GDPR Data Analytics Webinar Summary

Three Key Points

1. **The GDPR Increases Options for Organizations to Process Data:** Legitimate Interest is an available legal basis to enable GDPR compliant data analytics, artificial intelligence (AI), machine learning and digital transformation.

2. **Support for Global Data-Driven Business Beyond GDPR Compliance:** Pseudonymisation, as newly defined under the GDPR, supports the separation of the information value of data from the re-identifiability of individuals as necessary to support innovation for data-driven businesses.

3. **Controlled Re-Linking of Data Increases the Value of Data Analytics:** The ability of pseudonymisation to help support re-linking of data about individuals under controlled conditions distinguishes it from anonymisation, general statistical analysis, or complete de-Identification, which are not designed to support re-linkability of data.

Gartner predicts that by 2020, more than 40% of enterprise revenue will come from digital business. Similarly, IDC forecasts that by 2020, 50% of the Global 2000 will see a majority of their business come from their ability to create digitally-enhanced products, services, and experiences. ***Yet many data-driven operations underlying these projections rely on data analytics that are increasingly subject to restrictions on lawful data use such as contained in the GDPR and similar evolving regulations.***

Global Data Processing Regulations

- EU General Data Protection Regulation (GDPR) and draft ePrivacy Regulation
- Japanese Act on Protection of Personal Information ("APPI")
- Australian Privacy Amendment (Notifiable Data Breaches) Bill
- India's Supreme Court ruling that privacy is a fundamental right
- Canada PIPEDA and data breach notification obligations
- Israel Privacy Protection Regulations (Data Security)
- U.S. State (48) and Territory data breach notification obligations
- Federal Trade Commission ("FTC") settlement with Uber

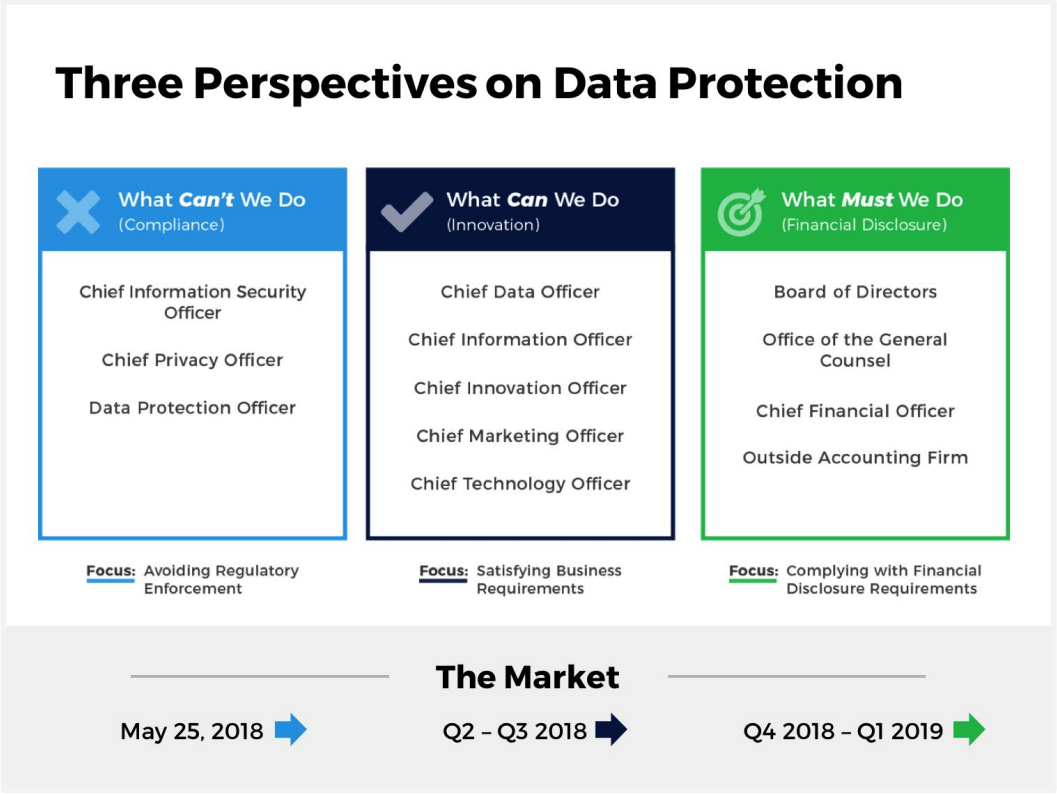
The GDPR is global in scope. If an organization processes records of any individuals in the EU to provide goods or services in the EU, the GDPR applies no matter where the organization is located. The GDPR is much more than a law pertaining to EU personal data – it is the leading wave of transformational data processing restrictions evolving around the globe.

