

# Synthetic data to assist brain research

APRIL 8, 2022 · COLLABORATION, HEALTH, RESEARCH

**There are two good reasons why researchers are interested in synthetic health data: artificially created data does not contain patient data and can be produced in large amounts for research purposes. VTT is developing a method for creating synthetic imaging data for brain research in an international Industrial-grade Verification and Validation of Evolving Systems (IVVES) project.**

[Lue tämä suomeksi](#)

“If you look very carefully, you can find qualitative details in the image that might make you doubt its authenticity. Anatomically speaking, however, the images are sufficiently credible and suitable for research purposes,” says VTT researcher **Harri Pölönen** when presenting an artificially created human brain MRI.

The brain structures shown in the MRI images of the human head are unique and, therefore, it is impossible to fully fade out the data about a person's identity without the anatomical information relevant for research also suffering. The supply of authentic MRI images is limited, and their use is subject to a permit. Artificially created data provides a solution to these problems.

**Jaakko Lähteenmäki**, Principal Researcher at VTT, describes the challenges associated with the acquisition of health data needed for research use: “In [a recent study](#), we utilised health data from approximately 7,000 people. The permit process related to its use and the compilation of data from different information systems took 16 months, and the cost of compiling the data amounted to 80,000 euros. Delays and costs could largely be avoided by utilising synthetic data. At the same time, we could also avoid the need to define the content of research accurately in advance, which would facilitate the development of new data-based innovations.”

Harri Pölönen has already created thousands of [synthetic 3D brain images](#) in the past year. Brain images can be used, for example, in the diagnostics of Alzheimer's disease or in research projects related to brain cancer. The research is being run by Philips of the Netherlands as part of the IVVES project.

One data set available on the Internet contains MRI images of approximately 350 Alzheimer's patients, and it has been widely utilised in the study of memory disorders. However, the development of AI methods, or neural networks, requires significantly more teaching material than a few hundred images. Pölönen uses this image bank to create a larger collection of artificial MRI images representing various memory disorders.

“For example, a synthetic data set of 10,000 images could already be used to train advanced neural networks that could be of assistance in diagnosing or monitoring Alzheimer's disease,” says Pölönen.

## From traditional image processing to neural network-based methods

“With the development of computing capacity and technology in recent years, neural networks can finally be used for doing all kind of useful things. Previously, machine learning algorithms were designed in detail based on the researcher's own analysis and imagination. Modern neural network-based algorithms learn independently the best possible model with a view to the end result from the teaching data entered into the computer,” says Pölönen.

“The neural network imitates the human brain in many ways. It contains a large number of memory locations with millions of neural connections of different strengths, just like the human brain,” says Pölönen. “Each neural network must be trained separately for the problem to be solved. Some networks are trained to identify a tumour or some other object, some to perhaps make diagnoses, and others to produce synthetic images,” he continues.

## A neural network learns its task in a couple of weeks

Generative adversarial network (GAN) is a popular neural network used for creating, for example, very authentic facial images. How is a neural network created with the help of a competing generative neural network?

Harri Pölönen simultaneously trains two neural networks to compete with each other: "To the first network, I feed the authentic MRI images as teaching material, from which it starts creating new images, which are as identical as possible without being fully identical. The second neural network, on the other hand, acts as a lie detector, trying to identify which MRI images are authentic and which synthetic ones created by the first network. In the case of 3D MRIs, the competition between the networks will take a couple of weeks until, hopefully, the lie detector network is no longer able to distinguish synthetic images created by its adversary from the authentic ones. The training of networks fails if the networks do not learn at the same pace. The lie detector must not be too perfect, as this means that the network creating synthetic images will never succeed in its task. The lie detector must not be too poor either, because then the neural network creating synthetic images becomes lazy, and the MRIs it creates do not look very real."

## More agile generation of synthetic data for corporate use

"Synthetic data has been used internationally for a few years, and it is suited for any kind of data. Brain images cannot be anonymised in any other way because every person has unique brain anatomy. Synthetic data can also be used to locate gaps and unevenness in data sets. It is also possible to artificially add deviations to the data that can be used to train or test the system," says Senior Researcher **Johan Plomp** from VTT.

Neural network experts work in a communal manner around the world, so the tools for building neural networks are freely available online in line with the open source principle. "The methods and technologies improve from week to week when development is carried out in a communal manner with the world's leading researchers and institutions. We also keep in touch and exchange ideas about the matter with other actors, including the top university MIT," says Pölönen.

"Over the past couple of years, we have been developing the creation of synthetic data. We have succeeded in improving the method, so the images have clearly improved in terms of resolution along the way. With modern computers, producing high-quality synthetic image data takes anything from days to weeks, but even that is much easier and faster than compiling authentic images and obtaining permits to use them," says Plomp. The aim is to bring more agility to making synthetic data available for corporate use. We still need a good operating model for this," says Plomp.

## IVVES - Verification and Validation of Evolving Systems project

The project is part of an extensive international [IVVES project](#), which is a cooperation project of the Eureka ITEA cluster. Its aim is to improve the reliability of adaptive industrial systems through various testing and quality assurance methods, such as solutions that improve the testability and transparency of machine learning systems, AI-based testing methods and data quality assurance methods. The project also develops development methods for machine learning solutions (MLOps) and methods for generating synthetic data.

“Trust is a very important factor that affects the introduction of machine learning methods, for example in the health sector, the banking world, the transport sector and the cyber future. Without verification methods, the enormous potential of artificial intelligence will remain unused,” says Plomp.

The IVVES project is coordinated by Philips of the Netherlands and involves 26 partners from five countries. The other Finnish participants in addition to VTT include F-Secure, Futurice, HeadAI, Solita, Techila and the University of Helsinki. In Finland, the project funding is obtained from Business Finland.

[FACEBOOK](#)
[TWITTER](#)
[LINKEDIN](#)
[REDDIT](#)
[TUMBLR](#)
[PINTEREST](#)
[0 LIKES](#)
[Previous](#)

Everyday choices: Ayush Bharti, can AI fix a dropped Wi-Fi connection?

[Next](#)

Neste and FCAI collaboration: Optimizing chemical reactors with AI

Finnish Center for Artificial Intelligence FCAI is a community of experts in artificial intelligence in Finland, initiated by Aalto University, University of Helsinki, and VTT Technical Research Centre of Finland. FCAI is one of the Academy of Finland's Finnish flagships, hubs of top-level research and impact.

[CONTACT](#)
[ABOUT](#)
[COLLABORATE](#)
[FAQ](#)
[RESEARCH](#)
[FOR RESEARCHERS](#)
[FOR MEDIA](#)
[IMPACT](#)
[FOR INDUSTRY](#)
[DIVERSITY](#)
[TIETOA MEISTÄ](#)
[PRIVACY](#)
[ACCESSIBILITY](#)

SUBSCRIBE TO FCAI  
NEWSLETTER

