The Undetectable Cyber Security Threat: Data Poisoning

by Tributech Solutions GmbH

Executive Summary

Gartner and other think tanks agree that data poisoning will be a major cyber security threat in the future. Firstly, because just a very little percentage of poisoned data can have a huge impact on the decision accuracy of an application or algorithm. Various research institutions have conducted data poisoning experiments that all confirm this fact. Secondly, companies currently do not have any mechanisms in place to detect poisoned data and existing technologies do not solve the problem as they only cover parts of the data pipeline.

What is Data Poisoning?

The term data poisoning initially was closely connected to training data for AI or ML algorithms. It describes the intentional manipulation of training data in order to influence the decision of the algorithm. However, over the years the term data poisoning has taken on a more general meaning. Nowadays it's being used for the intentional manipulation of any kind of data in order to influence the decision of an application or algorithm. In other words, data poisoning can be seen as the industrial "fake news" of the 21st century.

The reason why data poisoning is not such a common term is that data poisoning attacks do not have to be disclosed to the public, in most cases. The GDPR (General Data Protection Regulation) disclaims that only data breaches containing personal information have to be published. However, data poisoning attacks mostly do not target data that includes personal information. Hence, when companies fall victim to a data poisoning attack, they do not have to publish the incident, in most cases.

A Little Poisoned Data – Huge Impact

One of the reasons why data poisoning is so dangerous is that just a very small amount of poisoned data can have a huge impact on the decision accuracy. Different research institutions have conducted several data poisoning experiments that all support this statement.

1. Data Poisoning of a picture recognition platform

In 2020 the University of Maryland conducted a data poisoning experiment on the Google Auto ML platform, which is a picture recognition platform. In the process of this experiment, they tested the decision accuracy of the algorithm in detecting the right object on a picture and how much poisoned data would be needed for the algorithm to take a wrong decision.

Picture: Aaron Foster/Getty Images

Picture: Shutterstock

Figure 1: Data Poisoning: Google Auto ML

The result was that only 0.1% poisoned data lead to the algorithm identifying the frog as an airplane in nearly 80% of the cases. In other words: **0.1% poisoned data lead to an error increase of 80%.** (University of Maryland. 2020)

2. Data Poisoning of e-mail spam filters

In 2021 the Northeastern University executed a data poisoning attack on e-mail spam filters. The aim of this experiment was to manipulate the algorithm in order to classify e-mails including malware as safe. By inserting only 1% of poisoned data the algorithm marked malicious mails as safe in 90% of the cases. In other words: **1% of poisoned data lead to an error increase of 90%.** (Northeastern University, 2021)

3. Data poisoning of smart sensors

The most recent and most advanced data poisoning attack was conducted by the DARPA, which is a moonshot research arm of the Pentagon, focusing on defence and advanced research projects. In January 2022 the DARPA successfully executed a data poisoning attack on the electrical grid of the US. They hijacked a smart sensor and manipulated the data that was sent to the headquarters. However, the IT infrastructure of the headquarters had no mechanisms to detect the poisoned data. Eventually, the headquarters took wrong actions based on fake data, which lead to black out in Plum Island, New York. (Bloomberg, 2022)

How to Protect Yourself from Data Poisoning?

As seen in the described experiments the huge threat of data poisoning lies in the lack of detection: Currently, companies do not have any mechanisms in place to automatically detect poisoned data. At the moment some experts even say that once a data poisoning attack is detected, the damage is already done. The reason for the lack of detection is that existing technologies do not cover the whole data pipeline and data lifecycle.

Encryption is not enough

The growing extent to which businesses, infrastructures, data platforms, companies, etc. are interconnected means that data “moves” across system boundaries during its lifecycle and is used for different use cases. Encrypting data end-to-end helps ensure that data is unaltered when in transit or at rest. However, it has limited capabilities when

- Data is transferred across different systems
- There is a time delay when data is transferred from the source to the consumer or
- There are data scheme transformations

For example, a company that operates energy generators, collects production data and stores it in its infrastructure. A selection of this data is transferred to the energy provider for reconciliation and billing once a month. Although the data is encrypted end-to-end, it is not possible to guarantee that the data is unaltered due to the time delay between the data transfers and the transmission across system boundaries.

