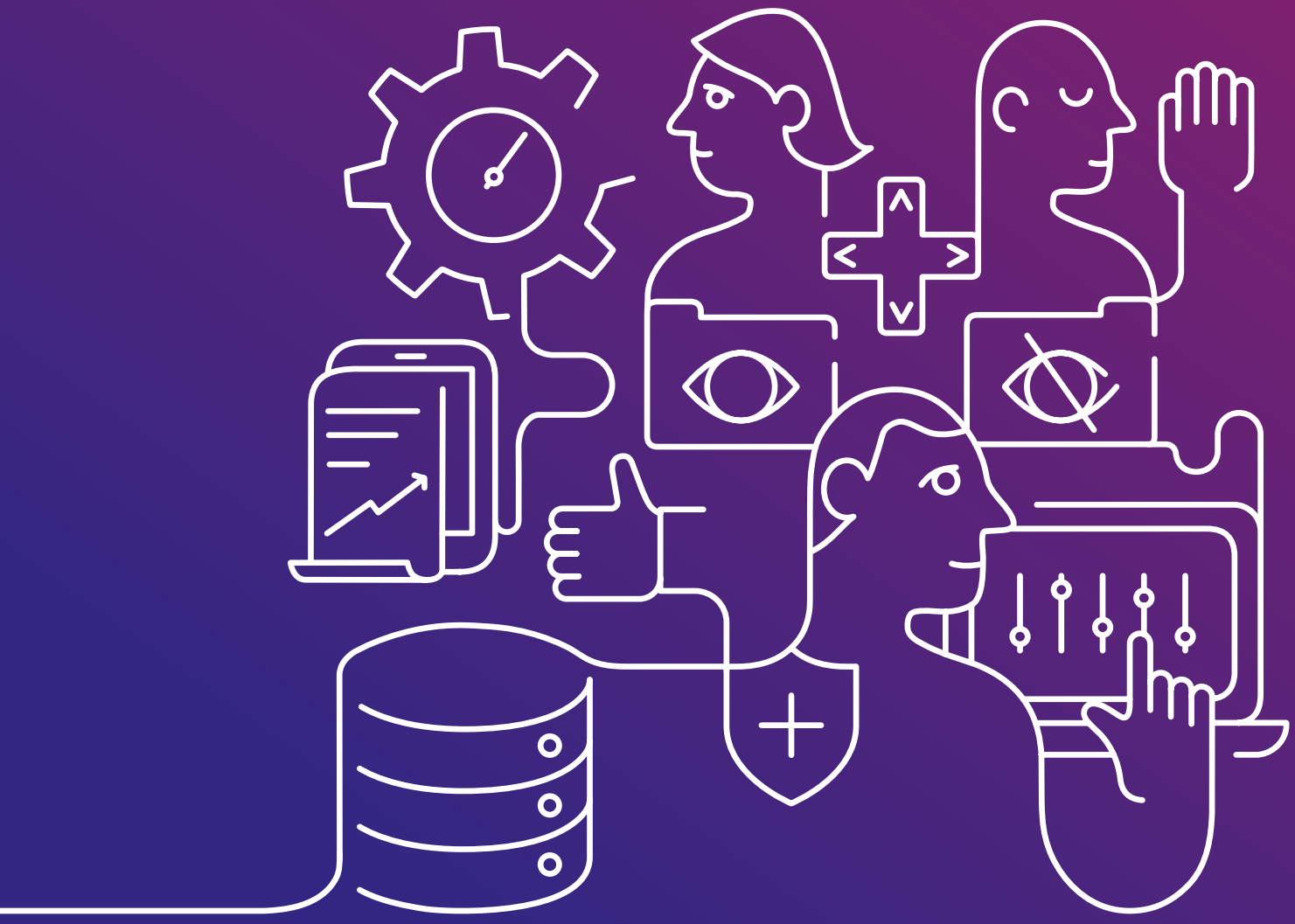
PRIVITAR

Enterprise data privacy with Privitar Publisher

Scalable and flexible data privacy execution and management



www.privitar.com

Why is data privacy important?

