Unified surveillance systems: Data mining with PeekYou, GPS and facial recognition

Received: 30th January, 2019

Jessica Berger
MLIS, CIPM is an information security and privacy consultant whose research on the cyber security of drones informed drone data privacy policy design for the cities of Boston and Toronto. As a former paediatric registered nurse and health columnist concerned with computer-based patient records and confidentiality, she has published articles addressing online safety and the privacy rights of children and patients. She holds a master of library and information science and is an International Association of Privacy Professionals (IAPP) certified privacy programme manager. Through San Jose State University, Jessica served as an information security and records manager intern, originating the Wikipedia Library's privacy programme including security measures around integrated apps, encryption protocols, social media and virtual private network (VPN). She currently offers cyber security analysis and privacy programme consulting in Massachusetts and Connecticut.

Tel.: (413) 527 1869; E-mail: jsberger@tutanota.com; Website: https://privacyprotectordotblog.wordpress.com/

Abstract Unified surveillance systems threaten to unlock a portal to mass surveillance, swift round ups, incarceration and deportation. By combining data from an array of people tracking technologies, governments and corporations can now instantly locate and monitor entire populations in real time. This paper unveils the mechanics of how these technologies are bound together within the fabric of our daily lives, silently invading our privacy. Research methods include examination of more than 29 articles, an exploration of PeekYou and an interview with an accomplished transportation professional from Seattle, Washington. Privacy-by-design offers the promise of freedom from constant corporate and government scrutiny. Ongoing coerced assent to mass surveillance need not remain our global fate. The United States Constitution, in particular the Third, Fourth and Fifth Amendments, offers citizens protection from the systematised misapplication of these invasive and largely covert programmes. As such, it is never too late to alter the course of history so that we amplify these technology's attributes while protecting all people from the abuse inherent in their utilisation.

KEYWORDS: artificial intelligence, bulk data collection, connected technologies, corporate surveillance, data capture, facial recognition software, geolocation data, government surveillance and privacy

INTRODUCTION

Since 2010, the Transportation Security Administration (TSA) has conducted warrantless surveillance of innocent air travelers targeted by the Department of Homeland Security (DHS). Merely sweating too much can land you on the DHS watch list.[1] Mass surveillance in the name of security crowds both our skies and highways. Linking Massachusetts Department of Transportation (Mass DOT) tracking policies to global positioning systems (GPS), Hausman,[2] explains that geo-location tracking involves multiple

modalities, of which GPS is but one. Dunn found that 'Google is tracking your location, even when the setting is turned off'.[3] Connected technologies provide ample opportunity for corporate and government spying. In fact, 100,000 people's movements were traced through cell phone data triangulation.[4] Amazon has profited from its facial recognition system, Rekognition, by selling facial images to the police. The American Civil Liberties Union (ACLU) explains, 'With Rekognition, a government can now build a system to automate the identification and tracking of anyone.'[5]

PURPOSE

This paper introduces the fundamentals of facial recognition software, GPS and PeekYou, with a focus upon the interwoven impacts these inventions have upon our privacy and freedom. Do these combined technologies really enable the tracking and monitoring of entire populations in real time? If so, this triumvirate could readily coordinate the distribution of goods and services to needy populations. In contradistinction, the pandemics of mass incarceration and slavery[6,7] position these technologies as harbingers of doom for those concerned with privacy and freedom. We will explore how these modalities function as a unified surveillance system. In response, we will highlight how the US Constitution offers citizens protection from the systematised misapplication of these invasive and largely covert systems.