Prior to the GDPR, the primary burden of risk for inadequate data protection in the EU was born principally by data subjects, due to limited recourse against data controllers and processors that collected and stored their data. However, this burden of risk is shifted by the GDPR's emphasis on rights of individual data subjects. As a result, a data subject's "consent" must be "freely given, specific, informed and an unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her" to serve as lawful basis for processing personal data. These GDPR requirements are impossible to satisfy with respect to iterative data analytics where successive

analysis, correlations and computations cannot be described with required specificity and unambiguity in advance at the time of consent.

To lawfully process iterative data analytics, new technical measures that help support alternate (non-consent) GDPR-compliant legal bases are required. After May 25, 2018, companies that continue to rely on broad-based consent will ***not*** comply with GDPR requirements. Failure to comply with GDPR obligations exposes parties, including co-data controller and data processor partners, to fines equal to the ***greater*** of 20 Million Euros or 4% of global gross revenues of the ultimate parent company, plus additional significant obligations, liability, and exposure.

The September 20, 2017 webinar – **Don't Lose Access to Data Analytics Under the GDPR** – featured presentations by Gwendal Le Grand – Director of Technology and Innovation at the French CNIL, Jules Polonetsky – Chief Executive Officer at the Future of Privacy Forum (FPF), and Gary LaFever - Chief Executive Officer at Anonos. The panelists discussed ways that global organizations can reconcile the growing importance of data analytics with increasingly complex and multi-jurisdictional restrictions on lawful data use.



Assessment of actions necessary to comply with new GDPR requirements is the initial focus of most organizations. Organizations later begin to appreciate that actions taken to comply with the GDPR are also relevant for financial disclosure

purposes since potential lawsuits, regulatory fines, and lost access to data, could harm operating results and therefore require disclosure. It is imperative that actions taken to avoid regulatory enforcement and comply with financial disclosure requirements do not contravene innovation requirements necessary for organizations to achieve business requirements.

The good news is that pseudonymisation, a new technical and organizational measure introduced under the GDPR, helps to support alternate (non-consent) legal bases necessary for iterative data analytics.

Webinar panelists discussed how GDPR requirements for new technical and organizational measures seek to both protect personal data as well as to enable privacy respectful innovation by enabling GDPR compliant analytics.

Slides for Gwendal Le Grand (Director of Technology and Innovation at the CNIL) (Three slides in total)

Gwendal Le Grand
Director of Technology
and Innovation

CNIL.

Anonymization vs. Pseudonymisation

Opinion 5/2014 on Anonymization Techniques: Singling Out, Linkability, Inference

Randomization	Generalization
Noise addition	Aggregation and k-anonymity
Permutation	L-diversity/t-proximity
Differential privacy	

Pseudonymisation is NOT Anonymization

- **Recital 26:** Principles of data protection do not apply to anonymous information vs. pseudonymised data are personal data (indirectly identifiable persons)
- **Recital 29:** Selective access to data within an organization to ensure that people only have access to data as needed for their jobs and no more. By using Pseudonymisation, only discrete elements need be made available to a person to support minimal use
- Pseudonymisation is different from traditional "access controls" that are enforced at a "person" level. In contrast to access controls that determine "who" can access data, Pseudonymisation enables selective access to control "what" each person can see and do.
- **Recital 28:** Pseudonymisation can reduce the risks to data subjects and help controllers / processors meet their obligations. It should mainly be seen as a security measure (cf. Article 25 on Data Protection by Design and by Default, Article 32 on Security) among others.

CNIL.

Legal Basis for Processing

- WP29 Opinion 6/2014 on Legitimate Interest
- GDPR Article 6
  - **Consent** is limited to what can be described with specificity and unambiguity in advance
  - **Necessary for Contract** is limited to requirements for performing contract and does not include ancillary processing.
  - **Controller Legal Obligations** are unlikely to support most Data Analytics.
  - **Data Subject Vital Interest** is unlikely to support most Data Analytics.
  - **Public Interest** is unlikely to support most Data Analytics.

