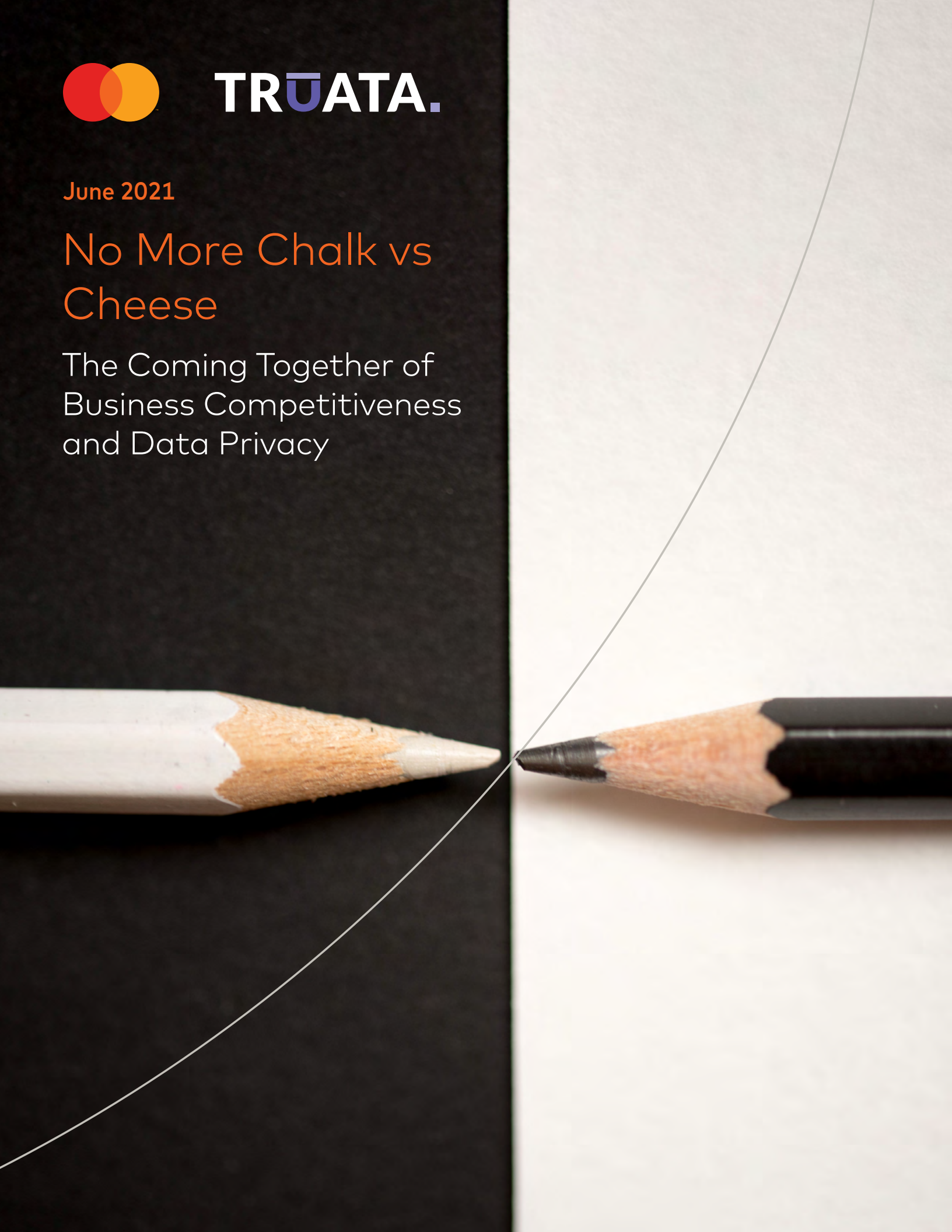
June 2021

No More Chalk vs Cheese

The Coming Together of Business Competitiveness and Data Privacy



Charles de Gaulle reportedly once jokingly asked: "How can you be expected to govern a country that has 246 kinds of cheese?" Half a century later, business leaders are more seriously asking how they can be expected to govern a business in the face of overwhelming data privacy requirements. As of February 2020, 128 of the world's 190-plus countries had a form of data privacy legislation in place.[1]

Many businesses no longer exclusively operate in one jurisdiction. A simple web presence can take a business across territorial limits. As such, business competitiveness may on the surface appear incompatible with data privacy. To borrow a British idiom, they're like chalk and cheese. But de Gaulle was proud of his country's cheese output, and the opportunities found in data privacy should be similarly celebrated.

