PRIVITAR

Share Anonymized Data Securely With Digital Watermarks

Working with Sensitive Data

Companies are eager to build innovative partnerships but are frequently reluctant to leverage the skills and expertise of third parties due to concerns over sharing data.

Responsible companies should protect individual privacy before sharing, by removing or obscuring the sensitive data through anonymization techniques. They should also control and track the distribution of sensitive datasets by inserting a digital watermark into the data.

Anonymization techniques reduce the risk of an adversary identifying one or more individuals in a dataset. Watermarking enables detection and attribution of unauthorized distribution or publishing of a sensitive dataset, and so acts as a deterrent against such unauthorized behavior.

The Privitar Data Privacy Platform™ embeds a digital watermark in anonymized data, which can be recovered should the data be leaked.

Why Use Watermarks?

- Anonymization of a dataset reduces the risk of re-identification but does not completely eliminate the possibility, which is why data sharing should be only be done under contractual restrictions. Watermarks reinforces this.
- The Privitar Platform stores an audit trail of who authorized data access to which user, and for what purpose, keyed by the watermark.
- Digital Watermarks can be embedded in any anonymized dataset.

How Does It Work?

Digital media commonly carry watermarks within metadata or redundant data in the file.

However, raw datasets (i.e. a tabular file, an extract from a relational or semi structured non-relational database, or the result of an interactive database query) typically don't contain such metadata or redundant data, and so digital watermarks require manipulating or perturbing the data itself, leading to an undesirable loss of information and utility of the dataset.

Privitar's privacy-preserving algorithms modify the raw dataset using a combination of tokenization, generalization and other perturbation techniques to resist re-identification of individuals in the dataset. This adds a small amount of noise to the data, enabling The Privitar Platform to also embed a watermark.

Privitar Publisher distributes the watermark redundantly throughout the dataset. This makes it difficult for an adversary to remove the watermark either by releasing only a subset of the data, re-ordering the data or by perturbing the anonymised data.

Benefits

> Watermarks provide a robust mechanism to record why an anonymized dataset has been made available, to whom it has been provided, and the date of creation.

> The Watermark survives when the data is moved between formats (relational to Excel to flat file) and also survives the data being reorganized or filtered.

> The presence of the Watermark ensures the data recipient takes their responsibilities seriously.

We're Privitar

We help organizations 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: UK +44 203 282 7136
US +1 857 347 4456
w: www.privitar.com