ARTICLE

Protection for 'Inferences Drawn': A Comparison Between the General Data Protection Regulation and the California Consumer Privacy Act

Jordan M. Blanke*

Inferences drawn from personal data have arguably become more dangerous to individual privacy than the vast collection and storage of the data itself. Recently there have been questions raised about whether the General Data Protection Regulation (GDPR) has sufficient protection for these inferences. Probably not surprisingly, and learning from this possible shortcoming, the California Consumer Privacy Act (CCPA) specifically includes 'inferences drawn' as part of its definition of personal information. This article explores the widespread use of inferential data and compares the protection provided under the GDPR and the CCPA for such inferences.

Keywords: privacy, data protection, inferences drawn, GDPR, CCPA

1 Introduction

A relatively new area of concern regarding privacy and data protection is the increasing use made of inferences drawn upon the vast amounts of personal information that is routinely collected. Ever since the early days of what has become the era of big data, people have massaged data with hopes of gaining insight into human behaviour. While the European Union has generally been far ahead of the United States in protecting data with respect to its collection and use, the sophistication of analytical tools may have outpaced the protection provided by the General Data Protection Regulation (GDPR). While the GDPR clearly provides protection for the data itself, there may not be sufficient protection for some of the inferences drawn from that data. Probably not surprisingly, the brand-new California Consumer Privacy Act (CCPA) specifically addresses this issue. While there are a number of articles that examine the similarities and the differences between the GDPR and the CCPA, this is one of the first that focusses on *inferences drawn* from data.

2 Part I

2.1 *Background*

One of the first reported instances of the analytical use of big data goes back to 2002, when J.P. Martin, a numbers-oriented executive at Canadian Tire, decided to see what kind of information he could extract from the data that his company collected from credit-card transactions.[1] What he found was both interesting and surprising and would eventually give rise to a whole slew of new industries. Among other things, he discovered that people who bought furniture pads, carbon-monoxide detectors, premium birdseed, or snow roof rakes were very good credit risks, but that people who bought cheap motor oil were not.[2] And one should not even consider extending credit to someone who would buy a *Mega Thruster Exhaust System.*[3]

Around the same time that Martin started playing with those numbers, we were introduced to the notion of predictive analysis by the 2002 film *Minority Report*,[4] based upon a 1956 short story of the same name by Phillip K. Dick.[5] In the film, a PreCrime division of the police

Notes

* Jordan 'Jody' Blanke is the Ernest L. Baskin, Jr. Distinguished Professor of Computer Science and Law at the Stetson-Hatcher School of Business at Mercer University in Atlanta. Email: BLANKE_J@mercer.edu.

1 Charles Duhigg, *What Does Your Credit-Card Company Know About You?*, The New York Times Magazine (12 May 2009).

2 *Ibid.*

3 *Ibid.*

4 *Minority Report* (20th Century Fox 2002).

5 Phillip K. Dick, *The Minority Report*, 4(6) Fantastic Universe (1956).

department is charged with preventing the future acts of would-be criminals before they are able to commit those crimes. While much of PreCrime's information comes from three so-called 'precogs' who are able to see the future, some of it comes from a collection of biometric data from iris scanners and from voluminous databases. In a famous scene from the movie, the lead character, who underwent an iris transplant in order to avoid detection of his true identity, is greeted in a store by a holographic image of a woman: 'Hello, Mr Yakamoto. Welcome back to the Gap. How did those assorted tank tops work out for you'?[6] The film is set in 2054. It certainly appears that we, as a society, in 2020, are way ahead of schedule, at least with regard to the ubiquitous collection of data and immediate access to massive databases from which numerous inferences can be drawn.

In what was probably the most widely publicized episode illustrating both the effect and the accuracy of predictive analysis, Target famously predicted, in 2012, that a teenaged girl was pregnant.[7] Its researchers discovered that women just beginning their second trimesters often bought unscented lotion and that during the first twenty weeks of a pregnancy, they often bought calcium, magnesium and zinc supplements. From these purchases, Target could infer that such a woman was likely pregnant. Accordingly, Target mailed a letter to the home of the teenaged girl who it predicted was pregnant, congratulating her on her pregnancy and offering her a variety of coupons and offers. The girl's father, who happened to open the letter, was not amused. After initially being outraged by Target's direct mailing, he later learned that Target was, in fact, correct about its inference. His daughter was pregnant.

In a study from 2013, researchers at Cambridge University were able to predict, with startling accuracy, a number of sensitive personal attributes on a the basis of Facebook *likes*.[8] Using data collected from 58,466 American volunteers through the myPersonality Facebook app and an average of 170 Facebook *likes* per volunteer, the researchers were able to predict the race of a person with 95% accuracy, the gender of a person with 93% accuracy, whether a person was a Democrat or a Republican 85% of the time, whether a person was Christian or Muslim 82% of the time, and whether a man or a woman was gay, 88% and 75% of the time, respectively. Cigarette use, alcohol use, and drug use were also predicted at fairly high rates of success, 73%, 70%, and 67%, respectively.

2.2 *Data Analytics Today*

There are a number of interrelated processes that are involved in modern data analytics:

(1) The collection of data, through either direct means, from people themselves, or indirect means, using a variety of collection and tracking technologies via the Internet, mobile phones or sensors (often referred to, collectively, as The Internet of Things);
(2) The building of profiles from the data collected; and
(3) The use of inferences drawn from the data collected to both build the profiles and to predict future behaviour based upon the data and the profiles.

2.3 *Collection of Data*

Companies have been collecting consumer data for decades, if not centuries. Until the advent of digital technology and the Internet, however, collection was fairly inefficient. As computers become more powerful and less expensive and database technology more sophisticated, the collection and aggregation of data increased exponentially.[9]

Even before the Internet became what it is today, there was concern about protecting personal data or information. In the United States, in 1970, the Fair Credit Reporting Act (FCRA) recognized the potential consequences of vast amounts of sensitive information collected and stored in databases.[10] In 1974, the Family Education Rights and Privacy Act gave students and their parents certain rights in their educational records.[11] By 1988 a number of concerned Congressmen learned that video stores actually kept records of which videotapes were rented and by whom, and very quickly passed the Video Privacy Protection Act (VPPA).[12] In 1995, around the time that the World Wide Web was beginning to become popular, the European Union passed the Directive on Data Protection, the first significant comprehensive law to

Notes

[6] *Minority Report*, *supra* n. 2.