DEFINITIONS AND HISTORY OF PEEKYOU, GPS AND FACIAL RECOGNITION SOFTWARE

General definitions related to data mining

'Chelsea Manning has compared life in the US to her time in prison because of surveillance systems, cameras and the presence of police'.[8]

Basic data mining terminology is essential to understanding how disparate information retrieval systems work together synergistically to track people. When people's movements, images, keystrokes or conversations are simply monitored and recorded, this is referred to as *data capture*.[9] Once data is captured by one system, it can be added to data from other systems. Combining data from different sources is referred to as *aggregation*.[10] Matching algorithms applied to this aggregated information pinpoint a specific individual. This process is called *data matching*.[11] Finally, we stumble into 'Data mining, (which) does not discover already existing information; (but) generates new information about existing information'.[12] Privacy, liberty, and legal concerns arise even in the face of claims that this information is anonymised. In separate reports, Weisinger,[13] Barocas[14] and Leonard[15] quoting Cavoukian and Jonas[16] concur that re-identification of individuals is inherent in the data mining process as numerous pieces of personal information are clustered into a set. This data is so specific in terms of activities, addresses and habits that it will re-identify individuals.

History and definition of PeekYou

Bell[17] interviews PeekYou advisory board member Marshall Sponder, who reveals what can be done by combining text analytics with geo-location data picked up from mobile devices. Sponder aspires to a career in politics that leverages his awareness of data mining.[18] Meanwhile, PeekYou went from Beta in 2006,[19] to live in 2007, claiming '50 million users'. This people search engine crawls the web, collocating online materials about a single individual into one profile. The patented algorithm allows 'PeekYou's search engine (to) calculate(s) the likelihood of any URL being associated with an individual'.[20]

Background and definition of Global Positioning System (GPS)

People connected to smartphones associated with Verizon can be found in real time because the phone's whereabouts can be triangulated 'every 7 second(s)'[21] in relation to the nearest mobile phone tower.[22–24] The inclusion of 'black boxes' throughout a car's operating systems allow for monitoring of its route, miles and speed.[25] Related to, but different from GPS, is the Massachusetts Department of Transportation's (Mass DOT) collection of data about drivers on toll roads. Here, automobiles pass under gantries that scan the EZPass to bill the driver for the toll. If drivers do not have an EZPass, a camera takes a photo of the car licence plate and a bill is sent to the driver.[26]

Linking Mass DOT tracking policies to GPS, Hausman explains that geolocation tracking involves multiple modalities, of which GPS is but one; Hausman relates that GPS 'consists of orbiting satellites…detected by a GPS receiver in a cell phone or other mobile device…(with) a clear line of sight to four or more satellites.'[27] Whereas in 2014, only under certain circumstances could a person be immediately found with near precision, now there exists an entire industry called People Tracking Technology.[28,29]

Radio frequency identification devices (RFIDs), vision analytics, raspberry pi and 3D spatial learning are just four of the 15 people tracking technologies available to retailers, government and employers.[30]

History and definitions of facial recognition

Innocent football fans attending the 2001 Tampa, Florida Super Bowl were the unwitting subjects of mass surveillance through facial recognition technology. These sports enthusiasts' faces became images in a *data capture*. Their facial images were *aggregated*, to then become subjects of *data matching* against a criminal database by law enforcement.[31,32] This covert manoeuver was purportedly conducted to assay the efficacy of facial recognition technology as a counter-terrorism tool. Nineteen sports fans faces matched up with those of 19 'petty criminals'.[33] The fact that an entire stadium of people were experimented upon without their consent and treated like 'mug shots' is itself a controversial topic. Agre explains that facial recognition software not only matches faces against a database of other faces, it also 'can extract facial expressions.'[34]

Now, new facial recognition technologies can even identify emotions.[35] Friedland reports that in the future, our government could spy on our faces from afar.[36] Given the ubiquity of surveillance cameras, it becomes impossible to avoid being tracked and monitored. Agre points out that unlike other biometrics, facial recognition technology renders individuals powerless with regards to choice.[37]

AGGREGATION OF DATA FROM PEEKYOU, GPS, AND FACIAL RECOGNITION APPS

By combining the information from PeekYou, GPS and facial recognition software, businesses can covertly uncover a new customer's preferences and assets. Hausman suggests that vendors might provide preferential service to those with rich profiles.[38] Barocas contends that this phenomenon predisposes to prejudicial treatment.[39] While shopkeepers may appreciate FaceFirst's identification of individuals with a history of shoplifting,[40] shoppers attempting to turn over a new leaf will find themselves pegged as a criminal in perpetuity. In fact, Event Technology's sales pitch lauds its ability to immediately identify people on watch lists through facial recognition.[41] Imagine having a medical issue that causes you to sweat, being put on a DHS watch list because of this 'unusual behaviour' and then subsequently being barred from events.