Data drives modern businesses. Across every industry there is enormous potential to turn the rich datasets organisations hold into groundbreaking insights, innovative products and scientific excellence. Availability of data is a key driver for many of the most exciting technology innovations that are emerging today, such as the growing importance of machine learning.

For all the benefits, the new data economy has ignited concern from citizens and governments. Consumers are sensitive to misuse of their personal data within organisations, and data breaches are a serious issue as rogue actors aim to steal highly valuable personal data. Acting without due consideration of privacy issues puts organisations at risk of regulatory sanctions, potential litigation and loss of reputation with customers.

As organisations seek to leverage data for exciting new projects, they must instigate tighter data governance controls that must cope with different use cases, using at-rest and streaming data, and on-premise and cloud deployments. Such governance projects must address the risks inherent in accessing personal data, but without stifling innovation.

This is the context in which Privacy Engineering techniques are emerging as key tools for all organisations across the public and private sectors that are using data to safely unlock insights, improve efficiency and create exciting new services.

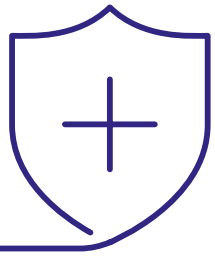
How can you balance a need for data accessibility with controls on the risk of disclosure?

The twin requirements of making personal data available for useful processing while addressing the risks of unauthorised disclosure requires dedicated processes, and implementing such processes is one of the key functions of privacy engineering solutions.

There are many data processing techniques that are useful in reducing the risk inherent in a dataset. The application of these techniques to an organisation's body of data is a fundamental capability of any privacy engineering solution. In addition, when applying any privacy-preserving transformation, it is important to understand the effect on usability. It is critical to ensure that the data, after being made more private, remains fit for purpose.

Also, every data privacy concern is contextual: a dataset may be released for several different purposes, so the privacy considerations on any one case must take into account its particular tolerance for risk and the amount of data required to fulfil it. Privacy engineering solutions must help understand the content of a data release, protect it against risks, and work to minimise its scope.

Furthermore, in organisations with ambitious data requirements, managing numerous individual data releases can be daunting. Data owners must ensure that every point of use meets corporate standards on data privacy. Privacy engineering solutions therefore have an important role in scaling an organisation's overarching privacy principles to every release, regardless of where or when processing actually happens.



How can Privitar Publisher help solve these issues?

Privitar Publisher is a comprehensive solution for data privacy protection and governance that helps organisations safely extract value from confidential data through a standardised privacy protection approach.

Publishers key features include:

- A comprehensive set of privacy enhancing techniques that can be assembled into reusable policies tailored to a wide range of use cases.
- Controlled data sharing with Protected Data Domains (PDDs), watermarking, and lineage.
- Centralised management across data on-premise, in the cloud and streaming, with audit.

Adapt to diverse use cases with flexible privacy policies

It is important to be adaptable to specific use cases when planning privacy protections. Publisher adopts a modular approach to its privacy policy system, allowing policies to be assembled flexibly from a wide-ranging set of privacy processing rules.

Publisher's capabilities include tokenisation, where sensitive values are replaced with randomly generated pseudonyms, and generalisation, where values are replaced with a lower-resolution 'blurred' form.

- In situations where protection of raw values is important, such as data preparation for testing, tokenisation is a good choice.
- To retain utility without disclosing exact values, generalisation can be used. This allows analysis of data to be performed and similar results obtained, but protects sensitive data.

Controlled data sharing with Protected Data Domains

Publisher uses several techniques to deliver its privacy capabilities and key among them is the ability to group data releases in Protected Data Domains (PDDs).

A PDD is a set of managed data releases with controlled, understandable privacy risk. It is the unit of data for privacy governance and management that records data lineage, permitted recipient, purpose and lifetime of the data, and what privacy protections have been applied.

PDDs are particularly suited for data sharing with internal and external parties. In accordance with data minimisation principles, PDDs allow safe sharing of just the essential data required for a task or project, and nothing more.

PDDs control direct dataset linkability: datasets published to different PDDs will not be directly linkable, thus reducing the overall risk of multiple dataset releases.