[7] Charles Duhigg, *How Companies Learn Your Secrets*, The New York Times Magazine (16 Feb. 2012).

[8] Michal Kosinski, David Stillwell & Thore Graepel, *Private Traits and Attributes Are Predictable from Digital Records of Human Behavior,* 110(15) Proc. Nat'l Acad. Sci. 5802–5805 (9 Apr. 2013), https://www.pnas.org/content/110/15/5802 (accessed 4 May 2020).

[9] *See generally* Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information* (2015); Viktor Mayer-Schonberger & Kenneth Cukier, *Big Data: A Revolution that Will Transform How We Live, Work, and Think* (2013); Scott R. Peppet, *Regulating the Internet of Things: First Steps Towards Managing Discrimination, Privacy, Security, & Consent*, 93 Tex. L. Rev. 85 (2014).

[10] 15 U.S.C. § 1681 (1970).

[11] 20 U.S.C. § 1232g (1974).

[12] 18 U.S.C. § 2710 (1988).

protect personal data.[13] In 1996, the Health Insurance Portability and Accountability Act (HIPAA) was passed, finally providing protection for medical records.[14] In 1998, the Children's Online Privacy Protection Act (COPPA) provided comprehensive privacy protection on the Internet – but only for children under the age of thirteen.[15]

In the early days of the Web, most data was collected directly from the user. Many web sites actually provided a legitimate opportunity for notice and consent. Unfortunately, as the Web matured, so did the technology, and many web sites began using a variety of tracking tools, like cookies, web beacons, tracking bugs, and pixel tags to surreptitiously collect personal information from users. Similarly with smartphones, a wide variety of personal information is collected from different components of the technology. Most users have no idea how much information is collected every time they use their telephone or any of the apps installed thereon.

2.4 *Profiles*

As it became easier and more cost efficient for companies to harvest information from the Web, smartphones, and other sources of data, the data broker industry grew enormously. In 2012 it was reported that Acxiom executives stated that 'its database contains information about 500 million active consumers worldwide, with more than 1,500 data points per person'.[16] By 2014 it was reported that Acxiom stated that '[f]or every consumer we have more than 5,000 attributes of customer data'.[17]

In 2017 Wolfie Christl prepared a report on *Corporate Surveillance in Everyday Life.*[18] On the web site discussing his findings, he summarized the astounding number of profiles maintained by *Large Online Platforms* (Facebook has profiles on 1.9 billion Facebook users, 1.2 billion WhatsApp users, and 600 million profiles on Instagram users; Google has profiles on 2 billion Android users, 1+ billion Gmail users, and 1+ billion YouTube users; Apple has profiles on 1 billion iOS users); *Credit Reporting Agencies* (Experian has credit data on 918 million people and marketing data on 700 million people; Equifax has data on 820 million people; and TransUnion has data on 1 billion people); and *Consumer Data Brokers* (Acxiom has data on 700 million people and from 1 billion cookies and mobile devices, and it manages 3.7 consumer profiles for its clients; and Oracle has data on 1 billion mobile users and 1.9 billion web site visitors, and provides access to 5 billion 'unique' customer IDs).[19]

One of the concerns about profiles is whether the data contained therein was collected legally. Another concern is whether the data should be maintained in the profile and if so, for how long. For these concerns, as will be discussed below, the GDPR provides way more protection than most US laws, other than the sectoral protection provided by the FCRA, FERPA, the VPPA, COPPA, and HIPAA. Also, as will be discussed below, Californians now have some legal rights in this regard, too.

Another concern about profiles is whether the data contained in them are accurate or correct. To some extent, this may be a double-edged sword. One tactic that people have used to attempt to combat the inequity of the relentless collection of data, is to purposely provide incorrect information, for example, providing different dates of birth or different telephone numbers to different web sites. While this strategy may successfully prevent an accurate profile from being maintained by data brokers, it is often the inadvertently inaccurate data that provides the most damage. Depending upon the applicable law, individuals may or may not have the right to inspect the data contained in their profiles, and they may or may not have the right to demand correction of inaccurate information. In an article describing one writer's frustration with trying to request information about herself from data brokers, she wrote that '[I was] equally irked by the reports that were wrong – data brokers who thought I was a single mother with no education – as I was by the ones that were correct – is it necessary for someone to track that I recently bought underwear online'?[20]

The accuracy of data maintained in a profile is also important when it comes to the inferences drawn from *incorrect* information. Certainly, the old computer adage, 'garbage in, garbage out' would apply.

Notes

[13] Directive 95/46 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995 O.J. (L 281) 1, 1 (EC).

[14] Pub. L. 104–191 (1996); 45 C.F.R. § 160.203 (2002).

[15] 15 U.S.C. § 6501 (1998).

[16] Natasha Singer, *Mapping, and Sharing, the Human Genome*, N.Y. Times (16 June 2012), https://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html (accessed 4 May 2020).

[17] Jeff Chester, *Acxiom: 'For Every Consumer We Have More than 5,000 Attributes of Customer Data'*, Center for Digital Democracy (10 Jan. 2014), https://www.democraticmedia.org/acxiom-every-consumer-we-have-more-5000-attributes-customer-data (accessed 4 May 2020).

[18] Wolfie Christl, *Corporate Surveillance in Everyday Life: How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions* (June 2017), https://crackedlabs.org/en/corporate-surveillance/#3 (accessed 4 May 2020).

[19] https://crackedlabs.org/en/corporate-surveillance/#3 (accessed 4 May 2020).

[20] Lois Beckett, *Everything We Know About What Data Brokers Know About You*, (13 June 2014), https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you (accessed 4 May 2020), (describing Julia Angwin's experience requesting her information from data brokers).