AN INTERVIEW: SEATTLE TRANSPORTATION PROFESSIONAL DISCUSSES PRIVACY BY DESIGN TRAFFIC TECHNOLOGY INNOVATIONS

Fortunately, municipal leaders are working to safeguard the privacy of citizens through adoption of new transportation technologies offering privacy by design. The following interview with an innovative transportation professional from Seattle, Washington, showcases this crucial development.

Q: Is there any way to remain anonymous while driving on toll roads?
A: There are two ways to look at this: Primarily, the answer is 'NO', at least not legally. The vehicle will be identified either through reading the toll tag or through a license plate reader (most systems incorporate these). The only way around this would be to not have a tag or a license plate. However, from another perspective, the answer is 'YES' in that inherently the system cannot positively identify the driver, only the registered owner of the vehicle or account.
Q: Is your license plate scanned if you don't have an EZPass?
A: This is the case for most systems. If there is a lane or option on the signage for 'pay by mail', they're reading the license plate and billing the registered owner.
Q: Is a record of when your EZPass travels below the gantries recorded and held?
A: Yes, at least until payment is received.
Q: Could you describe any privacy-by-design transportation innovations in your area?
A: There is a related situation regarding license plate readers. Wherever you see a system that says, 'X minutes to <someplace>', it's doing that using the technology. In this case it's a two-agent system: The first agent reads license plates optically. It doesn't store any image; it just directly creates a record of that license plate which is assigned a temporary record ID. That system holds the association between the license plate number and the ID, but not any geo-location or timestamp. The second agent receives only the record ID and adds the geo-location (based on the reader that did the detection) and timestamp. That way, one agent only knows the license plate number, but not where or when it was seen, and the other only knows where and when but not who. I would expect that the information would have a TTL (time to live) and be automatically deleted after that expires. This comes up regarding 'Connected Vehicles', too. The standard (for when/if it ever goes live) uses security certificates that are issued to vehicles that are only valid for 5 minutes after it is first used. In the same way, the certificate issuer and the system don't communicate the association between the certificate and vehicle to preserve anonymity.
(Transportation Professional from Seattle, WA, 9th May, 2019, personal communication)

SOCIAL AND PRIVACY ISSUES

Thompson and Thompson define privacy as 'The right to be let alone and the right of individuals to determine when, how, and how much information about themselves is released to others.'[42] Importantly, they emphasise that privacy includes freedom from intimidation to reveal personal secrets in order to attend to one's daily affairs.[43] Friedland notes that we must comply with numerous intrusive online privacy policies to merely conduct our basic banking and healthcare needs.[44] The same is true for social media. Citizens must either comply with these policies or forgo banking, medical care, social media and more; however, there has been some regulatory pushback.

While PeekYou profited from aggregating personal data, they earned the dubious distinction of being subpoenaed by the Federal Trade Commission (FTC). Timberg explains how the FTC highlighted PeekYou's unscrupulous practice of 'onboarding (which) allows markets to load offline information — from magazine subscriptions, store loyalty cards or government records — into cookies that digital advertisers use to target consumers for pitches.'[45]

