Data Regulation

COUNTERMEASURES HARNESS THE POWER OF A.I.

BY RYAN MCCRACKEN

cgn®

Good thinking. Globally.

Our world is more connected than ever. A user of technology products produces data with every interaction, transaction, and movement. Our digitized world affords endless opportunities to discover new insights, correlations, and discoveries through data science and artificial intelligence (A.I.). However, along with those benefits also come some serious risks. The power and control of data custodians is being scrutinized and whether limits should be placed on how data is used.

In response to these concerns, governments have begun to regulate data collection, use, and privacy. It is completely reasonable to set common standards related to data control and use, since the consequences of data being in the wrong hands is potentially enormous, like the consequence to an unwitting consumer from identity theft. However, taking regulations too far can stifle innovation and our ability to harness the power of technologies. Developers of technologies like artificial intelligence can utilize specific mathematical techniques in their algorithms to alleviate these security and privacy concerns.

A.I. is grounded in mathematics, and, simply described, is a string of algorithms chained seamlessly together. Differential privacy and cryptographic techniques can be utilized within the chain to provide information anonymity and a secure method of transfer. Cryptography is the study of techniques used for secure communication in the presence of adversarial actors. It uses a set of techniques to encrypt communication using complex algorithms to convert the information into a unique code. Differential privacy is like cryptography in that in uses complex mathematical techniques to mask information so that it may be transferred, stored, or used in a secure manner.

Differential Privacy uses a set of algorithmic rules that limit the disclosure of specific information within a database, or more complex techniques creating mathematical ‘noise’ or randomization within the algorithm. The addition of ‘noise’ or randomization to the output of an algorithm makes it difficult to reverse engineer and determine specific information from a dataset. As far as cryptography, two examples that could be utilized are asymmetric cryptography and blockchain. Asymmetric cryptography uses a public key to allow anyone to submit data. After the data is submitted and converted by the algorithm, the only way to decrypt the code is to use of a private key that is held securely by the administrator. This is an efficient method of security because the effort and investment in security is focused solely on the private key and not the entire dataset. Further, both differential privacy and asymmetric cryptography require simply adding another algorithm within an A.I. model.

Blockchain would be another effective method to ensure security and privacy, but its complexity would not make it a good fit for all applications. It is also a much more complex method and would be costly to maintain. Like asymmetric cryptography, it also converts information into cryptographic code. This code in blockchain is called hashes. The source of the complexity is the method in which blockchain stores information. It uses a structure called a merkle tree, which essentially creates a ‘hash of hashes’ or one code to represent an entire dataset. Further, it requires a network of computers to

validate the computations of the other users within the network, making this technique costly over time from the requisite computing power necessary to perform the validation. In sum, this is a much more complex solution to the same problem. Depending on the use and sensitivity of the information being used, blockchain might be a better solution because with its complexity comes additional security.

In sum, the scope of data regulations need not be cumbersome in order to achieve the desired level of security and privacy. The use of mathematical techniques like differential privacy, asymmetric cryptography, and blockchain will resolve many of the concerns these regulations attempt to address.

REFERENCES

Dwork, C. & Roth, A. (2014). *The Algorithmic Foundations of Differential Privacy*. Foundations and Trends. 9(Nos. 3-4), 211-407.

Official Journal of the European Union. (2016). *General Data Protection Regulation*. 59 (L119), 1-149.

Wallace, N. & Castro, D. (2018). *The Impact of the EU's New Data Protection Regulation on AI.* Center for Data Innovation. https://www.datainnovation.org/2018/03/the-impact-of-the-eus-new-data-protection-regulation-on-ai/, 1-37.

Illinois Personal Information Protection Act, 815 ILCS 530, et seq.

de Kruijff, J., & Weigand, H. (2017). *Understanding the blockchain using enterprise ontology*. In E. Dubois, & K. Pohl (Eds.), Proceedings of the 29th International Conference on Advanced Information Systems Engineering (CAiSE 2017) (Vol. 10253, pp. 29-43).

AL.Jeeva, Dr.V.Palanisamy, K.Kanagaram (2012). *Comparative Analysis Of Performance Efficiency And Security Measures Of Some Encryption Algorithms.* International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 3, May-Jun 2012, pp.3033-3037