2.5 *'Inferences Drawn'*

The technology of predictive analysis – as well as its use – has increased dramatically over the last several years. A report by Privacy International, *Examples of Data Points Used in Profiling*, describes a variety of inferences made about and from different types of data.[21] The report 'mostly draws from computer science literature to (1) show types of sensitive information that can be inferred through the analysis of common forms of data, [and] (2) illustrate concrete harms that these inferences can produce'.[22]

Among the examples discussed are those involving studies that inferred ones identity from publicly available information. In one study, information from newspaper articles 'uniquely and exactly matched medical records in the state database for thirty-five of the eighty-one cases (or 43%) found in 2011, thereby putting names to patient records'.[23] In another one, a researcher revisited a prior study to find a lower, albeit still significant, percentage (63%) of people identified solely from their gender, zip code and date of birth.[24]

Regarding inferences revealing personal information, one study demonstrated how Social Security numbers could be inferred from birth data and readily available information from data brokers and social network profiles.[25] The Kosinski study, discussed above, showed how Facebook likes can infer private traits and characteristics, such as gender, religion, political affiliation, and sexual orientation.[26] Another study showed how publicly available geographic information from Tweets could accurately infer 'average income based on one's neighborhood, average housing cost, debt, and other demographic information, such as political views'.[27]

Another important concern about data analytics and predictive analytics is that they may perpetuate or develop discriminatory practices.[28] One paper stated that 'rather than correcting for the apparent biases in the police data, the model reinforces these biases'.[29] Using a certain modelling tool in Oakland, 'black people would be targeted by predictive policing at roughly twice the rate of whites. Individuals classified as a race other than white or black would receive targeted policing at a rate 1.5 times that of whites'.[30] In other studies, discrimination effected Uber usage 'in Seattle through longer waiting times for African American passengers – as much as a 35% increase' and in Boston 'via more frequent cancellations against passengers when they used African American-sounding names'.[31]

Yet another issue that can dramatically affect the validity of an inference drawn is the importance of the difference between correlation and causation. Analytical tools today can reveal a great number of statistical correlations. Those correlations may or may not represent a causal relationship. While it seems rather likely, and makes some good sense, for example, that a person who is concerned enough about the condition of his or her floors would buy furniture pads, it seems less likely that the consumption of cheese is related to the number of people who die because they become entangled in their bedsheets, or that the consumption of more margarine may lead to a higher divorce rate in the state of Maine. Yet for each of these two latter examples, there is a very strong statistical correlation. There are many similar humorous examples of these strong, albeit dubious, relationships on a web site called *Spurious Correlations.*[32]

Inferences drawn from data can be problematic both in building a profile and in later extracting information from it. As data is collected about a person, inferences may be drawn from the data and stored as part of the profile as if it were independently collected data. Unless there is a distinction made in the profile about which data is *raw*

Notes

[21] Privacy International, *Examples of Data Points Used in Profiling* (2017), https://privacyinternational.org/sites/default/files/2018-04/data%20points%20used%20in%20tracking_0.pdf (accessed 4 May 2020).

[22] *Ibid.*

[23] *Ibid.*, at 3–4. *See* Latanya Sweeney, *Uniqueness of Simple Demographics in the U.S. Population*, Laboratory for Int'l Data Privacy, Working Paper LIDAP-WP4 (2000).

[24] *Ibid.*, at 4. *See* Philippe Golle, *Revisiting the Uniqueness of Simple Demographics in the US Population*, 5 ACM Workshop on Privacy Elec. Soc'y 77, 78 (2006).

[25] *Ibid.*, at 8. *See* Alesandro Acquisti & Ralph Gross, *Predicting Social Security Numbers from Public Data*, 106 Nat'l Acad. Sci. 10975 (2009).

[26] *Ibid.*, at 10. *See* Michal Kosinski, David Stillwell & Thore Graepel, *Private Traits and Attributes are Predictable from Digital Records of Human Behavior*, 110 Nat'l Acad. Sci. 5802 (2013).

[27] *Ibid.*, at 20. *See* Ilaria Liccardi, Alfie Abdul-Rahman & Min Chen, *I Know Where You Live: Inferring Details of People's Lives by Visualizing Publicly Shared Location Data*, Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 1–12 (May 2016), https://doi.org/10.1145/2858036.2858272 (accessed 4 May 2020).

[28] *See generally* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 Calif. L. Rev. 671 (2016); Tal Z. Zarsky, *Understanding Discrimination in the Scored Society*, 89 Wash. L. Rev. 1375 (2014); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 Wash. L. Rev. 1 (2014); James Grimmelmann & Daniel Westreich, *Incomprehensible Discrimination*, 7 Calif. L .Rev. Online 164, 170 (2017); Margot E. Kaminski & Andrew D. Selbst, *Opinion, The Legislation That Targets the Racist Impacts of Tech*, NY TIMES: THE PRIVACY PROJECT (7 May 2019), https://www.nytimes.com/2019/05/07/opinion/tech-racism-algorithms.html (accessed 4 May 2020).

[29] Privacy International, *supra* n. 21, at 28. *See* Kristian Lum & William Isaac, *To Predict and Serve?*, 13 Significance 14, 18 (2016).

[30] *Ibid.*

[31] *Ibid.*, at 30. *See* Yanbo Ge, Christopher R. Knittel, Don MacKenzie & Stephen Zoepf, *Racial and Gender Discrimination in Transportation Network Companies*, Nat'l Bureau Econ. Research, Working Paper No. 22776 (2016), https://www.nber.org/papers/w22776 (accessed 4 May 2020).

[32] https://www.tylervigen.com/spurious-correlations (accessed 4 May 2020).

and which is inferred, all of it may appear to be *raw*. Certainly, all of the problems already discussed about inferences drawn will apply to this data that has now been saved to the profile and will likely thereafter be considered as factual and verified data that becomes a permanent and persistent part of that profile.