- **Legitimate Interest** may be supported by Pseudonymisation, which also helps to support Data Analytics – and other secondary uses (further processing) – as compatible purposes.

CNIL.

Legitimate Interest Ground: Processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, **except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject.**

- Art 6(1)(a) - Consent
- Art 6(1)(b) – Necessary for Contract
- Art 6(1)(c) – Controller Legal Obligations
- Art 6(1)(d) – Data Subject Vital Interest
- Art 6(1)(e) – Public Interest
- Art 6(1)(f) – Legitimate Interest

See 'pizza' examples on pages 31, 32, and 33 of WP29 Opinion 6/2014 for examples of limitations.

Anonos

Benefits of Pseudonymisation

Article 4(5) definition: The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.

Article 5 – Principles Relating to Processing of Personal Data
Pseudonymisation can help to satisfy requirements under Article 5:

- **Alternate (Non-Consent) Legal Bases:** Reduce risk to data subjects in support of alternate (non-consent) legal bases for primary data uses [Art. 5(1)(a) – see also Art. 6(1)(b-e)];
- **Secondary Data Uses (Further Processing) – like Analytics:** Support secondary uses (further processing) of data – such as Data Analytics – as a compatible purpose not requiring consent [Art. 5(1)(b) – see also Art. 6(4)];
- **Selective Access:** Enforce selective access to data within an organization and when data is shared between organizations to enable data minimization by limiting access to only that data which is relevant for each authorized purpose [Art. 5(1)(c)];
- **Archiving:** Support archiving of data for public interest, scientific, or historical research and statistical purposes [Art. 5(1)(e) – see also Art. 89(1)]; and
- **Enhanced Security:** Enforce granular technical and organizational measures to help protect against unauthorized or unlawful processing, accidental loss, destruction or damage [Art. 5(1)(f) – see also Art. 25].

CNIL.

Slides for Jules Polonetsky (CEO at the Future of Privacy Forum (FPF))
(Six slides in total)