In addition, PDDs also include dataset watermarking technology that acts as a deterrent to users disclosing their datasets to others in an unauthorised way. If a breach occurs, recovering a watermark from a file involved in a breach can make it easier for an organisation to respond.

Secure collection and linking of datasets with Privitar SecureLink

Privitar SecureLink enables the collection and safe linking of datasets from multiple contributing organisations without ever revealing the private identifiers on which the data is linked, not even during processing.

The identifiers never leave the contributing organisations' boundaries unencrypted but thanks to an advanced encryption scheme can still be used by the central party to join data together and create a richer and more valuable dataset.

Standardised governance and management with Publisher Policy Manager

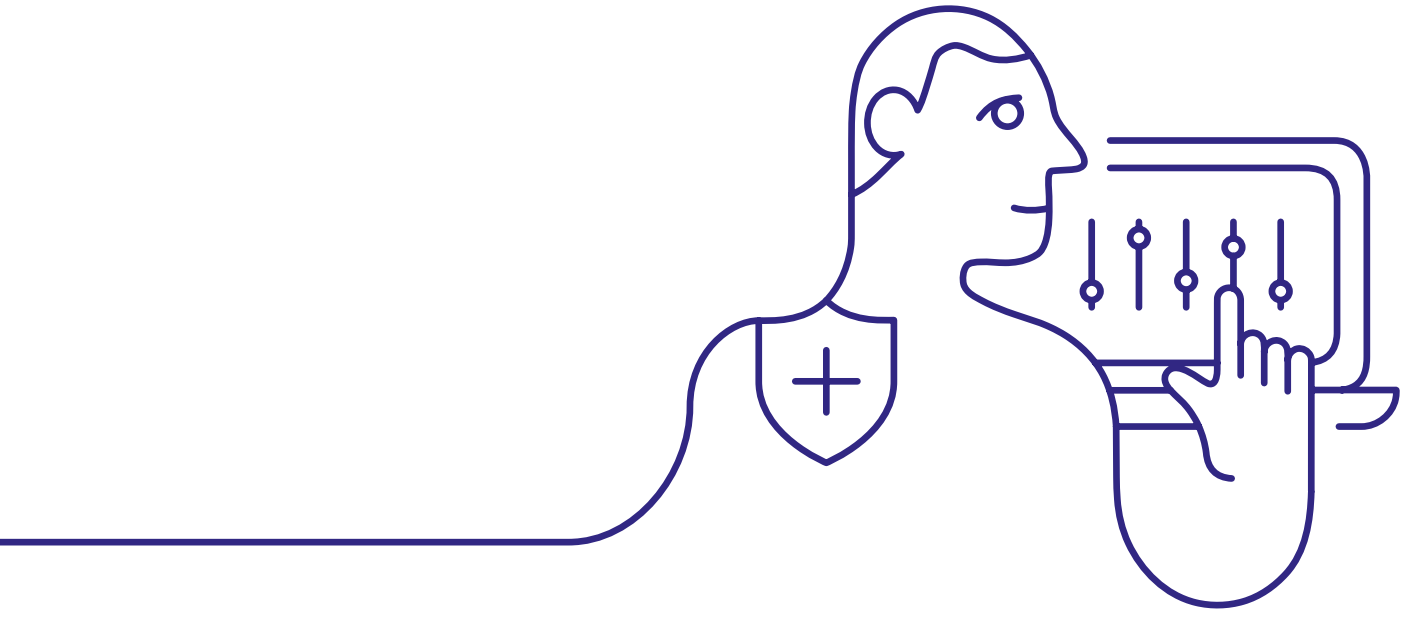
Privitar Publisher's Policy Manager allows policies to be applied consistently across multiple types of environment.. The same privacy-preserving transformations can be applied on static and streaming data, on premises or in the cloud.

It is important for data owners to have good governance oversight into privacy processing wherever it occurs. Publisher provides a central Policy Manager where all privacy processing is defined and monitored. The Policy Manager ensures a consistent application of privacy standards across the organisation. A team-based authorisation model allows control to be retained centrally or delegated to different departments or business units, while allowing for consistent rules and collaboration.

Every dataset that is managed by Publisher has a full recorded history providing an audit trail of where the data came from, how it has been adapted and to whom it is made available.

