Blogs
June 24, 2019
Pseudonymized Personal Information on Blockchain Not Sufficient Under CCPA

The California Consumer Privacy Act (CCPA) imposes new transparency and disclosure obligations on businesses' use, sale, and disclosure of consumer information. Businesses will need to honor requests from consumers to access their personal information, delete their personal information, and opt out of the sale of their personal information. "Personal information" is more broadly described in the CCPA than in any prior statute: that is, "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." In reaction, many blockchain-based businesses are relying on data obfuscation techniques in an attempt to avoid liability—including hashing personal information and other forms of encryption. Hashing is a form of encryption that converts raw data into a seemingly random string of values. Hashing is generally a "one-way" process; that is, by using the encryption key, an individual may reduce the same data to the same hash but will not be able to use the key to conjure the raw data from which it was reduced. By utilizing these techniques, businesses hope to render any personal information that may appear on-chain sufficiently deidentified so as to fall out of the scope of the CCPA's rigorous obligations. However, encryption and, by extension, hashing would not result in anonymization but rather pseudonymization. Anonymization is the process of making personal data completely and irreversibly non-personal. Pseudonymization, by comparison, is the process of making personal data non-personal, but nevertheless capable of reidentification when combined with other data or by exercising some process. Because many blockchains are public and transparent there are ample opportunities to associate the data across other data points, and to analyze the metadata around the ledger entries such that pseudonymous data risks reidentification. If an individual knows the personal information at issue and the hashing algorithm, they can simply perform the hashing operation on the off-chain data (assuming no further obfuscation techniques were applied to the input data), and then look for the corresponding hash on the blockchain to identify where that data was entered into the ledger. (Of course, it is possible that there is more than one raw data set that, if subjected to the encryption methodology, would result in the same hash. However, so long as the hash string is sufficiently large, the likelihood of that happening with any particular set of data is infinitesimally small.) More, encryption in any form is still vulnerable to hacking as technology grows more sophisticated, and therefore any use of these techniques is subject to possible regulatory scrutiny if not the CCPA's private right of action. Therefore, because there is always a risk of reidentification, as a best practice, blockchain companies should avoid writing pseudonymized or plain text personal data to the blockchain whenever it is practical to do so. To the extent that the use of pseudonymized cannot be avoided, blockchain companies should consider a risk-based approach to further reduce the risk of reidentification. The very public nature of blockchain transaction data increases the likelihood of linkability and necessitates the use of noise addition techniques like salting (i.e. adding a bit of random data to the input data to guarantee a unique output) before applying the hash algorithm to any personal data stored on the blockchain. Such techniques may help avoid the concern identified above regarding the re-hashing of the input data and identifying references on chain. Of course, the use of these techniques and others in a risk-based approach will depend on the particular facts and circumstances related to the data types employed, and it is ultimately the responsibility of the blockchain company and its users to decide on an appropriate approach.

Authors

Explore more in

Privacy & Security