Aimino
AI Applications with Synthetic Data and High Data Quality

Everyone is talking about Big Data, Artificial Intelligence, Machine Learning. Everyone says that automating tasks requires more data. However, the primary focus is often on optimizing models and collecting more and more data, while the models themselves and the amount of data are of secondary importance. The quality of the data, however, is crucial for the performance and value of AI applications. And the startup Aimino builds on this insight.

Combination of Real and Synthetic Data

Currently, the tedious collection and preparation of data sets for teaching the AI system accounts for 80 percent of all work. To change this, Aimino, a spinoff of KIT, relies on a data-centric approach and offers companies suitable services and tools for efficient cleaning of already existing datasets. In addition, Aimino enhances datasets with its synthetic data adapted to customers' purposes. With its expertise, Aimino can significantly improve the quality of data for companies, for example in the electronics, automotive, and pharmaceutical industries, thus reducing the workload while increasing the efficiency of the AI model.

Security through Automatically Generated Fake Data

There are numerous approaches to protecting against data breaches. The usual strategy is to maximize the security of potential entry points of attackers to make it as difficult as possible to access crucial data.

Aimino takes an alternative approach, namely protection using synthetic data. This is data that is automatically generated via artificial intelligence (AI). Synthetic data can be generated for various uses, such as visual or structured data found in spreadsheets, or sensor data. Using this synthetic data, companies can now train their models for rare scenarios. Collecting the corresponding real data would be much more costly.

Companies benefit twofold: In a comprehensive data set that contains a lot of newly generated credible fake data, real data almost gets lost. This reduces the value of this data set for attackers. It becomes disproportionately difficult for them to separate the real data from the fake data. Access to such a data set thus becomes less attractive. Should they still succeed, misuse of the data would be more time-consuming.

When companies use AI for security reasons, Aimino's synthetic data helps increase the accuracy and performance of AI. Generated data, which would otherwise be very costly and time-consuming to collect, can be used to increase the robustness of AI in many areas.

The well-done pattern in the figure on the left served as the basis for the synthetic generation of the four following samples: Scratched, flipped, colored, and bent samples.
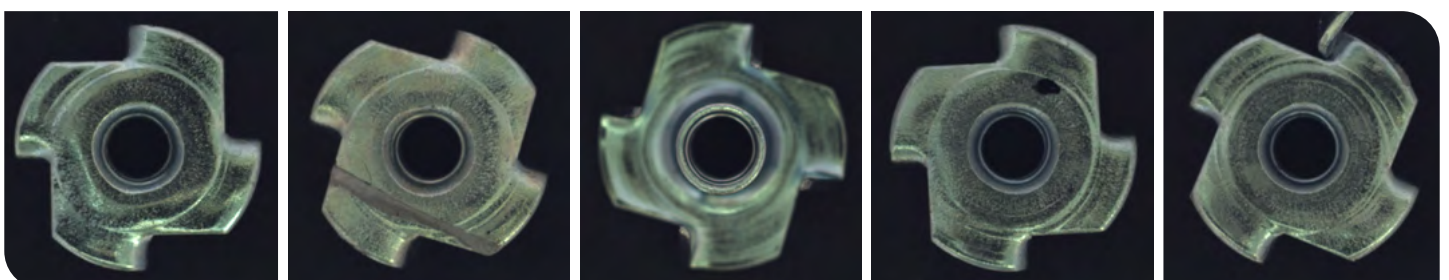
The Amelia Data Generation Tool

Karlsruhe-based startup Aminio makes practical AI applications accessible to everyone. With Amelia, a tool for generating data, companies need up to 95 percent less data for AI applications, thus saving time and costs while protecting their data. Aimino's technology requires only a few data points as the basis for the generation process. No encryption skills are required to use the Amelia tool. Aimino's synthetic data helps reduce the number of false alerts triggered by users' AI systems by up to 73 percent. Aimino helps AI-based companies overcome data scarcity and avoid errors by providing synthesized data and offering cleaning of existing data using the Amelia tool. With Amelia, AI engineers can generate customized synthetic data 50 to 250 times faster. Customers in manufacturing, for example, typically require synthetic defect images to improve the performance of their visual inspection system and validate such systems before release. Other applications, such as defective parts detection, require synthetic images with artifacts of the natural environment, such as different backgrounds and lighting conditions. Amelia can create all of these and also provides automatic data cleaning for direct generation of AI models.