Data privacy regulations are aimed to be pro-consumer; they're designed to protect individual consumer privacy rights by giving them control of their data. Still, pro-consumer doesn't have to mean anti-business. On the contrary, meeting consumer expectations usually reflects good business.[2] But when legally enforced, good business can seem like business as usual and hardly an opportunity for competitive differentiation. At worst, data privacy can end up being perceived as a regulatory burden.[3]

Taking data from potential liability to enhanced asset can be achieved through independent data anonymisation.[4] This paper explores that from the following perspectives:

- The challenge of data anonymisation
- An independent data anonymisation solution
- The added value of independent data anonymisation

"Business competitiveness may on the surface appear incompatible with data privacy. To borrow a British idiom, they're like chalk and cheese.

[1] "Data Protection and Privacy Legislation Worldwide." The United Nations Conference on Trade and Development, February 2020.

[2] "'Companies that Use Data in an Ethical Way will Win the Trust of Customers.'" Trūata, February 5, 2020.

[3] "Knowing More by Knowing Less: The True Value of Anonymous Data." Mastercard, 2020.

[4] "How Privacy-Driven Innovation Can Eliminate Data Waste." Trūata, March 2021.

The Challenge of Data Anonymisation

Personalisation paradox, consent conundrum, digital dichotomy. The terms are as alliterative as they are contradictory. Consumers want to benefit from businesses putting their personal data to work, but they don't want businesses to have insights into them personally.[5]

The situation is as trying for businesses as it is for consumers. On the one hand, businesses worldwide are left with unusably small sets of personal data with strict expiration dates on permitted usage. On the other, they have large sets of personal data that they aren't permitted to use.

Even the notion of "permission" is difficult to pin down. Standards vary between regions, particularly around the required regulatory threshold for personal data to be considered transformed into non-personal data. What's valid in one jurisdiction may not suffice in another.

In addition, the geographic boundaries between data privacy jurisdictions often border on irrelevant more than they border on each other. For example, *the Brussels effect*, coined after *the California effect* on regulation across the US, means that the impact of the General Data Protection Regulation (GDPR) extends well beyond its EU borders. The dedicated name for the effect reflects the economic clout of the region, but the situation is no different with the numerous other privacy regulations around the world.

Confusion over what is permissible means the only recourse for many businesses is to try to meet the most stringent regulations in existence anywhere in the world as a common denominator. Data anonymisation is supposed to help. The theory is simple: render the personal data unidentifiable by transforming it into non-personal data that falls outside of any privacy regulations around the world. Things are more challenging in reality. Many businesses who try to anonymise data internally fail to meet the required high standard.[6]

"Confusion over what is permissible means the only recourse for many businesses is to try to meet the most stringent regulations in existence anywhere in the world as a common denominator.

[5] "Global Consumer State of Mind Report." Trūata, 2021 (survey conducted across France, South Korea, Brazil, the UK and the US).

[6] "10 Misunderstandings Related to Anonymisation." Agencia Española de Protección de Datos & European Data Protection Supervisor, April 2021.

An Independent Anonymisation Solution

There exists one fundamental among a host of approaches to data anonymisation: independence. No matter how strong the combination of anonymisation solutions, retention of the source data and the modified data is inherently risky.