Even if inferences are not saved as part of one's profile, the inferences drawn from the data are all subject to the problems already discussed – the data may be inaccurate, the analytical tools or models may be faulty or may produce discriminatory results, the presumed relationship may be merely coincidental rather than causal, or the methods used or results produced may violate the law. As Omer Tene and Jules Polonetsky observed in 2013, in the world of big data, 'what calls for scrutiny is often not the accuracy of the *raw data* but rather the accuracy of the *inferences* drawn from the data'.[33]

3 Part II

3.1 *The GDPR*

The GDPR and its predecessor, Directive on Data Protection,[34] have been the bellwether standard for data protection for many years. Among the very many important recitals and definitions in the GDPR are these two definitions that are particularly relevant to this article:

> 'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[35]; and
>
> 'profiling' means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements;[36]

Among the many rights given to data subjects under the GDPR are the right to access,[37] the right to rectification,[38] the right to erasure (the 'right to be forgotten'),[39] the right to restriction of processing,[40] and the right to data portability.[41] Particularly relevant to this article, Article 21 provides that '[w]here personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to processing of personal data concerning him or her for such marketing, which includes *profiling* to the extent that it is related to such direct marketing'.[42] Also, Article 22 provides that the 'data subject shall have the right not to be subject to a decision based *solely* on *automated processing*, including *profiling*, which produces legal effects concerning him or her or similarly significantly affects him or her'.[43]

While the GDPR provides probably the most extensive coverage for data protection in the world, there are those who believe that it did not go far enough in protecting data inferences, which are often generated by algorithms and automated processing.[44] In *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*,[45] Sandra Wachter and Brett Mittelstadt present an argument that (1) the GDPR grants individuals 'little control or oversight over how their personal data is used to draw inferences about them',[46] (2) the GDPR 'provides insufficient protection against sensitive inferences … or the remedies to challenge inferences or important decisions based upon them',[47] (3) the European Court of Justice (ECJ) has

Notes

[33] Omer Tene & Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 11 Nw. J. Tech. & Intell. Prop. 239, 270 (2013).

[34] Directive 95/46, Art. 25, 1995 O.J. (L 281) 31, 56–57 (EC).

[35] GDPR, at Art. 4(1).

[36] GDPR, at Art. 4(4).

[37] GDPR, at Art. 15.

[38] GDPR, at Art. 16.

[39] GDPR, at Art. 17.

[40] GDPR, at Art. 18.

[41] GDPR, at Art. 20.

[42] GDPR, at Art. 21 (emphasis added).

[43] GDPR, at Art. 22(3) (emphasis added).

[44] Sandra Wachter & Brett Mittelsadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2019(2) Colum. Bus. L. Rev. 494 (2019); Howard Yu, *GDPR Isn't Enough to Protect Us in an Age of Smart Algorithms* (2018), https://theconversation.com/gdpr-isnt-enough-to-protect-us-in-an-age-of-smart-algorithms-97389 (accessed 4 May 2020).

[45] *Ibid.*

[46] *Ibid.*, at 494.

[47] *Ibid.*, at 495.

'made clear that data protection law is not intended to ensure the accuracy of decisions and the decision making processes involving personal data, or to make these processes fully transparent',[48] and that, therefore, (4) 'a new data protection right, "the right to reasonable inferences," is needed to help close the accountability gap currently posed by "high risk inferences"'.[49]

Wachter and Mittelstadt are concerned that '[p]ersistent records can be created through inferential analytics, consisting of unpredictable and potentially troubling inferences revealing information and predictions about private life, behaviors, and preferences that would otherwise remain private'.[50] They believe that the GDPR has sufficient mechanisms for managing the input side of processing, but may not have sufficient protection for the output side.[51]

The authors discuss that it is not clear whether – or how much – inferences are protected by the GDPR.[52] Under guidelines created by the old Article 29 Working Party, inferences would be considered as personal data under Article 4 of the GDPR. The Article 29 Working Party distinguished among four types of data:

(1) 'provided data' – data that is directly provided by the data subject to the data controller, like mailing address, user name, or age,
(2) 'observed data' – data that is indirectly or passively provided by the data subject, like raw data processed by a smart meter, geolocation data, or keystroke dynamics,
(3) 'derived data' – data that is created by the data controller based upon data provided by the data subject, like state of residence from zip code, and
(4) 'inferred data' – data that is created by the data controller based upon data provided by the data subject using some sort of analytics, like credit rating or health risk.[53]

Obviously, the last category is most relevant to this discussion. Applying a previous Article 29 Working Party opinion on the concept of personal data, one would probably conclude that inferred data is Article 4 personal data because the result of maintaining that data would likely have an impact on the rights and interests of the data subject.[54]

Furthermore, in other guidelines, the Article 29 Working Party discussed the interplay between automated decision-making and profiling, stating that 'the process of profiling is often invisible to the data subject. It works by creating derived or inferred data about the individuals – "new" data that has not been provided directly by data subjects'.[55] It also stated that such collection and use of this data would likely implicate several basic principles of the Directive, including (Article 5(1)(b)) further processing and purpose limitation, (Article 5(1)(c)) data minimization, and (Article 5(1)(d)) accuracy.[56]

The Article 29 Working Party also addressed, in another opinion, how various anonymization and pseudonymization techniques would affect inference risks or inference attacks.[57] It defined an inference risk as when there is a 'possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes'.[58] It discussed several anonymization techniques, including three randomization techniques – noise addition, permutation, and differential privacy – and two generalization techniques – aggregation and K-anonymity, and L-diversity/T-closeness,[59] and concluded that, while some of the techniques might be helpful in reducing such risks, inference risks still existed even with their use.[60] Similarly, no matter which type of pseudonymization is used (keys, hashing, tokens, and

Notes

[48] *Ibid.*

[49] *Ibid.*

[50] *Ibid.*, at 513.

[51] *Ibid.*, at 513–514.

[52] *Ibid.*, at 515–521.

[53] *Ibid.*, at 515–517; Art. 29 Data Prot. Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679*, WP251rev.01, 9–10 (6 Feb. 2018), https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053 (accessed 4 May 2020).

[54] *Ibid.*, at 515–517; Art. 29 Data Prot. Working Party, *Opinion 4/2007 on the Concept of Personal Data*, 01248/07/EN WP136, 8 (20 June 2007), https://www.clinicalstudydatarequest.com/Documents/Privacy-European-guidance.pdf (accessed 4 May 2020).