Jules Polonetsky
Chief Executive Officer

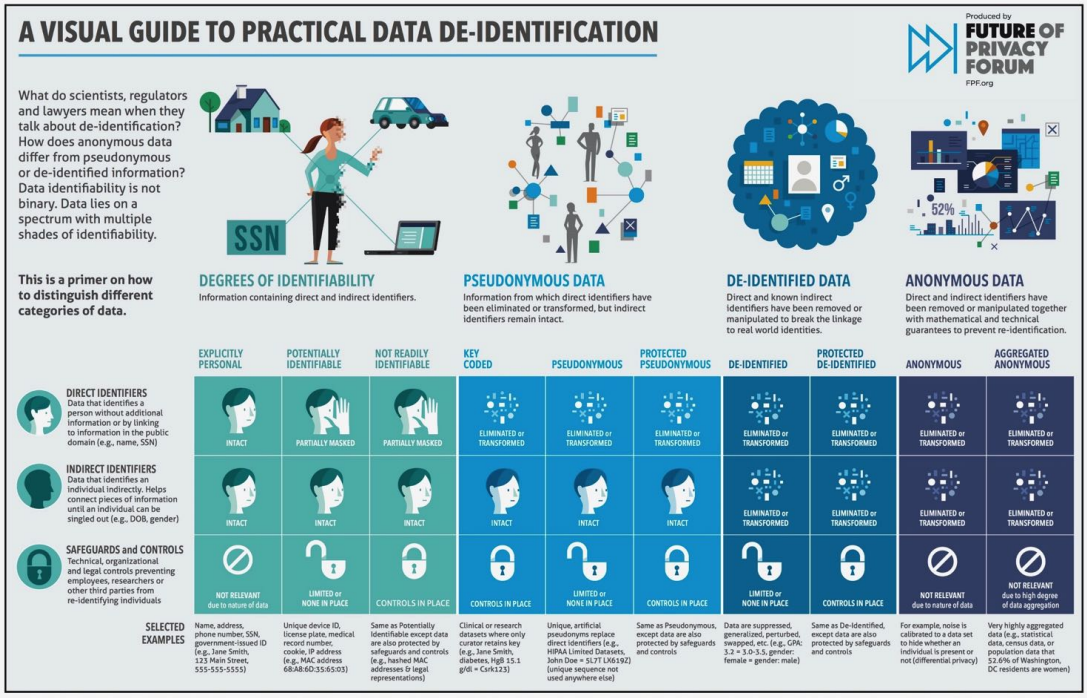
FUTURE OF PRIVACY FORUM

Anonos

Advances in Anonymization Technologies

- **Differential Privacy** - provides a quantifiable measure of the excess privacy risk any individual may incur due to their data being included in an analysis, as compared with their data not being included.
- **Synthetic Data** - mimics an original dataset, by replacing actual values from the original data with altered values that still retain many (but not necessarily all) important statistical properties of the real dataset.
- **Secure Multiparty Computation** - enables distributive computing that allows data owners to maintain their data "in their own silos" and yet compute results on the combined data.
- **Homomorphic Encryption** - provides the ability to generate aggregated reports about the comparisons between separate, fully encrypted, datasets without revealing the underlying raw data.

FUTURE OF PRIVACY FORUM



FUTURE OF PRIVACY FORUM https://fpf.org/2016/04/25/a-visual-guide-to-practical-data-de-identification/

Anonos

Pseudonymisation and the GDPR

Pseudonymisation - the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that:

1. Such additional information is kept separately, and
2. Is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person. Art. 4(5).

GDPR aims to "create incentives to apply pseudonymisation when processing personal data." Rec. 29.

Pseudonymisation is a measure that can:
Support a finding of "compatibility" for processing personal data beyond the original collection purposes. Art 6(4)(e).
Ensure respect for the principle of data minimisation when processing for scientific research. Art. 89(1).
Implement data protection by design and default. Art. 25(1).
Ensure a level of security appropriate to the risk. Art. 32(1)(a).

FUTURE OF PRIVACY FORUM

PII and De-Identification in the U.S.

Federal Trade Commission, Section 5 of the FTC Act
PII is data that can be "reasonably linked" to a particular person, computer, or device.
Three-Part Test for when data are not **"reasonably linked"**
- Reasonable steps to de-identify
- Public commitment
- Require any third parties to commit

Must achieve a "reasonable level of justified confidence"
- "Reasonable" depends on the circumstances, e.g., available methods and technologies, nature of the data, and use of the data.
- Variety of technical approaches may be reasonable, e.g., deletion or modification of data fields, addition of sufficient "noise" to data, statistical sampling, or use of aggregate or synthetic data.

HHS, HIPAA §164.514(b)
Expert Determination Method: A person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable:
(i) ...determines that the **risk is very small** that the information could be used, **alone or in combination** with other reasonably available information, **by an anticipated recipient** to identify an individual who is a subject of the information; and
(ii) Documents the methods and results of the analysis.

FUTURE OF PRIVACY FORUM

Evolving Identifiers in the U.S.

In the matter of Myspace (settled September 2012): Sharing of unique profile IDs that can be linked back to individuals via access to readily available look-up database is impermissible.

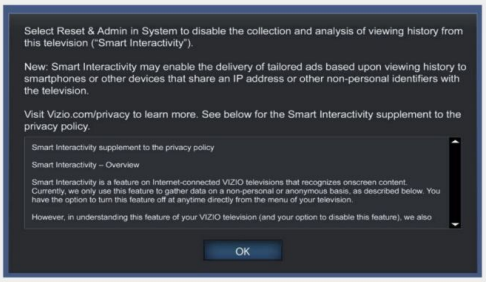
Jessica Rich, Director of the FTC Consumer Protection Bureau, speaking to the **Network Advertising Initiative annual summit** (May 2016): "In many cases, persistent identifiers such as device identifiers, MAC addresses, static IP addresses, or cookies meet [the FTC test]."

In the matter of Vizio (settled February 2017): Sharing of IP Address information involves improper sharing of PII even if the purpose is the creation of "anonymized" data.

Count I

15. As described in **Paragraph 14a**, Myspace represents, expressly or by implication, that it will not use or share a user's PII except as described in the privacy policy, including sharing that information with third parties, without first giving notice to and receiving permission from that user.