Personal data may be deidentified by replacing it with artificial identifiers knowns as pseudonyms. But reidentification can be achieved through *quasi-identifiers*, which are combinations of attributes within records that can be used to single out records that may identify an individual, link two or more records, or infer the value of an attribute from the value of others.[7]

Effective anonymisation solutions that randomise and generalise datasets can in theory remove all quasi-identifiers. But, even in a hypothetically infallible case, they can't eliminate the risk of reidentification when the possibility of partial access to elements of an original dataset remains.[8]

Fines aren't businesses' only concerns; a perception that they don't view their customers as valued individuals can result in longstanding reputational damage. Bring in an independent third party to conduct the anonymisation process, and the risk of reidentification is reduced to an insignificant level since the source data and the modified data no longer coexist in one place. In addition, the third party assumes responsibility for the anonymisation process, which mitigates a business's non-compliance risk.

> Even the deftest use of privacy-enhancing anonymisation techniques inevitably results in some loss in value.

[7] "Opinion 05/2014 on Anonymisation Techniques." Article 29 Data Protection Working Party, European Commission, 2014.

[8] "Estimating the Success of Re-identifications in Incomplete Datasets Using Generative models." Nature Communications, July 23, 2019.

An Independent Anonymisation Solution

Yet the independent application of the most suitable anonymisation solutions to a dataset requires a delicate balance if analytical value is to be preserved. Even the deftest use of privacy-enhancing anonymisation techniques inevitably results in some loss in value. The restoration of much, if not all, of that value comes from using large anonymised datasets to build and train robust models to apply to smaller consumer datasets that don't require anonymisation. The analytic outputs, or trained model code, can then provide valuable insights to businesses (figure 1).

Still, the process superficially appears to represent little more than a circuitous, albeit necessary, route to give businesses access to insights they would have been able to obtain directly with unrestricted use of personal data. What's missing from figure 1 is a recognition of the additional value created by the independent anonymisation process.