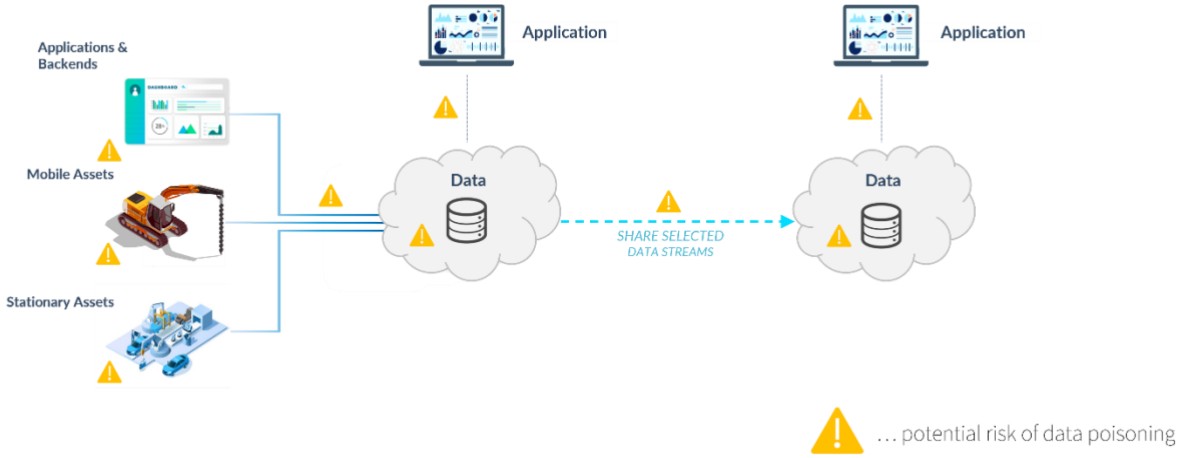

Figure 2: Attack points for data poisoning

Figure 2 illustrates a typical data pipeline and the potential attack points for data poisoning. The first attack point is directly at the data source, before the data is being collected. The second attack point is data in transit, when it is sent to some kind of database. Followed by the data at rest before it is being used by an application. Even though this is an example of first party data, a person looking at the dashboard of the application has no ability to verify if the data he/she is looking at is the *true data* from the desired data source. If data is being shared across infrastructures and systems, companies must blindly trust the data even more.

Tributech offers the **only** solution on the market that allows organizations to immediately detect poisoned data along the whole data pipeline and data lifecycle. Once data gets collected the patented data notarization service creates cryptographic proofs of each datapoint. No matter where the data is being transferred to, from then on it is possible to verify origin and integrity of any data point over the entire data lifecycle and thereby detect poisoned data immediately.

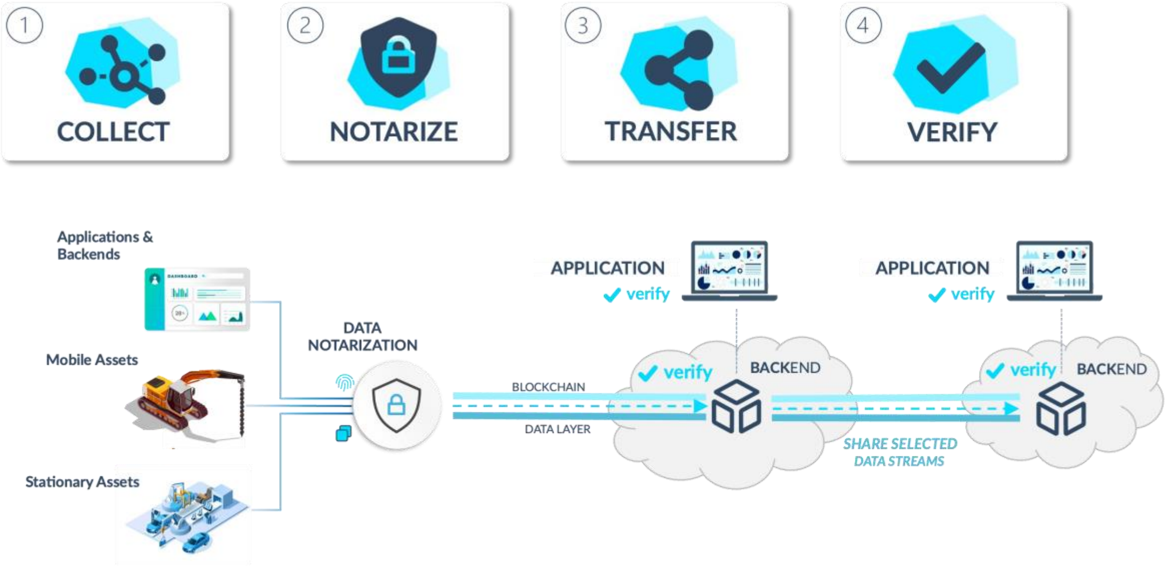

Figure 3: Tributech's technology

How Does the Data Notary Service Work?

In contrast to established proof and control mechanisms for assets within our physical world, there is a general lack of such mechanisms when it comes to digital assets. This becomes apparent when digital assets leave their company or system boundaries. Even if a system was set up according to the highest security standards, it becomes impossible to verify the origin and integrity of a dataset at any later point once it leaves the system's boundaries.