Connected technologies provide ample opportunity for government and corporate spying. On the corporate front, Alim et al. (2017), of The Electronic Frontier Foundation, found that Google collects and stores vast swaths of personal information about K-12 students who are forced to use their services as a condition of matriculation.[46] Tunick explains how US police favour GPS and video cameras over traditional stakeouts.[47] Friedland introduces the loophole concept of 'The Silent Subpoena'.[48] This legalises warrantless searches in which the government utilises aggregated data.[49] Warrantless searches are also possible through the use of *zero days* — intentional corporate cyber security flaws through which politicians breach encryption.[50] The legality of these surveillance tactics does not render them ethical or constitutional. Tunick corroborates this view, noting parallels between these techniques and those deployed by the Gestapo.[51]

The combined powers of PeekYou, GPS, and facial recognition do indeed allow the government to track mass groups of people in real time. In fact, 100,000 people's movements were traced through mobile phone data triangulation. Bayir reports that in addition to location tracking, the data from these same people's mobile phones was used to extrapolate information about their habits, friends and activities.[52] Data mining could potentially endanger innocent people. For instance, if one is mistaken for a terrorist, it could be possible to be placed on a no-fly list or treated to legalised brutalities such as extraordinary rendition.

DEMOCRACY AND CONSTITUTIONALITY OF THE TECHNOLOGIES

Friedland cites the Third Amendment's protection of citizens from US military home intrusion.[53] Today, this amendment should protect US citizens from government's bulk collection of our online data — much of which is created in citizens' own homes. According to findlaw.com, the Fourth Amendment guarantees the Constitutional right to personal privacy at home, at work and about our persons.[54] Henceforth, the police must obtain a search warrant before investigating a citizen's private affairs.[55] Yet, there are no search warrants attached to licence plate scanners, GPS signal data, facial recognition software, surveillance cameras and PeekYou. Have you ever heard this rap: 'You have the right to remain silent. You have the right to speak to an attorney. Anything you say can and will be used against you'? These rights embody the Fifth Amendment Miranda Rights, which provides those in police custody with protections from self-incrimination.[56] Fifth Amendment rights are subverted by the ubiquity of surveillance cameras when surreptitious information from facial recognition programmes establishes culpability. Friedland relays the import of the Miranda doctrine with regards to testimony obtained under duress.[57] Since we are forced to have our faces scanned, are tracked by GPS and monitored online, one could say that the entirety of our interconnected lives should be protected by the Third, Fourth and Fifth Amendments.

CONCLUSION

The combined surveillance capabilities of GPS, facial recognition apps, people tracking technologies and PeekYou are already covertly embedded within the fibre of our society. Nevertheless, by upholding preexisting Amendments and introducing new privacy legislation specific to these technologies, we can curtail the insidious spread of coerced assent to mass surveillance.[58,59] These technologies hold the potential for an unprecedented coordination of resources for the greater good. It is never too late to alter the course of history such that we amplify these technology's attributes while protecting all people from the abuse inherent in their utilisation.

