



How to share data while protecting sensitive information.

The advantages of sharing data for companies and the importance of data privacy.



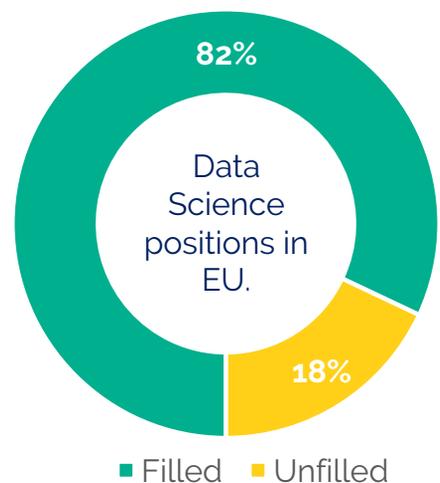
The value of sharing data between partners.

The value of customer data for internal operations, insights, and product development is undisputed. Customer data has manifested itself as a major resource for the modern business world.

Now, many companies have started exploring the potential of sharing data across organizations. The general consent across companies is that sharing data between partners provides a major collaborative and strategic advantage and sharing data for collaboration can provide value to companies in many ways.

Access external talent

The IDC¹ estimates that by 2020, almost 1 in 5 Data Science or Data Analytics positions will be unfilled in the EU, leaving companies to find other ways to access the much needed talent. In order to fill this skill gap, companies work with agencies or offer data science competitions via hackathons or platforms such as Kaggle to share data with external data scientists, where they can work on new challenges like for example building complex machine learning applications for companies.



Enrich your available information base

According to the Harvard Business Review², 64% of surveyed companies engage in data blending and use more than 5 different data sources for more complete information. Enriching internal data with external data sources also has the potential to unlock prior entirely unknown relationships and insights.

Leveraging multiple sources of data enables companies to explore new business opportunities as well as find relevant partners for new data-driven initiatives.

Foster collaborative product development

Next to finding new information and business opportunities, leveraging several sources of data from various partners can fuel data-driven applications of superior quality. For example, outside sources of high quality data provide better context for consumer products that leverage machine learning and in turn enable a higher quality product for end users.

Support scientific research

The scientific community heavily relies on access to necessary data. A prime example is research in the healthcare sector, where researchers frequently collaborate with insurers, hospitals, pharmaceutical companies and other institutions to obtain data for their research.

Necessary data requirements

Of course these use cases also come with specific requirements. Across all the mentioned scenarios, companies and collaborators have to make sure to share and work with truthful and granular data that is available and privacy-preserving.



Data needs to be complete and granular

Aggregated statistics and information still help business analysts to make decisions. In order to do so, query access to a database often serves the cause.

Nevertheless, with the rise of intelligent algorithms and the increasing demand of data scientists, far more detailed sources of insights are needed. Nowadays, data scientists require complete and highly granular datasets that allow them to perform complex analyses on top of such data.

Data needs to be available

Availability means that data can be used by authorized people when required.

Therefore, again, a simple query access to a database does not ensure that required data is always available when needed. Only by possessing an actual dataset on their own infrastructure, companies and data users can be assured that the data will be available at any given time.

Data needs to be privacy-preserving

According to our research, data privacy currently is one of the major reasons, companies refrain from sharing data with partners. Although willing to collaborate, companies currently often just don't see a secure way to share their data.

The need for data privacy has been traditionally solved by companies anonymizing their data. While seemingly simple, there are severe shortcomings of traditional anonymization methodologies.

The problem is that providing truly anonymized data is difficult. Anonymizing data in a privacy-preserving manner takes time, resources, and significant domain expertise. Traditional anonymization technologies usually have two problems. Either they are not sufficient and do not protect data properly enough, or they change data to an extent where it is barely usable for a variety of use cases. So, even if the generation of truly anonymous data has been successful, this often equals a significant loss in data utility, thus rendering an anonymous dataset useless.

Stalice provides a solution

Privacy-preserving synthetic data makes anonymizing data easy while maintaining data utility and data granularity. By leveraging the recent advances in machine learning and state-of-the-art privacy techniques, privacy-preserving synthetic data enables companies to release highly granular datasets with no risk of identifying a single individual.

This empowers companies to open up their new synthetic data in a GDPR-compliant manner and enable them to share actual datasets and not just information for product development, training new machine learning algorithms, and unlocking industry-wide insights.

Privacy-preserving synthetic data

Privacy-preserving synthetic data consists of entirely artificial and new data points and is guaranteed anonymous, while resembling the statistical properties and structure of the original dataset. It possess the following attributes:

- Properties and statistical information of the original data are mirrored in the synthetic data.
- The data structure of the original data is retained unaltered in the synthetic data.
- It is impossible to identify real individuals in privacy-preserving synthetic data.

Curious?

Get in touch with us to discuss how your company can also fully leverage data all the while protecting its privacy.

www.statice.ai

hello@statice.ai

+49 157 53072551



Stalice enables the open data economy. We provide you with privacy-preserving synthetic data that can be freely used and shared. Turn data privacy into an asset and **unlock your data.**

Stalice GmbH

www.stalice.ai

hello@stalice.ai

+49 157 53072551