What Is ‘Hashing,’ and Does It Help Avoid the Obligations Imposed by The New Privacy Regulations?

Article By

David A. Zetoony
Greenberg Traurig, LLP
Data Privacy Dish

Related Practices & Jurisdictions

- Administrative & Regulatory
- Communications, Media & Internet
- Corporate & Business Organizations
- Election Law / Legislative News
- California

Friday, April 2, 2021

Maybe.

“Hashing” refers to the process of using an algorithm to transform data of any size into a unique fixed-sized output (e.g., combination of numbers and letters). To put it in layman’s terms, some piece of information (e.g., a name) is run through an equation that creates a unique string of characters. Anytime the exact same name is run through the equation, the same unique string of characters is created. If a different name (or even the same name spelled differently) is run through the equation, an entirely different string of characters will emerge.

While the output of a hash cannot be immediately reversed to “decode” the information, if the range of input values that was submitted into the algorithm is known, they could be replayed through the algorithm until there is a matching
output. The matching output would then indicate what the initial input had been. For instance, if a Social Security Number was hashed, the number might be reverse engineered by hashing all possible Social Security Numbers and comparing the resulting values. When a match to the initial hash is found, the initial Social Security Number that created the hash string would be identified. The net result is that while hash functions are designed to mask personal data, they can be subject to brute force attacks.

Whether a hash value in and of itself is considered “personal information” depends upon the particular law or regulation at issue.

Information is not considered “personal information” under the CCPA if it has been “deidentified.”[1] Deidentification means that the data “cannot reasonably identify, relate to, describe, be capable of being associated with, or be linked, directly or indirectly, to a particular consumer.”[2] An argument could be made that, once hashed, information cannot reasonably be associated with an individual. That argument is strengthened under the CCPA if a business takes the following four steps to help ensure that the hashed data will not be re-identified:[3]

1. **Implement technical safeguards that prohibit reidentification.** Technical safeguards may include the process or techniques by which data has been deidentified. For example, this might include the strength of the hashing algorithm used or combining hashing with other techniques to further obfuscate information (e.g., salting).[4]

2. **Implement business processes that specifically prohibit reidentification.** This might include an internal policy or procedure that prevents employees or vendors from attempting to reidentify data or reverse hashed values.

3. **Implement business processes to prevent inadvertent release of deidentified information.** This might include a policy against disclosing hashed values to the public.

4. **Make no attempt to reidentify the information.** As a functional matter, this entails taking steps to prohibit reidentification by the business’s employees.

Pursuant to the CPRA slightly different deidentification factors will apply beginning in 2023.[5]

In comparison, in the context of the European GDPR, the Article 29 Working Party[6] considered hashing to be a technique for pseudonymization that “reduces the linkability of a dataset with the original identity of a data subject” and thus “is a useful security measure,” but is “not a method of anonymisation.”[7] In other words, from the perspective of the Article 29 Working Party, while hashing might be a useful security technique, it is not sufficient to convert personal data into deidentified data.

---


[4] Salting refers to the insertion of characters into data before it is hashed to make brute force reidentification more difficult.


[6] The Article 29 Working Party was the predecessor to the European Data Protection Board.


©2021 Greenberg Traurig, LLP. All rights reserved.

National Law Review, Volume XI, Number 92