16. In truth and in fact, as described in **Paragraphs 7 through 13**, in numerous instances Myspace provided the Friend ID of the viewing user to third-party advertisers who are not affiliated with Myspace. The Friend ID gives access to, at a minimum, the user's basic profile information, which for most users includes their full name. This use was not described in the privacy policy and Myspace did not receive permission from those users for such sharing. These facts would be material to consumers in their enrollment in and use of the Myspace service. Therefore, the representations set forth in **Paragraph 15** were and are false or misleading and constitute a deceptive act or practice.



FUTURE OF PRIVACY FORUM

Pseudonymisation in the U.S.

HIPAA Limited Data Sets: 45 C.F.R. § 164.514(e)(2).
A limited data set is protected health information from which certain specified direct identifiers of individuals and their relatives, household members, and employers have been removed.
A limited data set may be used and disclosed for research, health care operations, and public health purposes
Data recipient must enter into a data use agreement promising specified safeguards for the protected health information within the limited data set.

FERPA research exception: 34 C.F.R. 99.31(b)(2)(i)-(iii).

- Educational agencies, institutions, and parties that receive education records from them may release "de-identified student-level data for the purpose of education research…
- ….by attaching a code to each record that may allow the recipient to match information received from the same source."
- Code-creator must not disclose "any information about how it generates and assigns a record code, or that would allow a recipient to identify a student based on a record code"
- Record code cannot be used for any other purpose, and must not be able to ascertain student PII, and cannot be based on a student's social security number or other personal information.

FUTURE OF PRIVACY FORUM

Anonos

Slides for Gary LaFever
(CEO & Co-Founder at Anonos)
(Four slides in total)

Gary LaFever
Chief Executive Officer

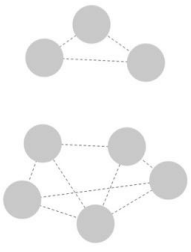
Anonos

Obfuscating Connections Between Data Elements

Re-Identification is Possible

These nodes represent data elements related to two different data subjects that are capable of being tracked, profiled and analyzed by third parties.

Each element can be be easily associated with and re-identified for each of the data subjects.

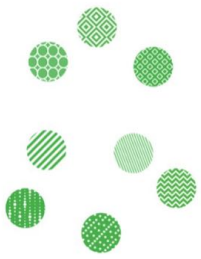


Anonos

Dynamic De-Identification

These nodes represent the same data elements that can be retained without loss of context necessary to support data analytics.

This is done by obfuscating connections between each of the data subjects and the data elements using dynamic de-identification to achieve controlled linkability of data.



Readily Linkable vs. Controlled Linkable

Re-Identification is Possible

Readily Linkable Persistent Identifiers

Identity of data subjects serves as the centralized means of tracking, processing and analyzing:

"Who"
each user is

"What"
a user does (or does not do)

"Where"
a user is when they do (or do not) do certain things

"When"
a user takes (or does not take) specific action(s)

"Why"
a user does (or does not do) certain things

Dynamic De-Identification

Controlled Linkability / Non-Persistent Pseudonymised Tokens

Highest practical level of de-identification for any given data

Optimizes the balance between maintaining the utility of data and protecting data privacy and security

Data values can be used selectively

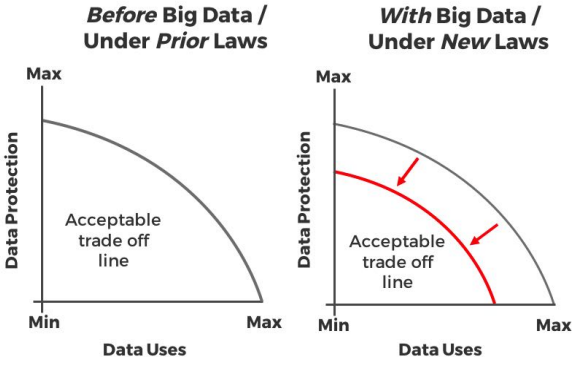
Anonos

Have Your Cake and Eat it Too™

HAVE YOUR CAKE - avoid liability, complying with regulations, ensuring trust (Data Protection)