[55] Article 29 Data Prot. Working Party, *Guidelines on the Right to Data Portability*, 16/EN, WP242rev.01, 9–11 (13 Dec. 2016), https://ec.europa.eu/newsroom/document.cfm?doc_id=44099 (accessed 4 May 2020).

[56] *Ibid.*, at 9–15. *See generally* Margot E. Kaminski, *Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability*, 92 S. Cal. L. Rev. 1529 (2019); Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 Fordham L. Rev. 1085, 1100–1105 (2018). Deven R. Desai & Joshua A. Kroll, *Trust But Verify: A Guide to Algorithms and the Law*, 31 Harv. J.L. Tech. 1, 5–6 (2017); Neil M. Richards & Woodrow Hartzog, *Privacy's Trust Gap: A Review*, 126 Yale L.J. 1180 (2017).

[57] Article 29 Data Prot. Working Party, *Opinion on Anonymisation Techniques*, WP216 (10 Apr. 2014), https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/index_en.htm (accessed 4 May 2020).

[58] *Ibid.*, at 11–12.

[59] *Ibid.*, at 11–19.

[60] *Ibid.*, at 23–24.

combinations thereof),[61] inference risks likely still exist.[62] Clearly, the Article 29 Working Party anticipated that inferences should be considered to be personal data.

More recently, the Council of Europe adopted Convention 108+ (Convention).[63] While not binding on the EU, the Convention elaborates on several basic principles of the GDPR that are largely consistent with the GDPR and the views of the Article 29 Working Party and likely to be followed by Member States. While not specifically addressing inferences, the Convention has several related provisions regarding anonymization and pseudonymization:

> 'Identifiable individual' means a person who can be directly or indirectly identified. An individual is not considered 'identifiable' if his or her identification would require unreasonable time, effort or resources.[64]
>
> The use of a pseudonym or of any digital identifier/digital identity does not lead to anonymization of the data as the data subject can still be identifiable or individualised. Pseudonymous data is thus to be considered as personal data and is covered by the provisions of the Convention.[65]
>
> Data is to be considered as anonymous only as long as it is impossible to re-identify the data subject or if such re-identification would require unreasonable time, effort or resources, taking into consideration the available technology at the time of the processing and technological developments.[66]
>
> When data is made anonymous, appropriate means should be put in place to avoid re-identification of data subjects, in particular, all technical means should be implemented in order to guarantee that the individual is not, or is no longer, identifiable.[67]
>
> 'Data processing' starts from the collection of personal data and covers all operations performed on personal data, whether partially or totally automated.[68]

Wachter and Mittelstadt argue that '[w]hile the legally non-binding guidelines of the Article 29 Working Party clearly endorse the view that inferences are personal data, the legally binding jurisprudence of the European Court of Justice (ECJ) is less generous in its interpretation'.[69] In a case deciding two separate applications for a residence permit, the ECJ distinguished between the personal data contained in the legal analysis and the legal analysis itself.[70] The ECJ permitted the release of a summary containing all the personal data considered in the decision, rather than a full-text of the decision.[71] Wachter and Mittelstadt are concerned about this because the 'ECJ's judgement makes clear that … the analysis and constituent inferences are not considered personal data'.[72] They are also troubled because the ECJ seemed to rule that 'data protection law in general, and the right of access in particular, are not designed to provide full transparency in decision-making involving personal data, or to guarantee 'good administrative practices'.[73] Finally, and probably most significantly, Wachter and Mittelstadt fear that, 'according to the EJC, when a private company draws inferences from collected data or makes decisions based on them, even if the final inferences or decisions are seen as personal data, data subjects are unable to rectify them under data protection law'.[74]

In a later case, the ECJ seemed a little more willing to expand the scope of the definition of personal data, holding that the 'use of the expression "any information" in the definition of the concept of "personal data", within [the Directive] reflects the aim of the EU legislature to assign a wide scope to that concept … provided that it "relates" to the data subject'.[75] The ECJ held that both the answers

Notes

[61] *Ibid.*, at 20–23.

[62] *Ibid.*, at 23–24.

[63] Council of Europe, Convention 108+: Convention for the protection of individuals with regard to the processing of personal data (18 May 2018), https://www.coe.int/en/web/data-protection/convention108/modernised (accessed 4 May 2020).

[64] Explanatory Report of Convention 108+, para. 17, https://rm.coe.int/cets-223-explanatory-report-to-the-protocol-amending-the-convention-fo/16808ac91a (accessed 4 May 2020).

[65] *Ibid.*, para. 18.

[66] *Ibid.*, para. 19.

[67] *Ibid.*, para. 20. *See* Paul Ohm, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, 57 UCLA L. Rev. 1701 (2010); Ira Rubinstein & Woodrow Hartzog, *Anonymization and Risk*, 91 Wash. L. Rev. 703 (2016).

[68] *Ibid.*, para. 21.

[69] Wachter & Mittelstadt, *supra* n. 44, at 521.

[70] Joined Cases C-141 & 372/12, *YS, M and S v. Minister voor Immigratie, Itegratie en Asiel*, 2014 E.C.R. I-2081, para. 48.

[71] *Ibid.*, para. 59.

[72] Wachter & Mittelstadt, *supra* n. 44, at 524.

[73] *Ibid.*, at 529 (citing the ECJ decision at para. 47).

[74] *Ibid.*, at 531.

[75] Case C-434/16, *Peter Nowak v. Data Prot. Comm'r*, 2017 E.C.R. I-994, para. 34.

submitted by a candidate on an accountancy exam and the comments of the examiner evaluating those answers were personal data of the candidate because they 'constitute information relating to that candidate'.[76] The ECJ further stated that giving:

> a candidate a right of access to those answers and to those comments … serves the purpose of … guaranteeing … that candidate's right to privacy with regard to processing the data relating to him … irrespective of whether the candidate does or does not have such a right of access under the national legislation applicable to the examination procedure.[77]

This 2017 decision may indicate a willingness of the ECJ to be a bit more expansive in its view of what constitutes personal data.