References

1. NPR (2018) 'Former TSA administrator discusses "Quiet Skies" surveillance program', All Things Considered, available at: https://www.npr.org/2018/07/30/634087400/former-tsa-administrator-discusses-quiet-skies-surveillance-program (accessed 13th May, 2019).
2. Hausman, S. (2014) 'Proliferation of tracking technology reaps security benefits, sparks privacy concerns', SecurityInfoWatch.Com, available at: https://www.securityinfowatch.com/alarms-monitoring/asset-and-gps-tracking/article/11350683/dr-steven-hausman-examines-the-benefits-and-pitfalls-of-the-proliferation-of-tracking-technology (accessed 9th May, 2019). See p. 1, para 1.
3. Dunn, J. E. (2018) 'Google is tracking your location, even when the setting is turned off', *Naked Security by Sophos*, available at: https://nakedsecurity.sophos.com/2018/08/15/google-is-tracking-your-location-even-when-the-setting-is-turned-off/ (accessed 9th May, 2019).
4. Bayir, M. A. (2010) 'Enabling location aware smartphone applications via mobility profiling', Doctoral dissertation, State University of New York at Buffalo. See p. 7 & p. 9.
5. Kan, M. (2018) 'Will Amazon's facial-recognition tech enable mass surveillance?', *PC Mag*, available at: https://www.pcmag.com/news/361346/will-amazons-facial-recognition-tech-enable-mass-surveillance (accessed 9th May, 2019).
6. Gibbons, A. and Marsh, S. (2018) 'Chelsea Manning says life in the US is like being in prison', *The Guardian*, available at: https://www.theguardian.com/us-news/2018/oct/01/chelsea-manning-life-us-like-prison-uk-visit (accessed 8th May, 2019).
7. Ochab, E. U. (2018) 'Human trafficking is a pandemic of the 21st century', *Forbes*, available at: https://www.forbes.com/sites/ewelinaochab/2018/07/26/human-trafficking-is-a-pandemic-of-the-21st-century/#4b0102906195 (accessed 8th May, 2019).
8. Gibbons & Marsh, ref. 6 above, p. 1.
9. Barocas, S. (2014) 'Panic inducing: Data mining, fairness, and privacy', PhD dissertation, New York University, USA.
10. *Ibid*.
11. *Ibid*.
12. *Ibid*., p. 19.
13. Weisinger, D. (2017) 'Big data and data anonymization: is anonymization an illusion?', Formtek, para 4, available at: https://formtek.com/blog/big-data-and-data-anonymization-is-anonymization-an-illusion/ (accessed 10th May, 2019).
14. Barocas, ref. 9 above, p. 8.
15. Leonard, P. (2014) 'Customer data analytics: Privacy settings for "big data" business. *International Data Privacy Law*, Vol. 4, No. 1, pp. 53–68, doi: http://dx.doi.org/10.1093/idpl/ipt032. See p. 63.
16. Cavoukian, A. and Jonas, J. (2012) 'Privacy by design in the age of big data', 8th June, available at: https://jeffjonas.typepad.com/Privacy-by-Design-in-the-Era-of-Big-Data.pdf (accessed 9th May, 2019).
17. Bell, G. (2012) 'Interview with Marshall Sponder, author of social media analytics', *Strategic Direction*, Vol. 28, No. 6, pp. 32–35, available at: https://www.emeraldinsight.com/doi/abs/10.1108/02580541211224102 para 1 (accessed 10th May, 2019).
18. *Ibid*., para 15.
19. Hussey, M. (2007) 'PeekYou emerges from stealth mode with 50 million profiles in beta, providing easy and efficient people search capabilities', *PRNewswire*, 17th July, available at: http://michaelhussey.com/2007/07/17/peekyoucom-beta-launch/ (accessed 7th May, 2019).
20. *Ibid*.
21. Hardin, N. (2017) 'Cell phone surveillance: Tactics, litigation, and next steps', p. 9, para 1, available at: https://vae.fd.org/sites/vae.fd.org/files/training/April_2018/03%20Cell%20Phone%20Surveillance.pdf (accessed 10th May, 2019). Note: this is an unpublished PDF uploaded by Nicole Hardin for public use; she is an assistant federal public defender for the Middle District of Tampa, FL, USA.
22. McMurrer, S. and Newburn, M. (2018) 'Next Generation 9-1-1 update', Board of Supervisors IT Committee Meeting, available at: https://www.fairfaxcounty.gov/boardofsupervisors/sites/boardofsupervisors/files/assets/meeting-materials/2018/oct09-it-next-generation-911-presentation.pdf (accessed 9th May, 2019).
23. Friedland, S. I. (2015) 'I spy: The new self-cybersurveillance', *Washington and Lee Law Review*, Vol. 72, No. 3, pp. 1459–1501, available at: https://heinonline.org/HOL/LandingPage?handle=hein.journals/waslee72&div=35&id=&page=&t=1557365135 (accessed 9th May, 2019), see p.1471.
24. Hausman, ref. 2 above.
25. Friedland, ref. 23 above, p. 1463 & pp. 1473–1474.
26. MassDOT (2016) 'All electronic tolling now activated along I-90', 29th October, archives, available at: http://blog.mass.gov/transportation/massdot-highway/all-electronic-tolling-now-activated-along-i-90/ (accessed 9th May, 2019).
27. Hausman, ref. 2 above, p. 1, para 1.
28. Max, R. (2018) 'People tracking: 15 technologies in 2018 — What's people tracking technology? *Behavior Analytics Retail*, available at: https://behavioranalyticsretail.com/technologies-tracking-people/ (accessed 10th May, 2019).
29. Hausman, ref. 2 above.
30. Max, ref. 28 above.
31. Celentino, J. C. (2016) 'Face-to-face with facial recognition evidence: admissibility under the post-Crawford Confrontation Clause', *Michigan Law Review*, Vol. 114, No. 7, pp. 1317–1353. See p. 1322.
32. Agre, P. (2001) 'Your face is not a barcode — Arguments against automatic face recognition in public places', available at: http://polaris.gseis.ucla.edu/pagre/bar-code.html (accessed 10th May, 2019).
33. Celentino, ref. 33 above, p. 1322.
34. Agre, ref. 32 above, para 11.
35. Gemalto (2018) 'The top 7 trends for facial recognition in 2018', available at: https://www.gemalto.com/govt/biometrics/facial-recognition (accessed 10th May, 2019).
36. Friedland, ref. 23 above, p. 1489.

