# Flexible Infrastructure for Mass Spectrometry Data

Laurent Gatto[1], Johannes Rainer[2], Sebastian Gibb[3]

[1] Computational Biology and Bioinformatics Group, de Duve Institute, UCLouvain
[2] Institute for Biomedicine, Eurac Research
[3] Department of Anaesthesiology and Intensive Care, University Medicine Greifswald

## Introduction

### Mass Spectrometry (MS) Data

- Paired measurements:
  - intensity (counts) vs *m/z* (mass-to-charge ratio).
  - intensity vs retention time.
- Additional various metatdata.

### Problems

- Vendor-specific, closed file formats.
- Various open-formats, e.g. mzML, mzXML, mz5, CDF, MGF, CSV
- Spectra libraries/Annotation information in databases, e.g. SQL
- Multiple R-packages require different file formats and/or data structures.

### Solution

`Spectra` offers a common, flexible data structure and interface.

## Data Structure

| class | storage | writeable |
|---|---|---|
| MsBackendDataFrame[1] | in-memory | yes |
| MsBackendHdf5Peaks[1] | on-disk | yes |
| MsBackendMzR[1] | on-disk | no |
| MsBackendRawFileReader[2] | on-disk | no |

## Interface

1

2

## Contact

rformassspectrometry/Spectra

## References

1. Gatto, L., Rainer, J. & Gibb, S. *Spectra: Spectra infrastructure for mass spectrometry data*.

2. Panse, C. & Kockmann, T. *MsBackendRawFileReader: Bridging spectra and thermofinnigan raw files*. (2019).