Automated de-identified data provisioning with Publisher REST APIs

Privitar Publisher offers a rich set of REST APIs that can be used to automate data provisioning flow and re-use metadata that may already be stored in the organisation. Thanks to its REST APIs, Privitar Publisher's policy management capabilities can be applied to large numbers of datasets and policies in an automated, standardised, predictable and audited manner.



How does Privitar Publisher work?

Privacy policies apply data protection controls such as tokenisation and generalisation as data is published into Protected Data Domains (PDDs). Datasets within a PDD optionally retain referential integrity, but are never directly linkable to another PDD. This separation enables the data owner to isolate risk between PDDs and to understand the implications of publishing or sharing a PDD.

Policies are created by data owners to be appropriate for the use case in question, and Publisher applies any data transformation needed to deliver the policy using processing engines that scale depending on the volume of data that needs to be protected.

Publisher can apply privacy protection to data in three ways:

Data flow and streaming

Publisher integrates easily with data flow tools such as Apache NiFi, Apache Kafka and Confluent Platform, so that de-identification drops in easily as a step in a live workflow.

Data flow and streaming frameworks support multi-node operation for scalability and fault-tolerance, and Publisher takes advantage being able to apply a policy in parallel to multiple data items.

Hadoop

Publisher can apply policies to large Big Data repositories using Apache Spark as a scalable and fault-tolerant processing engine.

Publisher supports cloud and on-premise deployments, and all major Hadoop distributions.

On-demand

Publisher provides a web-based secure API that can be used to perform de-identification within an organisation's own applications and services.

When operating in this way, the benefits of centralised policies and PDDs are still present because the API service synchronises with the central Policy Manager.

What are the key benefits of Privitar Publisher?

Publisher is applicable to a wide range of use cases and can be deployed to support adherence to both industry and governmental regulations, including GDPR.

Although each use case is unique, Publisher has the following major benefits:

Enabling safe data utilisation

Publisher enables its clients to de-identify data using the latest privacy techniques.

Its rich and flexible tool set of data protection features allows organisations to reduce risks from data releases while preserving utility of the de-identified data for their specific use cases and analyses.

Simplifying data privacy governance

Publisher allows an organisation to define privacy policies that determine how data must be protected across an organisation. This allows organisations to understand and control how protection is being applied, even when large amounts of data are being processed for many uses.

Publisher provides a single Policy Manager interface for defining and monitoring privacy governance, centralised across a data estate.

Providing linkage control

Appropriately managing data linkage is an important consideration in effective data protection strategies.

Publisher offers powerful tools to safely collect and join data from various sources, as well as control linkage of disseminated datasets.

Operating at enterprise complexity

Publisher accommodates the needs of organisations operating at large-scale enterprise complexity.

This encompasses the ability to work in the context of advanced organisational and data governance structures, meeting the high security requirements of enterprises, as well as enabling integration and automation of data protection workflows.

Summary

Protecting data privacy is an essential element of information governance for any data-driven organisation. It's not just the internet giants that need to redress their approach to data privacy; the issue is a board level concern across public and private sectors and a broad range of industry verticals.

Organisations must address privacy or face significant issues further down the line. What most are striving for is to find a balance between protecting privacy while still retaining utility in the use of the personal data they process. This requires a fundamental first step that acknowledges that privacy must be a core consideration and must be designed and engineered into data processes.

Privitar Publisher was created to help organisations achieve this balance. It is a deployed solution that is already helping to protects millions of personal data entries through a principled method of understanding and significantly reducing privacy risk..

As organisations across the world seek to exploit the potential offered by advanced uses of data, but also the regulatory and public concerns, addressing privacy risk is a critical hurdle to clear.

We're Privitar

We help organisations engineer privacy-preserving data operations, using advanced data privacy techniques that protect sensitive information while retaining data utility. Our software accelerates and automates privacy-safe data provisioning, helping our customers get more business value from their data, generate data-driven insights, and drive innovation.

Contact us:

e: info@privitar.com
t: +44 203 282 7136
w: www.privitar.com

PRIVITAR

@PrivitarGlobal

www.privitar.com