Digital Ethics for Biometric Applications in a Smart City

LYNNETTE H. X. NG, Centre for Informed Democracy & Social Cybersecurity, Carnegie Mellon University
ABIGAIL C. M. LIM, Centre for Trusted Internet and Community, National University of Singapore
ADRIAN X. W. LIM, School of Computing, National University of Singapore
ARAZ TAEIHAGH, Lee Kuan Yew School of Public Policy, National University of Singapore and Centre for Trusted Internet and Community, National University of Singapore

From border control using fingerprints to law enforcement with video surveillance to self-activating devices via voice identification, biometric data is used in many applications in the contemporary context of a Smart City. Biometric data consists of human characteristics that can identify one person from others. Given the advent of big data and the ability to collect large amounts of data about people, data sources ranging from fingerprints to typing patterns can build an identifying profile of a person. In this article, we examine different types of biometric data used in a smart city based on a framework that differentiates between profile initialization and identification processes. Then, we discuss digital ethics within the usage of biometric data along the lines of data permissibility and renewability. Finally, we provide suggestions for improving biometric data collection and processing in the modern smart city.

CCS Concepts: • **Social and professional topics** → **Government technology policy**; **Governmental regulations**; **User characteristics;** • **Security and privacy** → **Human and societal aspects of security and privacy;** • **Computing methodologies** → **Artificial intelligence**;

Additional Key Words and Phrases: Biometric, artificial intelligence, ethics, smart city, governance

ACM Reference format:
Lynnette H. X. Ng, Abigail C. M. Lim, Adrian X. W. Lim, and Araz Taeihagh. 2023. Digital Ethics for Biometric Applications in a Smart City. *Digit. Gov. Res. Pract.* 4, 4, Article 26 (December 2023), 6 pages.
https://doi.org/10.1145/3630261

1 INTRODUCTION

A biometric trait is a measurable characteristic of a human person. This characteristic differs sufficiently from person to person, harnessed toward the identification of an individual [1]. There are generally two broad

This research is supported by the Centre for Trusted Internet and Community of the National University of Singapore, Project No. CTIC-RP-20-02, Grant No. A-0003503-08-00.
Authors' addresses: L. H. X. Ng, Centre for Informed Democracy & Social Cybersecurity, Carnegie Mellon University, 5000 Forbes Avenue Pittsburgh Pennsylvania, USA, zip 15213; e-mail: lynnetteng@cmu.edu; A. C. M. Lim, Centre for Trusted Internet and Community, National University of Singapore; Innovation 4.0, #04-04 3 Research Link Singapore 117602; e-mail: abi-gayle.writes@gmail.com; A. X. W. Lim, School of Computing, National University of Singapore; COM1, 13 Computing Dr, Singapore 117417; e-mail: adrianlimxw@gmail.com; A. Taeihagh (Corresponding author), Lee Kuan Yew School of Public Policy, National University of Singapore, 469B Bukit Timah, Rd, Li Ka Shing Bldg., Level 2, #02-10, 259771, Singapore; e-mail: spparaz@nus.edu.sg.

2639-0175/2023/12-ART26
https://doi.org/10.1145/3630261

categories of biometric traits: physiological and behavioural traits. Physiological characteristics of the human body, such as fingerprints and facial features, are universal human characteristics that are present and non-renewable throughout a person's lifetime. Behavioural traits, such as voice and writing style, provide some information about individuals but are usually renewable and might not have a sufficiently high variance to differentiate two individuals [2, 3].

Smart cities have a network of interconnected devices that access, store, and transmit personal information [4]. Biometric traits provide a wealth of data to drive smart-city applications in the following areas: identification, authentication, surveillance, and personalization. Physiological characteristics such as fingerprints and palm prints are commonly used in forensics and crime-scene investigations to identify suspects [5]; and facial features are used in user authentication in smartphones [6]. Behavioural traits like gait patterns are used during surveillance by law enforcement agencies [7]; and voice patterns and modulation are used in personalization of smart-home voice-controlled devices like Siri [8].

