

## Pressemitteilung

Universität Luxemburg - Université du Luxembourg

Raphaël Cayrol

19.12.2022

<http://idw-online.de/de/news806869>

Forschungsergebnisse, Wissenschaftliche Publikationen  
Chemie, Informationstechnik  
überregional



### Gaining unprecedented view of small molecules by machine learning

**A new tool to identify small molecules offers benefits for diagnostics, drug discovery and fundamental research. A new machine learning model will help scientists identify small molecules, with applications in medicine, drug discovery and environmental chemistry. Developed by researchers at Aalto University and the University of Luxembourg, the model was trained with data from dozens of laboratories to become one of the most accurate tools for identifying small molecules.**

Thousands of different small molecules, known as metabolites, transport energy and transmit cellular information throughout the human body. Because they are so small, metabolites are difficult to distinguish from each other in a blood sample analysis – but identifying these molecules is important to understand how exercise, nutrition, alcohol use and metabolic disorders affect wellbeing.

Metabolites are normally identified by analysing their mass and retention time with a separation technique called liquid chromatography followed by mass spectrometry. This technique first separates metabolites by running the sample through a column, which results in different flow rates – or retention times – through the measurement device. Mass spectrometry is then used to fine-tune the identification process by sorting metabolites according to their mass. Researchers can also break metabolites into smaller pieces to analyse their composition using a technique called tandem mass spectrometry.

“Even the best methods can’t identify more than 40% of the molecules in samples without making some additional assumptions about the candidate molecules,” says Professor Juho Rousu of Aalto University.

Now, Rousu’s group has developed a novel machine learning model to identify small molecules. It was recently published in Nature Machine Intelligence.

“This new open-source model offers the whole research community an enriched view of small molecules. It will help research into methods to identify metabolic disorders, such as diabetes, or even cancer,” says Rousu.

The new approach elegantly sidesteps one of the challenges facing conventional methods. Because the retention times of molecules vary from lab to lab, data cannot be compared between labs. Eric Bach, a doctoral student at Aalto, came up with an alternative during his PhD research that solved the problem.

“Our research shows that while absolute retention times may vary, the retention order is stable across measurements by different labs,” Bach explains. “This allowed us to merge all publicly available data on metabolites for the first time ever and feed it into our machine learning model.”

With the incorporation of data from dozens of laboratories around the globe, the machine learning model is accurate enough to distinguish between mirror image molecules, known as stereochemical variants. So far, identification tools

have not been able to tell stereochemical variants apart, and the new capability is expected to open up new avenues in drug design and other fields.

“The fact that using stereochemistry improved the identification performance is a revelation for all developers of metabolite identification methods,” says Emma Schymanski, Associate Professor at the Luxembourg Centre for Systems Biomedicine (LCSB) of the University of Luxembourg. “This method could also be used to help identify and trace micropollutants in the environment or characterise new metabolites in plant cells.”

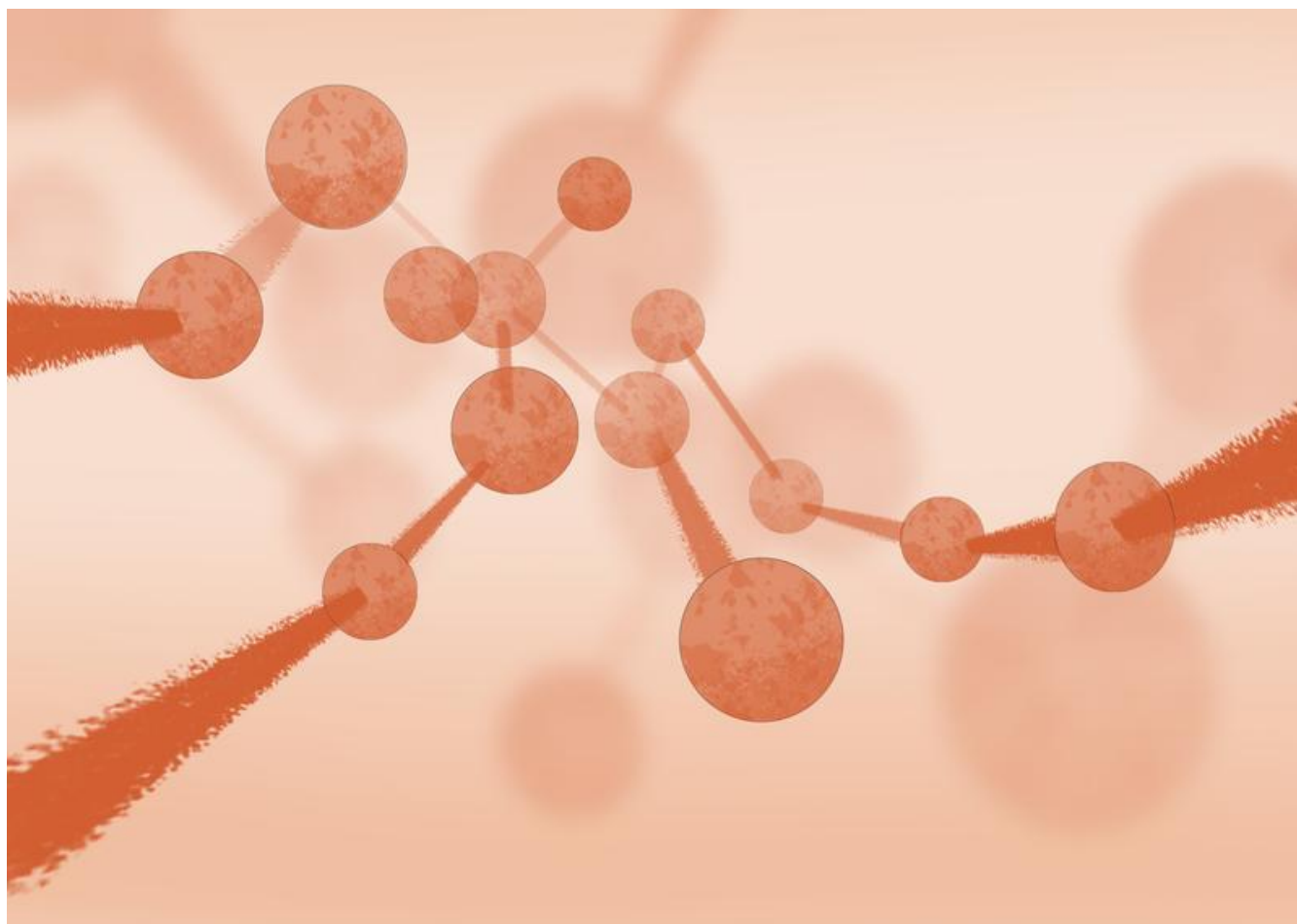
wissenschaftliche Ansprechpartner:

Prof. Juho Rousu - <https://www.aalto.fi/en/people/juho-rousu>

Associate Prof. Emma Schymanski - [https://www.uni.lu/lcsb/people/emma\\_schymanski](https://www.uni.lu/lcsb/people/emma_schymanski)

Originalpublikation:

<https://doi.org/10.1038/s42256-022-00577-2>



Small molecules - illustration  
Matti Ahlgren  
Aalto University

D