In discussing the ideological and historical bases for a right to reasonable inferences, Wachter and Mittelstadt state that data protection is merely one segment of the broader right to privacy, a right which 'addresses personal and family life, economic relations, and more broadly an individual's ability to freely express her personality without fear of ramifications'.[78] They question whether current EU data protection law is sufficient to protect against the 'novel risks of automated decision-making and profiling' in the age of big data.[79] They state that '[i]ronically, inferences receive the least protection of all the types of data addressed in data protection law, and yet now pose perhaps the greatest risks in terms of privacy and discrimination'.[80]

It remains to be seen whether or not the ECJ will consider inferences to be personal data and permit access and rectification. It may determine that inferences drawn on personal data are more like legal analysis, and, therefore, not personal data, or more like comments about exam answers, and therefore, personal data. It would seem that current trends in data protection law would favour the latter. What may be more problematic, however, also discussed by Wachter and Mittelstadt, is whether an inference made by an algorithm or process that is claimed to be a trade secret may prohibit a person from accessing or rectifying his or her data.[81] This may end up being a greater obstacle than the mere absence of specific provisions for inferences drawn under the GDPR or subsequent interpretations by the ECJ regarding the scope of the definition of personal data. As Wachter and Mittelstadt conclude, 'it is safe to assume that derived and inferred data will be covered by the Trade Secret Directive'.[82]

3.2 *California Consumer Privacy Act*

In 2018 California unanimously passed the CCPA, the most ambitious and comprehensive piece of privacy legislation in the history of the United States.[83] It was passed by the legislature in rather short order because of the threat of an even more rigorous model that would have appeared later that year as a ballot initiative in California.[84] While the bill was passed in June 2018, the new law did not take effect until 1 January 2020.

In the months leading up to 1 January 2020, several important things happened. First, in September, the group that had speared the ballot initiative in 2018 announced that it had filed with the state to place another initiative on the November 2020 ballot.[85] It would ask the residents of California to approve the California Privacy Rights and Enforcement Act of 2020, a collection of amendments that would dramatically strengthen the CCPA.[86] Second, in October, the Attorney General of California issued the proposed regulations required by the CCPA. If anything, these regulations further strengthened the provisions of the law.[87] And finally, in November, Microsoft, in a significant move, announced that it would extend the core protections of the CCPA to all customers in the U.S:

Notes

[76] *Ibid.*, para. 42.

[77] *Ibid.*, para. 56.

[78] Wachter & Mittelstadt, *supra* n. 44, at 573.

[79] *Ibid.*, at 574.

[80] *Ibid.*, at 575.

[81] *Ibid.*, at 606–610.

[82] *Ibid.*, at 609. *See* Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the Protection of Undisclosed Know-How and Business Information (Trade Secrets) Against Their Unlawful Acquisition, Use and Disclosure, 2016 O.J. (L 157) 1.

[83] A.B. 375, 2017–18 St. Assemb., Reg. Sess. (Cal. 2018) (hereinafter 'CCPA'). Interestingly, the law was passed unanimously under threat of a very popular and more restrictive law being passed by way of ballot initiative later in the year.

[84] *See* Jordan M. Blanke, *Top Ten Reasons to Be Optimistic About Privacy Law*, 55 Idaho L. Rev. 281, 306–307 (2019); Daisuke Wakabayashi, *California Passes Sweeping Law to Protect Online Privacy*, N.Y. Times (28 June 2018), https://www.nytimes.com/2018/06/28/technology/california-online-privacy-law.html (accessed 4 May 2020); Issie Lapowsky, *California Unanimously Passes Historic Privacy Bill*, WIRED (28 June 2018), https://www.wired.com/story/california-unanimously-passes-historic-privacy-bill/ (accessed 4 May 2020).

[85] A Letter from Alastair Mactaggart, Founder & Chair of Californians for Consumer Privacy (25 Sept. 2019), https://www.caprivacy.org/ (accessed 4 May 2020).

[86] *See* Jeremy Greenberg, *CCPA 2.0? A New California Ballot Initiative Is Introduced*, Future of Privacy Forum (26 Sept. 2019), https://fpf.org/2019/09/26/ccpa-2-0-a-new-california-ballot-initiative-is-introduced/ (accessed 4 May 2020), The full text of the ballot initiative is https://www.caprivacy.org/CPREA2020 (accessed 4 May 2020).

[87] David Kessler, Jeewon Kim Serrato, Susan Ross, Anna Rudawski & Max Kellogg, *Mic Drop: California AG Releases Long-Awaited CCPA Rulemaking* (11 Oct. 2019), https://www.dataprotectionreport.com/2019/10/mic-drop-california-ag-releases-long-awaited-ccpa-rulemaking/ (accessed 4 May 2020).

> We are strong supporters of California's new law and the expansion of privacy protections in the United States that it represents. Our approach to privacy starts with the belief that privacy is a fundamental human right and includes our commitment to provide robust protection for every individual. This is why, in 2018, we were the first company to voluntarily extend the core data privacy rights included in the European Union's General Data Protection Regulation (GDPR) to customers around the world, not just to those in the EU who are covered by the regulation. Similarly, we will extend CCPA's core rights for people to control their data to all our customers in the U.S. …
>
> In addition to guaranteeing the rights of individuals to control their personal information, we believe privacy laws should be further strengthened by placing more robust accountability requirements on companies. This includes making companies minimize the data they collect about people, specify the purposes for which they are collecting and using people's data, *and making them more responsible for analyzing and improving data systems to ensure that they use personal data appropriately.* Indeed, we are calling upon policymakers in other states and in Congress to build upon the progress made by California and go further by incorporating robust requirements that will make companies more responsible for the data they collect and use, and other key rights from GDPR.[88]

When the GDPR became effective in 2018, many companies decided that it was easier to have one, rather than multiple, data privacy policies, and non-EU citizens around the world benefitted from those GDPR-based policies. Similarly now, many non-California residents are optimistic that this 'peer pressure' will provide them with stronger protection for their personal information.[89] Microsoft has taken a major step in that direction.