2 BIOMETRIC DATA TRAITS FOR SMART CITY APPLICATIONS

With big data technology, almost any characteristic that can be harvested from a human can be used to form identification patterns to build a profile for a human [9]. Formulating each data collected as a pattern-of-life formulation, essentially the formation of patterns from sufficient data points of a particular trait of a human, can act as a biometric data characteristic for the identification of a human. For example, a sole fingerprint can identify a person, and so can gait patterns made up of many data points of a person's walking speed, stride, and timing of walk identify a person. In fact, in this digital social media age, even the online posting time and the posting language used can aid in distinguishing a person; this idea is harnessed by classifying automated and human accounts on social media through their temporal and linguistic properties [10].

A biometric application operates in two key stages: initialization and identification. At initialization, the biometric trait is measured from the human and transformed into a machine-readable format (i.e., vector or image) for storage within a database. At the identification stage, the same biometric trait is measured from the target human, converted into a machine-readable format then compared against the other stored templates within the database before reporting the person's identifier if found [11]. In the context of data required, we can then characterize biometric traits by the two key stages in a biometric application: in terms of the amount of data required for initialization and identification. The ideal biometric trait is one that requires low amount of data for both stages. Such examples are fingerprints and iris, and therefore they are most commonly used for critical applications such as border access control identification.

Adapting the biometric trait classification put forth by Raju and Udayashankara [3], Table 1 profiles several human traits in terms of segregating the human body into six different regions (hand, facial, ocular, medico-chemical, behavioural, soft) and profile the amount of data required for initialization and identification (high/low) and examples of biometric applications in a smart city. We note that soft biometrics like gender and hair colour are insufficient by themselves in distinguishing people due to their lack of uniqueness; more than one person can have the same trait. They are usually used to supplement traits obtained from other regions [11].

3 DIGITAL ETHICS IN BIOMETRIC DATA

We look at digital ethics in biometric data collection and usage through the frame of two attributes: the permissibility of data collection during the initialization stage and the renewability of biometric trait at the identification stage.

Ethical biometric data collection for smart cities should ensure that the information gathered during the initialization stage of an application is permissible by having data subjects provide consent before their data is harvested. A permissible trait requires specialized equipment and the physical presence of the person to harvest the data; a non-permissible trait means the data can also be harvested via observations. Data collected from

Table 1. Characterization of the Types of Biometric Data Traits

Region (adapted from Reference [3])	**Biometric data trait**	**Amount of data required for Initialization**	**Amount of data required for Identification**	**Example Applications**
Hand	Fingerprint	Low	Low	Border access controls, office security controls, theme parks (e.g., Disney World) [11]
	Palm print	Low	Low	Crime scene forensics [5]
Face	Facial features	High	High	Video surveillance [12], authentication in smartphones (e.g., Face ID) [6]
Ocular	Iris	Low	Low	Security systems [13]
Medico-Chemical	Electro-Cardio-Grams	High	High	Wearable devices for sports competitions and performances [14], critical health care [15]
	Heart rate	High	High	Wearable devices for sports competitions and performances [14], prediction of emotions in students to improve the learning process [16]
Behavioural	Voice	High	High	Smart home voice control (e.g., Siri, Alexa) [8], automated phone operating systems (e.g., telebanking) [3]
	Gait	High	High	Surveillance [7]
	Writing style	High	High	Crime scene forensics [17]
Soft	Gender, Hair colour, Height	Low	Unable to identify a person by itself	Ancillary information for forensic evaluation [18], smart home personalization [4]

non-permissible traits through observations can infringe on a person's privacy as no explicit consent is given. The right to privacy is one of the fundamental rights of human beings set out in the Universal Declaration of Human Rights [19]. Data collection through observation, such as harvesting people's gait or typing patterns, can often be used to identify individuals. This is not only limited to the physical space. A person's digital presence and profile can also be tracked through posts, images, and friends' information from social media platforms.

