

Newsenselab able to make medical data available for research while guaranteeing patients' anonymity

After implementing the Statice data anonymization software, the Newsenselab team validated the use of synthetic data to anonymize medical data without breaching the privacy of their application's users. They successfully anonymized more than 170 000 migraine symptom data points while maintaining a high utility for research. The EU Horizon 2020 framework program funded this project.

M-Sense, a leading migraine monitoring application

Since 2016, the Newsenselab team has dedicated themselves to empowering people suffering from headaches and migraines, helping them become experts of their own health. They developed a **monitoring and health assistance migraine mobile application**.

The M-sense application allows its users to better understand and reduce their symptoms. It also offers therapy methods to proactively combat the headaches. Aware of the medical information's sensitivity, the Newsenselab team is firmly attached to its user data privacy.

As part of the healthcare ecosystem, they are also committed to **helping the scientific community better understand the disease's patterns and causes**.

"M-sense started as an offshoot of a research project from the Humboldt University Berlin, and we stay true to our origins. We aim to provide fresh impetus for future migraine research."

Dr. Rer. Nat. Markus Dahlem,
Executive Director, CEO & Co-Founder of Newsenselab

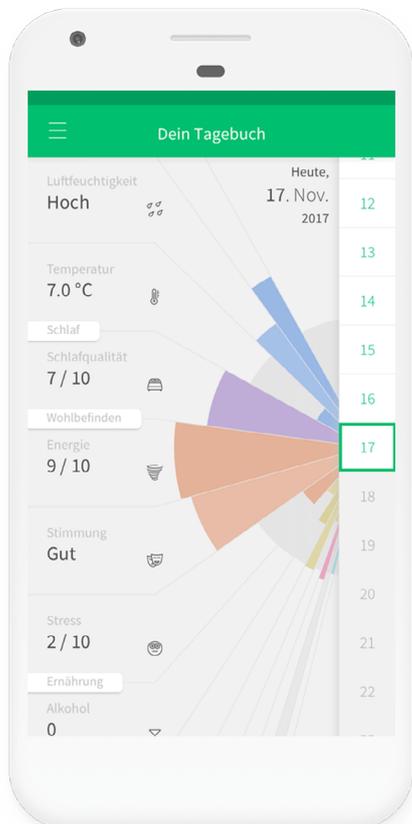
But due to the personal nature of most of their data and their engagement towards user's privacy, the Newsenselab team couldn't share the relevant information they wanted with their research partners.

The challenges of sharing medical user data for research purposes

In a field where data is hard to come by, healthcare applications represent a new data source to fuel medical research. For migraine research, **symptoms data can help researchers** better understand the diseases' patterns and causes, contributing to diagnosis and treatment research.

However, from the beginning, the Newsenselab team made a priority of protecting its users' privacy. Besides, M-sense had to comply with data regulations, as should European companies processing personal healthcare data.

The General Data Protection Regulation strictly regulates the processing of personal data of medical nature. And in Germany, the Digital Healthcare Act (DVG) adds up specific data privacy and security criteria for digital health applications like M-sense. As a result, most of their user data, including symptoms frequency and nature over time, could not be shared with research partners.



The M-sense application by Newsenselab

Simply stripping the data off its personally identifying information (PII) was not a practical approach for the team.

Even without PII, datasets are potentially able to be linked to open data sources to re-identify individuals, which is not considered a safe enough approach in the eyes of the GDPR. It compromised the usability and the safety of the data.

Either information wasn't anonymized enough to remove re-identification risks, which wasn't an option for the Newsenselab team, or the masking process altered the data too much to allow any analysis. Thus, the data remained unleveraged. And while users' privacy was protected, the team could not contribute to migraines' global research.

The Newsenselab team validated the use of synthetic data technology as a means of protecting data privacy

In healthcare research, **the statistical information in the data is more valuable than personal information** per se. However, in the context of privacy-preserving research, it's rarely possible to access the statistical information in the data without breaching privacy. And when data is made private through traditional protection means, it is often at the cost

of the data utility. One way of overcoming this dilemma is to use synthetic data, which the M-sense team decided to do.

The way privacy-preserving synthetic data works is **by generating artificial data points that mimic an original dataset's statistical properties** but contain none of its personal information. For the team, synthetic data represented an opportunity to contribute to migraine research by providing data that would **entirely protect their user privacy.**

As part of a research project, Newsenselab validated the use of synthetic data as a privacy protection mechanism for user medical data. The team implemented the Statice anonymization software, **running more than 170 000 data points through the privacy-preserving machine learning models**, to generate an artificial dataset consisting of migraine symptoms over data in a 10-dimensional space.

This new data had statistical and structural values highly similar to the original data. Still, it did not compromise the personal privacy of any individual from the original data.

Generating statistically accurate yet anonymous healthcare data

With Statice's solution, the Newsenselab team could get a new dataset generated from their user data. On this data, Statice ran privacy and utility evaluations to verify that it provided an **equivalent statistical value and the highest level of privacy protection**.

The synthetic data proved to maintain the statistical integrity of the original data. It also showed no one-to-one relationship with the sensitive medical data, ensuring no-one could tell whether a user had been part of the original dataset.

"Statice's data privacy solution offers the possibility to advance fundamental research while completely protecting the data of users."

Simon Scholler,
Head of Research at Newsenselab."

The M-sense team validated the ability to share data for research purposes in a privacy-preserving manner. This approach complies with existing personal and health data regulations and guarantees the highest level of privacy to their users. For Newsenselab, privacy-preserving synthetic data could contribute to accelerating medical research through collaboration with other healthcare providers and research groups.

Learn more and get in touch with us at www.statice.ai

About Statice

[Statice](#) develops state-of-the-art data privacy technology that helps companies double-down on data-driven innovation while safeguarding the privacy of individuals.

Thanks to the privacy guarantees of the Statice data anonymization software, companies generate privacy-preserving synthetic data compliant for any type of data integration, processing, and dissemination.

With Statice, enterprises from the financial, insurance, and healthcare industries can drive data agility and unlock the creation of value along their data lifecycle. Safely train machine learning models, finally process your data in the cloud or easily share it with partners with Statice.

About Newsenselab

With [M-sense Migraine](#), the Berlin-based health start-up Newsenselab GmbH has developed the first digital health application (DiGA) against migraine that can be prescribed by doctors and follows highest security standards. The app has been tested by the German Federal Institute for Drugs and Medical Devices (BfArM) and will now be reimbursed by health insurance companies within the framework of the Digital Healthcare Act (DVG).

M-sense Migraine combines various non-drug therapy methods from clinical practice, sourced from the guidelines of modern migraine treatment - personalized, motivating and always available to its users. The aim of the DiGA is to offer individualised therapy with the goal of attack reduction by analyzing individual pain patterns.

Newsenselab is also conducting clinical studies with partners such as the Charité Universitätsmedizin Berlin, as well as the University Hospitals of Halle and Rostock to investigate the effectiveness of DiGAs in the treatment of chronic diseases. Currently, participants from all over Germany are being sought for the ongoing migraine study EMMA, one of the first clinical studies to be conducted completely online.