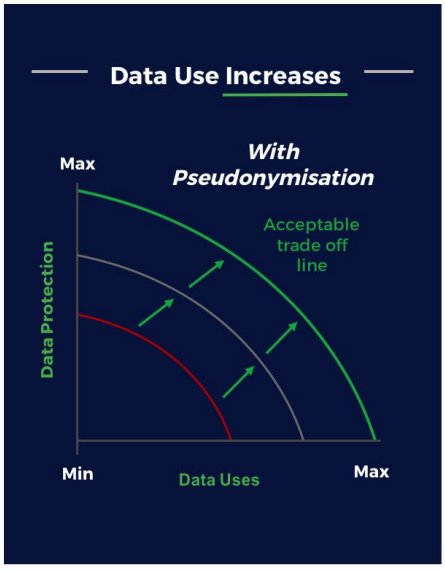
EAT IT TOO - process analytics, AI, personalization, maximize data value and sharing (Data Uses)

Data Use Decreases

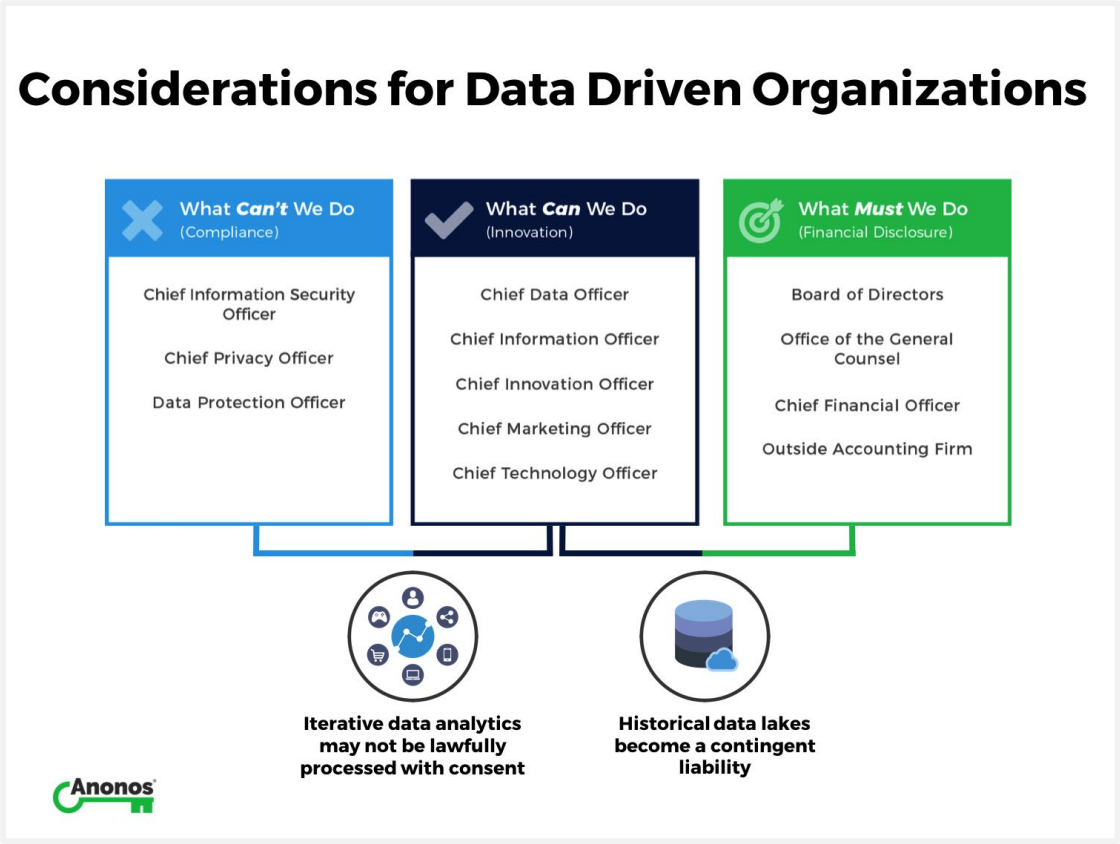


Anonos

Data Use Increases



Anonos



Note:

Webinar panelists' answers to questions – those asked and answered during the live event as well as those submitted during but not answered until after the event – will be provided to webinar registrants following coordination of answers from panelists.

Contact us at **BigPrivacy@anonos.com** to learn about saving your data and enabling protected data analytics under the GDPR.

To learn how BigPrivacy technology is used in your industry, visit **anonos.com/usecases**

Anonos®

BigPrivacy® Unlocks Data