A person's personal profile and temporal and spatial movements should be kept private and surveyed with a proper warrant [20]. Biometric data that has been exposed to consumer technologies can be passed without the knowledge or consent of consumers to the third parties. Several companies, such as Amazon Ring and Family Tree DNA, have passed on their consumer data to law enforcement agencies without prior consent from their customers [21, 22]. Ethical biometric data collection should involve an opt-in process for people to allow biometric applications to harvest and store a template of their traits, and explicit consent should be obtained from application users before information is passed on to other entities. Regulations such as the **Biometric Information Privacy Act (BIPA)** have been developed to aid in this aspect. The BIPA requires private entities to have a written policy on the purpose and the time period the data is kept and obtain written consent from the application users [23]. The General Data Protection Regulation imposes tough obligations such as the mandatory performance of privacy impact assessments and the requirement for user consent for biometric applications [24].

At the identification stage of a biometric application, the application used to identify a person should not rely solely on non-renewable biometric traits, for if that trait is stolen, then a person's identity is stolen as well. A renewable trait means the trait can be changed across time; a non-renewable trait means that it does not change throughout a person's lifetime. Examples of non-renewable traits are fingerprints and palmprints; and that of renewable traits are gait and writing style. Should a person's fingerprint be lifted off a surface and a

Table 2. Permissibility and Renewability of Biometric Data

Region (adapted from Reference [3])	**Biometric Data Trait**	**Permissible**	**Renewable**
Hand	Fingerprint	Yes	No
	Palm print	Yes	No
Face	Facial features	No	No
Ocular	Iris	Yes	No
Medico-Chemical	Electro-Cardio-Grams	Yes	No
	Heart rate	Yes	No
Behavioural	Voice	No	Yes
	Gait	No	Yes
	Writing style	No	Yes
Soft	Gender, Hair colour, Height	No	Yes

mould made of it, the person's identity can be impersonated to fool fingerprint-based authentication devices. Instead, the identification should rely on both renewable and non-renewable data sources to ensure accurate identification, since such a combination is less likely to be impersonated. The combination of a fingerprint and a behavioural trait like gait is less likely to be accurately impersonated and thus allows the person to preserve his identity. Renewable biometric data templates thus preserve a person's identity and make the system less prone to identity theft [25].

Table 2 profiles the biometric data traits analysed in the previous section and sets out their permissibility and renewability factors. Most physiological traits are generally permissible and require the presence of the data subject for data collection to happen. The person needs to be physically present for the biometric application to capture his fingerprints or iris information at the initialisation stage. Some, though, like facial features, can be non-permissible, because they can be extracted from observed images through computer vision algorithms. Behavioural traits generally require high amounts of data for initialisation and a long period of data capture and can be obtained through observation of images or video feeds. Thus, these traits are generally non-permissible. While consent can be explicitly given, biometric data such as gait patterns can be harvested from prolonged observation of an individual. Physiological traits that are derived from the hand, face, ocular and medico-chemical regions are non-renewable, as these traits are not consciously controlled by a person and will remain with the person for life. If these traits require low data for identification, once exposed, they can make a person prone to identity theft or misidentification. Behavioural traits are renewable as they can be consciously altered and do change over a person's lifetime or lifestyle. For example, while a person's writing style can remain relatively constant for a short period of time, it can change through external influences like attending writing classes.

However, requiring consent for permissible and renewable data traits results in security application trade-offs. Requiring consent prior to data collection is a trade-off with security: those with bad intentions may not be captured within the databases as they opt out of the data collection; therefore, security systems will not have a record of their biometric traits and will be unable to identify them. A fine line must be drawn between regulating the permissibility of biometric data collection through consent, protecting citizens' privacy as data subjects and security through identification and surveillance using biometric traits. Using renewable behavioural traits means that one's behaviour can be obfuscated or mimicked through deliberate changes, which causes one profile to look similar to another, resulting in false positive hits from the application or inconclusive results. Therefore, multiple traits need to be used to increase identification accuracy. There are several multimodal biometric applications that have been developed that combine renewable and non-renewable data sources, e.g., face features and speech [3], fingerprint and voice [26], face features and gait [27].