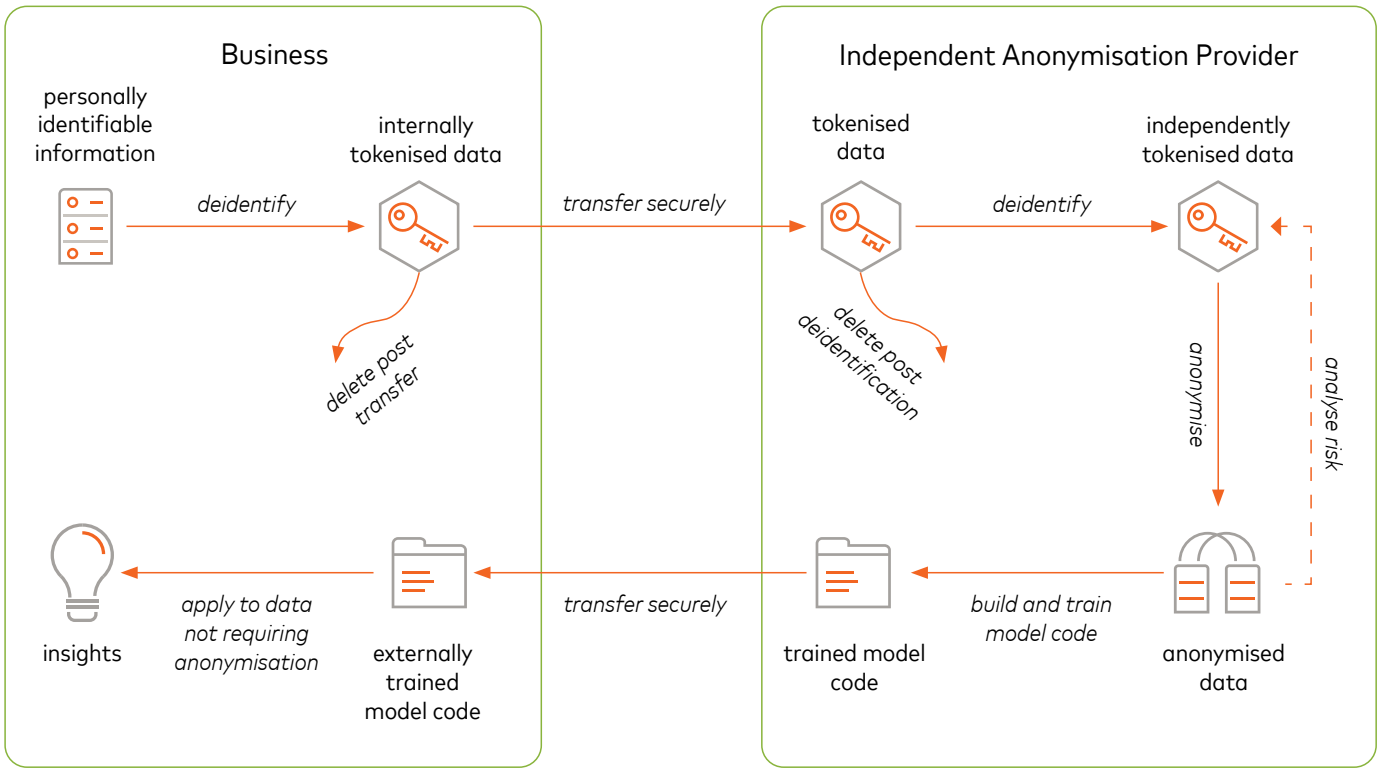


Figure 1: An Independent Anonymisation Process

The Added Value of Independent Data Anonymisation

The primary business cases for data anonymisation are to comply with stringent privacy regulations and to meet global consumer concerns. Roughly three-quarters of consumers believe they should own their *digital selves* or online identities, and similar numbers believe independent data protection regulation is more important than ever.[9]

In addition to an ethical imperative, appealing to consumer sentiment is important from a brand perspective: 76% of global consumers believe that businesses need to do more to protect their data privacy online, 69% are more likely to be loyal to a brand they trust to use data responsibly, 57% have stopped using brands that follow them online, and 60% believe they would spend more money with a brand they trust with personal data.[10]

It's no coincidence that data privacy matters more to consumers than any other technology affecting their purchase experiences other than cybersecurity, which it equals in ranking.[11] But meeting evolving consumer demand is a baseline expectation, and consumers and businesses don't appear to perceive the benefits of sharing data equally. While 60% of businesses believe consumers think the value they receive in exchange for sharing their data is worthwhile, only 44% of consumers believe that to be the case.[12] Businesses need to offer more.

The power of independent data anonymisation to meet regulations and consumer concerns comes from its ability to render personal data no longer personal. The transformed data falls outside data protection regulations and so is subject to fewer restrictions than personal data. That means non-personal data can be used for stable long-term modelling because certain rules on personal data will not apply: data retention limits, data minimisation requirements around only collecting what's necessary, and data subject rights such as a consumer's right to be forgotten. Access to current and historical data makes it easier to identify trends and seasonal patterns, which are essential in activities such as lifetime value modelling and using recommendation engines (figure 2).

> "It's the independently anonymised data that in some ways pulls up the value of the personal data that's available for use without anonymisation.

[9] "Global Consumer State of Mind Report." Trūata, 2021 (survey conducted across France, South Korea, Brazil, the UK and the US).

[10] "Global Consumer State of Mind Report." Trūata, 2021 (survey conducted across France, South Korea, Brazil, the UK and the US).

[11] "The Value of Experience: Customer Needs Top the Innovation Agenda." Harvard Business Review Analytic Services (sponsored by Mastercard), 2021.

[12] "The Great Data Exchange: What Businesses and Consumers Value in the Digital Economy." Harvard Business Review Analytic Services (sponsored by Mastercard), 2020.

The Added Value of Independent Data Anonymisation

So, contrary to what might be expected, it's the independently anonymised data with its increased value for stable long-term modelling that in some ways pulls up the value of the personal data that's available for use without anonymisation. Although antitrust or other restrictions may still apply, independent anonymisation grants businesses access to an ever-increasing mass of transformed data and the ability to conduct repeat analyses to boost the quality of insights obtained. A clear use case is artificial intelligence because of its insatiable demand for ethically sourced quality data.[13]

Independent anonymisation can also avoid a *Swiss cheese effect*, which refers to the holes in incomplete datasets that result from businesses not having access to data that requires anonymisation. The unrepresentative datasets can lead to inaccurate and biased business intelligence.