37. Agre, ref. 32 above, para 7.
38. Hausman, ref. 2 above, p. 1, para 4.
39. Barocas, ref. 9 above, p. 8.
40. Hausman, ref. 2 above.
41. Zenus (2018) 'Facial recognition and events — a comprehensive guide', Event — The Intelligence Platform to Run Better Events, available at: https://www.eventmanagerblog.com/facial-recognition-guide-2018 (accessed 9th May, 2019).
42. Thompson, C.W. and Thompson, D. R. (2007) 'Identity management', *IEEE Internet Computing*, Vol. 11, No. 3, pp. 82–85, doi: http://dx.doi.org/10.1109/MIC.2007.60.
43. *Ibid*.
44. Friedland, ref. 23 above, p. 1475.
45. Timberg, C. (2014) 'What do firms know about you? FTC would pull back the curtain', *The Washington Post*, 28th May, p. 2.
46. Alim, F., Cardozo, N., Gebhart, G., Gullo, K. and Kalia, A. (2017) 'Spying on students: School-issued devices and student privacy', The Electronic Frontier Foundation, available at: https://www.eff.org/wp/school-issued-devicesand-student-privacy (accessed 10th May, 2019).
47. Tunick, M. (2009) 'Privacy in public places: Do GPS and video surveillance provide plain views?', *Social Theory and Practice*, Vol. 35, No. 4, pp. 597–622. See p. 597.
48. Friedland, ref. 23 above, p. 1466.
49. *Ibid*.
50. *Ibid*., pp.1465–1466.
51. Tunick, ref. 47 above, p. 611.
52. Bayir, ref. 4 above.
53. Friedland, ref. 23 above, p. 1498.
54. FindLaw.com (2016) 'Search and seizure and the Fourth Amendment', Thomson Reuters, available at: http://criminal.findlaw.com/criminal-rights/search-and-seizure-and-the-fourth-amendment.html (accessed 10th May, 2019).
55. Friedland, ref. 23 above, p. 1466.
56. 'What is the definition of the Miranda Doctrine?' (2016) IAC Publishing Labs Company, available at: https://www.reference.com/government-politics/definition-miranda-doctrine-530313a00f9e5aee (accessed 10th May, 2019).
57. Friedland, ref. 23 above, p. 1492.
58. Hausman, ref. 2 above, p. 1, para 1.
59. Agre, ref. 32 above, para 12.