4 CONCLUSION

Biometric data is widely used in smart city applications and will continue to increase in usage as research into measuring and harnessing human traits progresses. Within this article, we put forth that most traits harvested from humans can form a profile of a person, and we examined different types of biometric data in terms of the data required for initialization and identification. We further discussed ethical issues regarding data collection in terms of data permissibility and renewability. An ideal biometric data collection for initialization should involve that data collected is permissible, and that user consent is provided before data is extracted. The ideal biometric data for identification should involve a mixture of renewable and non-renewable data to prevent identity fraud and increased accuracy. We hope this discussion serves as a stepping board toward more ethical biometric data collection in humankind's effort toward a smart city.

REFERENCES

[1] V. Andronikou, A. Yannopoulos, and T. Varvarigou. 2008. Biometric profiling: Opportunities and risks. In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, M. Hildebrandt and S. Gutwirth, Eds., Springer, Dordrecht, 131–145. DOI: 10.1007/978-1-4020-6914-7_7

[2] X. Lu and A. K. Jain. 2004. Ethnicity identification from face images. In *Biometric Technology for Human Identification*. SPIE, 114–123. DOI: 10.1117/12.542847

[3] A. S. Raju and V. Udayashankara. 2014. Biometric person authentication: A review. In *Proceedings of the International Conference on Contemporary Computing and Informatics (IC3I'14)*. IEEE, 575–580. DOI: 10.1109/IC3I.2014.7019771

[4] A. Ross, S. Banerjee, and A. Chowdhury. 2020. Security in smart cities: A brief review of digital forensic schemes for biometric data. *Pattern Recogn. Lett.* 138 (2020), 346–354. DOI: 10.1016/j.patrec.2020.07.009

[5] Y. Chen, S. C. Dass, and A. K. Jain. 2005. Fingerprint quality indices for predicting authentication performance. In *Audio- and Video-based Biometric Person Authentication*, T. Kanade, A. Jain, and N. K. Ratha, Eds., In Lecture Notes in Computer Science, vol. 3546. Springer, Berlin, 160–170. DOI: 10.1007/11527923_17

[6] A. Bud. 2018. Facing the future: The impact of apple FaceID. *Biometric Technol. Today* 1 (2018), 5–7. DOI: 10.1016/S0969-4765(18)30010-9

[7] I. Bouchrika, M. Goffredo, J. Carter, and M. Nixon. 2011. On using gait in forensic biometrics. *J. Forensic Sci.* 56, 4 (2011), 882–889, DOI: 10.1111/j.1556-4029.2011.01793.x

[8] H. Feng, K. Fawaz, and K. G. Shin. 2018. Wearable technology brings security to Alexa and Siri. *GetMobile: Mobile Comp. Comm.* 22, 1 (2018), 35–38, DOI: 10.1145/3229316.3229328

[9] A. K. Jain, A. Ross, and S. Prabhakar. 2004. An introduction to biometric recognition. *IEEE Trans. Circ. Syst. Video Technol.* 14, 1 (2004), 4–20. DOI: 10.1109/TCSVT.2003.818349

[10] L. H. X. Ng and K. M. Carley. 2023. BotBuster: Multi-platform bot detection using a mixture of experts. In *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM'23)*. 686–697. DOI: 10.1609/icwsm.v17i1.22179

[11] A. K. Jain, K. Nandakumar, and A. Ross. 2016. 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recogn. Lett.* 79 (2016), 80–105. DOI: 10.1016/j.patrec.2015.12.013

