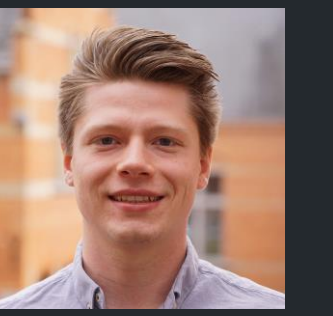




Search or jump to...

Ralf Gabriels (@RalfG)



HUPO-PSI / SpectralLibraryFormat

Checkout

Discuss

Contribute

Code

Issues 4

Pull requests 0

Wiki

Insights

Spectral Library Format: A standard format to exchange/distribute spectral libraries/archives

hupo-psi

standards

file-format

mass-spectrometry

proteomics

big-data

78 commits

1 branch

0 packages

0 releases

16 contributors

Apache-2.0

README.md

Introduction

Spectral libraries allow us to boost peptide identification workflows by extracting knowledge embedded in previously identified spectra. However, these libraries are provided in many different file formats and software tools usually require very specific input formats as well. This drastically hinders the usability of spectral libraries. The solution is a standardized, but flexible spectral library file format.

A flexible standard file format

Unified framework

4 levels of (meta)data
Library → Spectrum → Peak → Peak interpretation
Controlled vocabulary with metadata
Part of the PSI-MS ontology
All items defined by
Accession, name, value_accession, value, cv_param_group

Multiple representations

Text based → Human readable
JSON → Perfect for APIs
CSV, TSV → Easily read
HDF → Fast and efficient storage
...

Some examples

Metadata snippet of text-based format:

```
MS:1008014|spectrum index=500
MS:1008013|spectrum name=AAAVDPTPAAPAR/2_0
MS:1008010|molecular mass=1208.6510
[1]MS:1008030|number of enzymatic termini=2
[1]MS:1001045|cleavage agent name=MS:1001251|Trypsin
MS:1001471|peptide modification details=0
...
```

JSON metadata item:

```
{
  "accession": "MS:1001045",
  "cv_param_group": "1",
  "name": "cleavage agent name",
  "value": "Trypsin",
  "value_accession": "MS:1001251"
},
```

HDF group structure:

```
└─ library (group)
   └─ spectrum_001 (group)
   └─ spectrum_002 (group)
   └─ ...
   └─ spectrum_n (group)
      └─ intensity (array)
      └─ mz (array)
```

Contributing

Join the discussions and development on github.com/HUPO-PSI/SpectralLibraryFormat!

The HUPO PSI-MS team: Nuno Bandeira, Wout Bittremieux, Jeremy Carver, Matthew Chambers, Eric Deutsch, Ralf Gabriels, Shin Kawano, Henry Lam, Tytus Mak, Yasset Perez-Riverol, Ben Pullman, Vagisha Sharma, Jim Shofstahl, Tim Van Den Bossche, Juan Antonio Vizcaino, and Yunping Zhu.

