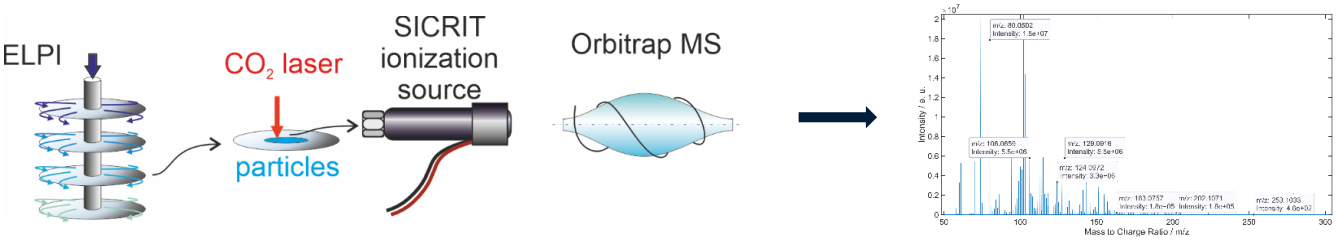
Lehrstuhl für Analytik
TUM School of Natural Sciences
Technische Universität München

Interdisciplinary project proposal:

Title: Development of a data pipeline for chemometric evaluation of high-resolution mass spectrometry data in environmental aerosol monitoring

Description:

The chemical composition of aerosol particles is a vital information when assessing the toxicity of an aerosol system. The current standard in organic analysis is high performance liquid chromatography (HPLC) coupled with mass spectrometry, which takes up to an hour per sample with an additional extraction time of 24 hours to prepare the samples. Our group developed a drastically faster method to desorb the analytes via Laser desorption and introduce them to high resolution mass spectrometry. Due to this advancement, a high throughput of samples becomes possible, which enables a continuous sampling and monitoring of the aerosol composition in the environment. This would be suited to help identify trends in aerosol compositions, recognize the occurrence of new toxic substances and monitor the levels of toxic particles.



The new situation with a high number of mass spectra, each with millions of datapoints and a proprietary data format necessitates us to explore possibilities for a data pipeline to make the information gathered more accessible to us, a group of scientists from varying backgrounds with foundational programming skills, mainly in Matlab and Python.

Your task in this project:

In order to develop a scientifically sound data pipeline, we already started with the theoretical aspects. Raw data parsing and conversion to mzML files is currently done via MSconvert, a C#-based open source program. Your role will first be the programming of a script to read, align and denoise the mzML data and convert the information to more user-friendly intensity-matrices with uniform x-axes. Following this, various approaches to chemometric data analysis have to be compared in their reliability for the use in PCA and similar classification techniques. One goal of our cooperation is to show you the application of your work and to further your understanding of the importance of statistics in the evaluation of scientific data. Concerning the programming languages for your project, we aim for a script that is read- and usable by the average PhD student in natural science and are open to suggestions from your side. An accessible coding structure, documentation, and output formats easily usable with Matlab/Python/R are key in achieving this. A GUI on the other hand is optional. We are always open to play to your strengths and addapt the scope of the internship.

Contact details:
Institut für Wasserchemie
Lasers and Particles Group
Haisch@tum.de

Lehrstuhl für Analytik
TUM School of Natural Sciences
Technische Universität München

Start:

Since the project is a constant part of our work, you can start anytime.

Location:

Our Team is located in Großhadern, München in Georg-Heberer-Straße 11. Office space can be provided, but working in home office might be easier for you.

Corresponding lectures:

Analytical chemistry offers lectures during the winter- and summer-semester. Possible topics include:

Current Research in the Instrumental Analysis of Trace Components 1 (CH3122/NAT0170)
Analytische Chemie (CH0107)
Fortgeschrittene analytische Verfahren (CH4115)

Organisation:

Simply contact Prof. Dr. Christoph Haisch via email (Haisch@tum.de),
or Felix Ludwig (Felix.Ludwig@tum.de)

Contact details:
Institut für Wasserchemie
Lasers and Particles Group
Haisch@tum.de