed by Yallen (2020). Implementation of an international biometric and data privacy regulation would provide for more equal treatment across countries, simplify guidelines for businesses as to what regulations are applicable, and potentially reduce the strain on state and national courts if separate international enforcement was considered.

With respect to the technical project, future work would be concentrated on implementing encryption methods and redesigning for easier production. In the capstone course, the projects are limited in processor board options, but the creation of a specialized FPGA board would be more secure. Continuing the development of the vocal harmonizer, a device with a simple level of manipulation of the voiceprint, would provide an appreciation of the depth of

knowledge companies are able to gain with more sophisticated devices and AI algorithms. If the vocal harmonizer project were modified to meet the proposed consent requirements for the collection and storage of voiceprint biometrics, the necessary ethical considerations of consumer privacy in design could be considered a case study for future research in the STS analysis.

WORKS CITED

Alexander, C. (2020, August). *National Biometric Information Privacy Act proposed by US lawmakers*. Iron Mountain. www.ironmountain.com/

Biometrics. (2002). In J. A. Malonis (Ed.), *Gale Encyclopedia of E-Commerce* (Vol. 1, pp. 64-65). Gale.

Callon, M. (1984). Some elements of a sociology of translation: domestication of the scallops and the fisherman of St Brieuc Bay. *The Sociological Review*, *32*(1), 196-233. doi: 10.1111/j.1467-954X.1984.tb00113.x

Fairhurst, M., Li, C., Da Costa-Abreu, M. (2017, November). Predictive biometrics: a review and analysis of predicting personal characteristics from biometric data. *IET Biometrics,* 6(6), 369-378. doi: 10.1049/iet-bmt.2016.0169

Gustad, L. (2020). *Actor Network of Voiceprint Biometric Identification Technologies.* [Figure 2]. *STS Research Paper: Division of Responsibilities for the Security of Voiceprint Biometric Collection and Use* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.

Gustad, L. (2021). *Current Progression of Engineering Tasks.* [Figure 3]. *STS Research Paper: Division of Responsibilities for the Security of Voiceprint Biometric Collection and Use* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.

Gustad, L. (2021). *Proposed Design Paradigm to Address National Legislation*. [Figure 4]. *STS Research Paper: Division of Responsibilities for the Security of Voiceprint Biometric Collection and Use* (Unpublished undergraduate thesis). School of Engineering and Applied Science, University of Virginia. Charlottesville, VA.

Latour, B. (2005). *Reassembling the social: An introduction to Actor-Network-Theory*. Oxford, NY: Oxford University Press. doi: 10.1080/10967490701515606

Martin, M., & Schinzinger, R. (2009a). Commitment to safety. In *Introduction to Engineering Ethics* (2nd ed.). McGraw-Hill Higher Education.

Martin, M., & Schinzinger, R. (2009b). Ethics and professionalism. In *Introduction to Engineering Ethics* (2nd ed.). McGraw-Hill Higher Education.

Prescott, N. (2020, January 15). The anatomy of biometric laws: What U.S. companies need to know in 2020. *The National Law Review.* Retrieved from https://www.natlawreview.com/

Privacy Rights Clearinghouse (2020, December). *California Privacy Rights Act: An overview.* Privacy Rights Clearinghouse. https://privacyrights.org/

Rosenthal, J. & Oberly, D. (2020, February). Biometric privacy in 2020: The current legal landscape. *Law360.* Retrieved from https://www.blankrome.com/

Thales Group. (2020, September). *Biometrics: definition, trends, use cases, laws and latest*. Thales: Building a future we can all trust. https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/inspired/biometrics

Tschider, C. (2018). Regulating the internet of things: Discrimination, privacy, and cybersecurity in the artificial intelligence age. *Denver Law Review*, *96*(1), 87–143.

VISA (2017). *Goddbye, passwords. Hello, biometrics* [Presentation slides]. Retrieved from https://usa.visa.com/dam/VCOM/global/visa-everywhere/documents/visa-biometrics-payments-study.pdf

Wood, L. (2019, December). *Worldwide voice biometrics markets, forecast to 2024- Emerging lucrative opportunities within the BFSI vertical.* Business Wire. https://www.businesswire.com/

Yallen, J. (2020). Untangling the privacy law web: Why the California Consumer Privacy Act furthers the need for federal preemptive legislation. *Loyola of Los Angeles Law Review, 53*(4), 787-825. Retrieved from https://digitalcommons.lmu.edu/llr/