The CCPA provides many rights that are commonplace to EU residents, but very novel among US residents.[90] Among other rights, a Californian consumer[91]:

(1) can request from any business that collects personal information about the consumer to disclose to the consumer the categories and specific pieces of information collected[92];
(2) can expect that a business that collects personal information inform the consumer of the categories of information collected and the purposes for which that information is used[93];
(3) can expect that a business shall not collect additional categories of personal information without providing the consumer appropriate notice[94];
(4) can request a business to delete any personal information about the consumer that the business collected from the consumer[95];
(5) can request a business that sells the consumer's information to disclose the categories of personal information collected about the consumer, the categories of personal information that the business sold about the consumer, and the categories of personal information that the business disclosed about the consumer for a business purpose[96]; and
(6) can direct a business that sells personal information about the consumer not to sell that information.[97]

At the heart of the CCPA is its extremely broad and comprehensive definition of 'personal information':

> 'Personal information' means information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household. Personal information includes, but is not limited to, the following:
>
> (A) Identifiers such as a real name, alias, postal address, unique personal identifier, online identifier Internet

Notes

[88] Julie Brill, Corporate Vice President for Global Privacy and Regulatory Affairs and Chief Privacy Officer, Microsoft Corp., *Microsoft Will Honor California's New Privacy Rights Throughout the United States* (11 Nov. 2019) (emphasis added), https://blogs.microsoft.com/on-the-issues/2019/11/11/microsoft-california-privacy-rights/ (accessed 4 May 2020), *See* Makena Kelly, *Microsoft Vows to 'Honor' California's Sweeping Privacy Law Across Entire US*, The Verge (11 Nov. 2019), https://www.theverge.com/2019/11/11/20960113/microsoft-ccpa-privacy-law-california-congress-regulation (accessed 4 May 2020).

[89] Blanke, *supra* n. 84, at 303–308. Not surprisingly, after the passage of the CCPA, a number of other US states introduced legislation based upon, or variants of, the CCPA. Eleven states introduced legislation in 2018–2019 pertaining to the protection of personal data; eight of them were at least loosely modelled after the CCPA.

[90] For excellent discussions of the similarities and differences between the GDPR and the CCPA, *see* Anupam Chander, Margot E. Kaminski & William McGeveran, *Catalyzing Privacy Law* (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3433922 (accessed 4 May 2020); Woodrow Hartzog & Neil Richards, *Privacy's Constitutional Moment*, 61 B.C.L. Rev. (forthcoming 2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3441502. *See also* James Q. Whitman, *The Two Western Cultures of Privacy: Dignity Versus Liberty*, 113 Yale L.J. 1151 (2004); Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?*, 3 Int'l Data Privacy L. 74 (2013); Tal Z. Zarsky, *Incompatible: The GDPR in the Age of Big Data*, 47 Seton Hall L. Rev. 995 (2017).

[91] CCPA, at § 1798.140 (g).

[92] CCPA, at § 1798.100 (a).

[93] CCPA, at § 1798.100 (b).

[94] *Ibid.*

[95] CCPA, at § 1798.105 (a).

[96] CCPA, at § 1798.115 (a).

[97] CCPA, at § 1798.120 (a).

> Protocol address, email address, account name, social security number, driver's license number, passport number, or other similar identifiers. ...
> (D) Commercial information, including records of personal property, products or services purchased, obtained, or considered, or other purchasing or consuming histories or tendencies.
> (E) Biometric information.
> (F) Internet or other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding a consumer's interaction with an Internet Web site, application, or advertisement.
> (G) Geolocation data.
> (H) Audio, electronic, visual, thermal, olfactory, or similar information.
> (I) Professional or employment-related information.
> (J) Education information … .
> (K) *Inferences drawn* from any of the information identified in this subdivision to create a *profile* about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes.[98]

In contrast to the GDPR, inferences drawn are clearly defined to be personal information under the CCPA. There is also a definition for 'infer' and 'inference', which means 'the derivation of information, data, assumptions, or conclusions from facts, evidence, or another source of information or data'.[99]

There is a definition for a 'unique identifier' or 'unique personal identifier', which means:

> a persistent identifier that can be used to recognize a consumer, a family, or a device that is linked to a consumer or family, over time and across different services, including, but not limited to, a device identifier; an Internet Protocol address; cookies, beacons, pixel tags, mobile ad identifiers, or similar technology; customer number, unique pseudonym, or user alias; telephone numbers, or other forms of persistent or probabilistic identifiers that can be used to identify a particular consumer or device.[100]

'Unique personal identifier' is one of those items specifically included within the broad definition of 'personal identifier'[101] and also specifically referenced in the relatively short definition of a 'consumer', which means 'a natural person who is a California resident … however identified, including by any *unique identifier*'.[102]

This definition is particularly important since it appears to be directed at the practice, apparently frequently used today by data brokers and others, of pretending to de-identify or anonymize data by removing familiar identifiers like name, email address, or phone number, but replacing them with other identifiers – or pseudonyms – that can be used to re-identify the person and link to other accounts using that information.[103] As Wolfie Chritl explained in *Corporate Surveillance in Everyday Life: How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions:*

> In theory, hashing is a one-way operation and cannot be reversed. However, in most cases it is misleading to consider this data as 'de-identified' or even 'anonymized' for several reasons. First, these hashed identifiers still persistently refer to unique individuals and therefore cannot be considered as anonymized, but rather should be understood as **pseudonyms.** Second, most companies use identical and deterministic methods to create – or calculate – these unique codes; therefore, they can match and link profiles as soon as one of these pseudonymous identifiers appears within the digital data ecosystem.[104]
>
> Similarly, some large data companies such as Acxiom, Experian, and Oracle have introduced their own **proprietary identifiers** for people, which are used to link their extensive consumer profile information with data managed by other companies, and then to link it with the advertising data ecosystems around the globe. Often, these companies use two different identifiers, one for data that they see as 'personally-identifiable information', such as names, and one that is used for other digital profile data. However, both identifiers are linkable.[105]

While there is not a specific definition for 'profile' or 'profiling' in the CCPA, the term is used twice. The first time as part of the 'inferences drawn' definition, where it is prohibited to use inferences drawn from any type of personal information 'to create a profile about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes,

Notes

[98] CCPA, at § 1798.140 (o) (1) (emphasis added).