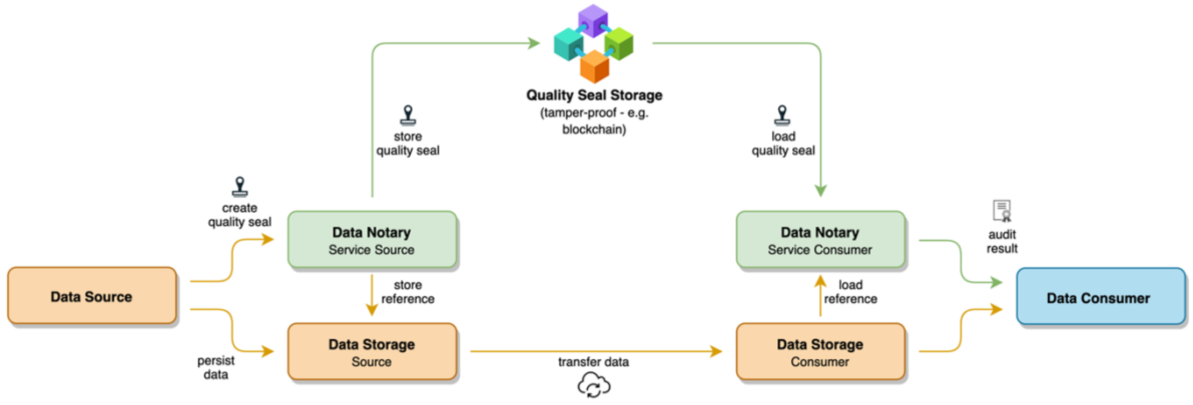

Figure 4: Data Notary

Data Quality Seal

The lacking component for this task is a single source of truth. A single source of truth needs to store tamper-proof records in form of a data quality seal for data Integrity, authenticity and other quality parameters of data, for all users across the data life cycle.

Data Authenticity

Data authenticity deals with the source of data and the required proof thereof. So-called public key certificates, digital certificates which can reliably identify their owner via a publicly available cryptographic key, are an ideal tool for proofing data source and ownership. These public keys are one of two integral components of public-key cryptography, private keys form the second. Private keys are held by only their individual owner and get used to create digital signatures which enable the proof of data ownership.

A Data Notary Service therefore enriches data with a public key, in order to make data authenticity provable. The private key, in addition, ensures that only the original data source, respectively the data owner, can control that certain data was indeed signed by itself and no other party can falsely claim original ownership.

Data Integrity

Data Integrity is understood as the original, untampered state of data. So-called hash functions, which are well established in the world of cryptography, form the basis for proving data integrity as part of a data notary service.
In essence, these cryptographic functions utilize complex mathematical algorithms in order to calculate a unique output value for any data provided as input. Even minor changes to the data provided, will result in a different output value, ultimately making any changes or manipulations of the original data traceable. Another major benefit of these hash functions is that the volume of input data does not affect the volume or size of the output value.
By utilizing these hash functions, a data notary service can therefore create a sort of digital fingerprint for any kind and volume of data like for example sensor data, documents, configuration files, raw data and more.

Data Source State

The configuration of a data source, as well as environmental factors, can have significant impact on the generated data and its quality. In order to ensure continuous data quality and an added level of traceability, it is important to also keep track of any such changes in the source configuration and its environment. These factors and any source configuration changes get tracked via the so-called data source state, which is the third and final compound of our data quality seal.

Tamper-proof Storage for Data Quality Seals

In order to guarantee the validation of data authenticity and data integrity in the long run, it is important to store these data quality seals in a tamper-proof and persistent manner. Data Quality Seals need to get created as early as possible in the data life cycle, as close to the data source as possible and further must not be mutable after they got stored. A main requirement for a data quality seal storage solution therefore is the immutable, append-only character of any data entries. This makes innovative tamper-proof storage technologies like for example Blockchain, a distributed ledger technology, an ideal choice for this task.

Per default, Tributech’s data notary solution integrates a private permissioned blockchain technology that is based on the Tendermint Protocol. It is optimized for high throughput to allow highly scalable and cost-efficient data notarization for any kind of data.

Data Audits

Auditable data can be seen as a warranty promise and plays a key role when it comes to business decisions and relationships. The verification of this warranty (proof of integrity, authenticity and additional quality parameters) is taken care of by data audits, which allow you to check the formerly introduced data quality seals and further creates audit reports for your analysis. Not only does it allow the verification of data internally, it also enables customers, suppliers and partners to verify data that they are consuming from your data sources.

Conclusion

Data Poisoning presents a huge danger for two main reasons:

- **Lack of detection**: Currently, companies have no mechanisms to detect poisoned data.
- **Small amount – huge impact**: Only a very small percentage of poisoned data massively influences the decision accuracy

Hence, whenever data is used for business actions or critical decisions, it is of absolute importance that this data can be verified and trusted. The audit of data quality seals is a quality assurance measure that typically gets executed prior to any such (critical) action, in order to ensure no decisions are based on poisoned data.

If you want to be protected against the undetectable cyber security threat of data poisoning, scan the QR-code and get in touch with us!