[12] J. Greener and L. Naegler. 2022. Between containment and crackdown in Geylang, Singapore: Urban crime control as the statecrafting of migrant exclusion. *Urban Studies* 59, 12 (2022), 2565–2581. DOI: 10.1177/00420980211034681

[13] S. Liu and M. Silverman. 2001. A practical guide to biometric security technology. *IT Profess.* 3, 1 (2001), 27–32. DOI: 10.1109/6294.899930

[14] M. Kos and I. Kramberger. 2017. A wearable device and system for movement and biometric data acquisition for sports applications. *IEEE Access* 5 (2017), 6411–6420. DOI: 10.1109/ACCESS.2017.2675538

[15] E. Kalai Zaghouani, A. Benzina, and R. Attia. 2017. ECG-based authentication for e-healthcare systems: Towards a secured ECG features transmission. In *Proceedings of the 13th International Wireless Communications and Mobile Computing Conference (IWCMC'17)*. 1777–1783. DOI: 10.1109/IWCMC.2017.7986553

[16] M. Hernandez-de-Menendez, R. Morales-Menendez, C. A. Escobar, and J. Arinez. 2021. Biometric applications in education. *Int. J. Interact. Des. Manuf.* 15, 2 (2021), 365–380. DOI: 10.1007/s12008-021-00760-6

[17] R. V. Yampolskiy and V. Govindaraju. 2008. Behavioural biometrics: A survey and classification. *Int. J. Biomet.* 1, 1 (2008), 81–113. DOI: 10.1504/IJBM.2008.018665

[18] M. Tistarelli, E. Grosso, and D. Meuwly. 2014. Biometrics in forensic science: Challenges, lessons, and new technologies. In *Biometric Authentication*, V. Cantoni, D. Dimov, and M. Tistarelli, Eds., In Lecture Notes in Computer Science. Springer International Publishing, Cham, 153–164. DOI: 10.1007/978-3-319-13386-7_12

[19] A. Nautsch et al. 2019. Preserving privacy in speaker and speech characterisation. *Comput. Speech Lang.* 58 (2019), 441–480. DOI: 10.1016/j.csl.2019.06.001

[20] B. Manby. 2021. The sustainable development goals and legal identity for all: First, do no harm. *World Dev.* 139, 105343, DOI: 10.1016/j.worlddev.2020.105343

[21] D. Calacci, J. J. Shen, and A. Pentland. 2022. The cop in your neighbor's doorbell: Amazon ring and the spread of participatory mass surveillance. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2 (2022), 1–47. DOI: 10.1145/3555125

[22] S. Skeva, M. H. Larmuseau, and M. Shabani. 2020. Review of policies of companies and databases regarding access to customers' genealogy data for law enforcement purposes. *Personal. Med.* 17, 2 (2020), 141–153. DOI: 10.2217/pme-2019-0100

[23] R.-J. Yew and A. Xiang. 2022. Regulating facial processing technologies: Tensions between legal and technical considerations in the application of Illinois BIPA. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*. ACM, 1017–1027. DOI: 10.1145/3531146.3533163

[24] M. Monajemi. 2017. Privacy regulation in the age of biometrics that deal with a new world order of information. *U. Miami Int'l Comp. L. Rev.* 25 (2017), 371.

[25] T. A. M. Kevenaar, G. J. Schrijen, M. van der Veen, A. H. M. Akkermans, and F. Zuo. 2005. Face recognition with renewable and privacy preserving binary templates. In *Proceedings of the 4thIEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)*. 21–26. DOI: 10.1109/AUTOID.2005.24

[26] M. A. Kowtko. 2014. Biometric authentication for older adults. In *Proceedings of the IEEE Long Island Systems, Applications and Technology Conference (LISAT'14)*. 1–6. DOI: 10.1109/LISAT.2014.6845213

[27] X. Zhou and B. Bhanu. 2006. Feature fusion of face and gait for human recognition at a distance in video. In *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*. 529–532. DOI: 10.1109/ICPR.2006.556

Received 14 July 2023; accepted 16 October 2023