[99] CCPA, at § 1798.140 (m).

[100] CCPA, at § 1798.140 (x).

[101] CCPA, at § 1798.140 (o) (1) (A).

[102] CCPA, at § 1798.140 (g) (emphasis added).

[103] *See supra* text accompanying nn. 58–63 for a discussion of the Art. 29 Working Party opinion about anonymization and pseudonymization.

[104] Christl, *supra* n. 18, at 69.

[105] *Ibid.*, at 68.

intelligence, abilities, and aptitudes'.[106] The second time as part of a definition for permitted 'business purposes' as long as that business purpose is a '[s]hort-term, transient use, provided that the personal information is not disclosed to another third party and is *not used to build a profile* about a consumer or otherwise alter an individual consumer's experience outside the current interaction'.[107]

It is important to acknowledge the unprecedented position of power bestowed upon the office of Attorney General of the State of California. In effect, that position is now the most powerful privacy regulator in the United States. By virtue of the authority to promulgate regulations and to enforce the provisions of the CCPA, the Attorney General will likely shape a good bit of data privacy policy and protection in the United States for the near future. It will be interesting to see how aggressively – and broadly – the terms of the CCPA are enforced.

Clearly, one of the issues addressed by the CCPA is protection from the ever-increasing use of data analytics/predictive analytics/inferential analytics to create enormous – and privacy-invasive – profiles that are anything but anonymous or de-identified. In this regard, California had the benefit of the experiences of the EU under the Data Protection Directive and the GDPR in crafting its legislation.

3.3 *Comparing Protection for Inferences Under the GDPR and the CCPA*

The GDPR does not have specific language about inferences and inferences drawn. As Wachter and Mittelstadt discussed, this may be problematic. In order for data to be protected under the GDPR, it must be considered to be personal data. While the old Article 29 Working Party carefully distinguished among provided data, observed data, derived data, and inferred data, it is really only the last category that is of concern. All of the others should clearly be considered as personal data, even derived data, because that kind of information is usually easily verifiable, particularly if it involves relationships among data where one item functionally determines another.

Inferred data, however, is often created by applying some sort of data analytics, whether we call it predictive analytics or inferential analytics, and it is usually applied at the back end or the output side of the process. As Wachter and Mittelstadt contend, while the GDPR has great protection for the input side, it is lacking on the output side.[108] It is not clear how inferred data will be treated under the GDPR. Absent any modifications to the language of the GDPR itself, much will depend upon whether the ECJ extends the definition of personal data to include inferred data. There is also concern about whether these output side analytic processes may be protected from examination by data subjects under the new EU Trade Secrets Directive.[109]

Even though the GDPR contains a definition for profiling, which appears to address some of the issues pertaining to inferences, it probably has little practical application. Article 22(3) of the GDPR provides that the 'data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'.[110] There are two problems. First, data controllers may contend that the decision is not based *solely* on automated processing,[111] and second, as Wachter and Mittelstadt discussed, this right to contest under Article 22(3):

> appears to be a mere procedural right to reverse decisions or impactful profiling made using inaccurate or incomplete input data. It is unlikely to compel data controllers to revise automated decisions based on inferences unless sector-specific decision-making standards or other provisions in data protection law have been infringed.[112]

Because California had the benefit of a couple of more years to see where technology was headed, it was able to include much tighter definitions.[113] The CCPA is extremely well drafted. Its definitions are superb. The

Notes

[106] CCPA, at § 1798.140 (o) (1) (K).

[107] CCPA, at § 1798.140 (d) (4) (emphasis added). Probably not surprisingly, there is even more attention paid to profiling in the proposed 2020 ballot initiative, in which the terms 'profile' and 'profiling' appear twelve times. There is a specific definition for 'profiling' and new provisions regarding sensitive data profiling and cross-context behavioural advertising. https://www.caprivacy.org/CPREA2020 (accessed 4 May 2020).

[108] Wachter & Mittelstadt, *supra* n. 44, at 513–514.

[109] *Ibid.*, at 606–610.

[110] GDPR, at Art. 22(3).

[111] This is an issue that is addressed in the 2020 California ballot initiative, because of its potential use as a loophole. The ballot initiative contains a proposed amendment requesting the Attorney General to adopt regulations 'to further define "profiling", to reflect that it is intended to cover the process of making a decision without any human involvement, but to prohibit any deliberate insertion of, or reliance on, a minor step involving a natural person, as a reason for not identifying the process as automated. Additionally, the creation of a process by a natural person or persons, or the entry of data by a natural person, shall not render the process exempt from the definition of "profiling"'. https://www.caprivacy.org/CPREA2020 (accessed 4 May 2020).

[112] Wachter & Mittelstadt, *supra* n. 44, at 571.

[113] Even in the two years since the CCPA was drafted, there are numerous enhancements to the language that is contained in the proposed 2020 ballot initiative. This will always be the case – which makes regulating new technology so difficult. By the time the law has tried to catch up with technology, technology has already moved several steps forward.

definition for personal information is very broad and includes the definition for inferences drawn, so unlike under the GDPR, there is no question that inferences derived from other pieces of personal information are considered themselves to be personal information, and thus subject to all the provisions in the CCPA that apply to personal information.

The definition for unique personal identifier is also extremely important as it is directed at preventing the charade involving supposedly anonymous or de-identified data that is quite easily re-identified by using persistent pseudonyms or proprietary identifiers. So while not inferences themselves, the unique personal identifiers could otherwise be used to facilitate the linking of profiles or of supposedly anonymized or de-identified data.

Certainly, if one looks at the purposes behind and goals for both the GDPR and the CCPA, one would conclude that both should provide protection for inferences drawn from personal data or personal information. Hopefully, this will be the case, but because legislation must be precise and exacting, there are some challenges under the GDPR that seem not to exist under the CCPA. Obviously, only time will tell how effective the CCPA and its enforcement by the Attorney General of California will be, and whether the GDPR will close the possible gap in coverage, regarding inferences drawn.