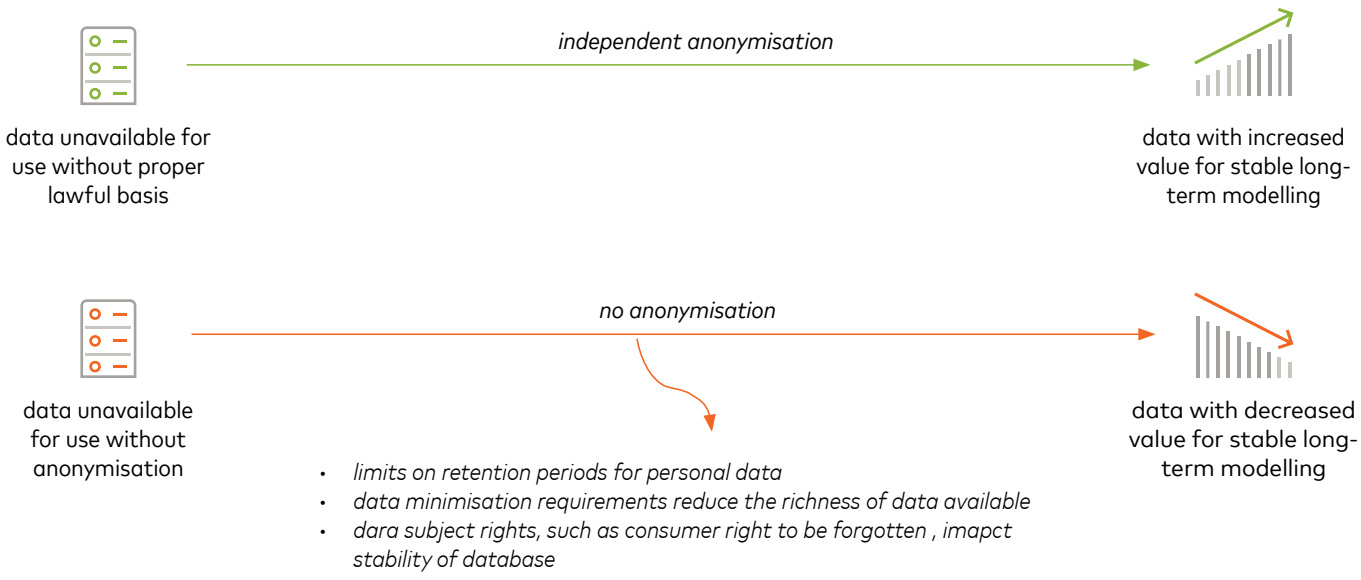


Figure 2: The Value of Data Anonymisation for Stable Long-Term Modelling

[13] "The Secret Advantage to Winning the AI Race." Trūata, November 2020.

Relishing the Regulations

Independent data anonymisation can preserve and even enhance the analytical value of data while protecting consumer privacy. It does so by embracing the spirit of data privacy regulations, which are designed to help consumers worldwide control and benefit from their own personal data while allowing businesses to compete more effectively.

De Gaulle's 246 kinds of cheese are surely a delight for connoisseurs who appreciate them. In the absence of a global data privacy standard, the same might now be said for the multitudinous data privacy laws that need no longer leave businesses with a chalky aftertaste.

Mastercard's Data Responsibility Principles exist to guide our activities as responsible data stewards and those of other like-minded organisations. The principles are foundational to Mastercard Advisors in the provision of consulting services and to Trūata—a company founded by Mastercard and IBM that meets high regulatory standards—in the provision of independent anonymisation, de-identification tools and privacy-enhanced analytics. Contact an expert to learn more:

Aoife Sexton
Chief Privacy Officer and Chief of Product Innovation, Trūata
aoife.sexton@truata.com

Adam Szonyi
Principal, Business Development, Data & Services, Mastercard
adam.szonyi@mastercard.com

More Resources

Global Consumer State of Mind Report 2021

Knowing More by Knowing Less: The True Value of Anonymous Data

The Great Data Exchange: What Businesses and Consumers Value in the Digital Economy