

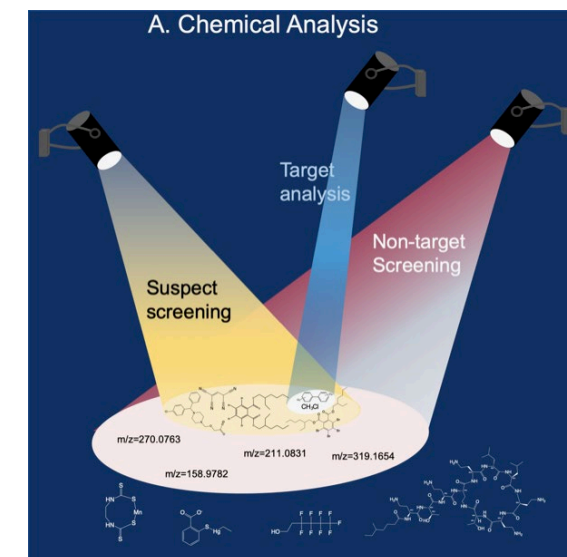
Chemical contamination of the sea – (management of the knowns and) Research on the unknowns

Assoc. Prof. Dr. Emma L. Schymanski
(plus many, many colleagues and collaborators!)

Environmental Cheminformatics Group,
Luxembourg Centre for Systems Biomedicine, University of Luxembourg

Email: emma.schymanski@uni.lu and [@ESchymanski](https://twitter.com/ESchymanski)

Web: https://www.uni.lu/lcsb/research/environmental_cheminformatics/

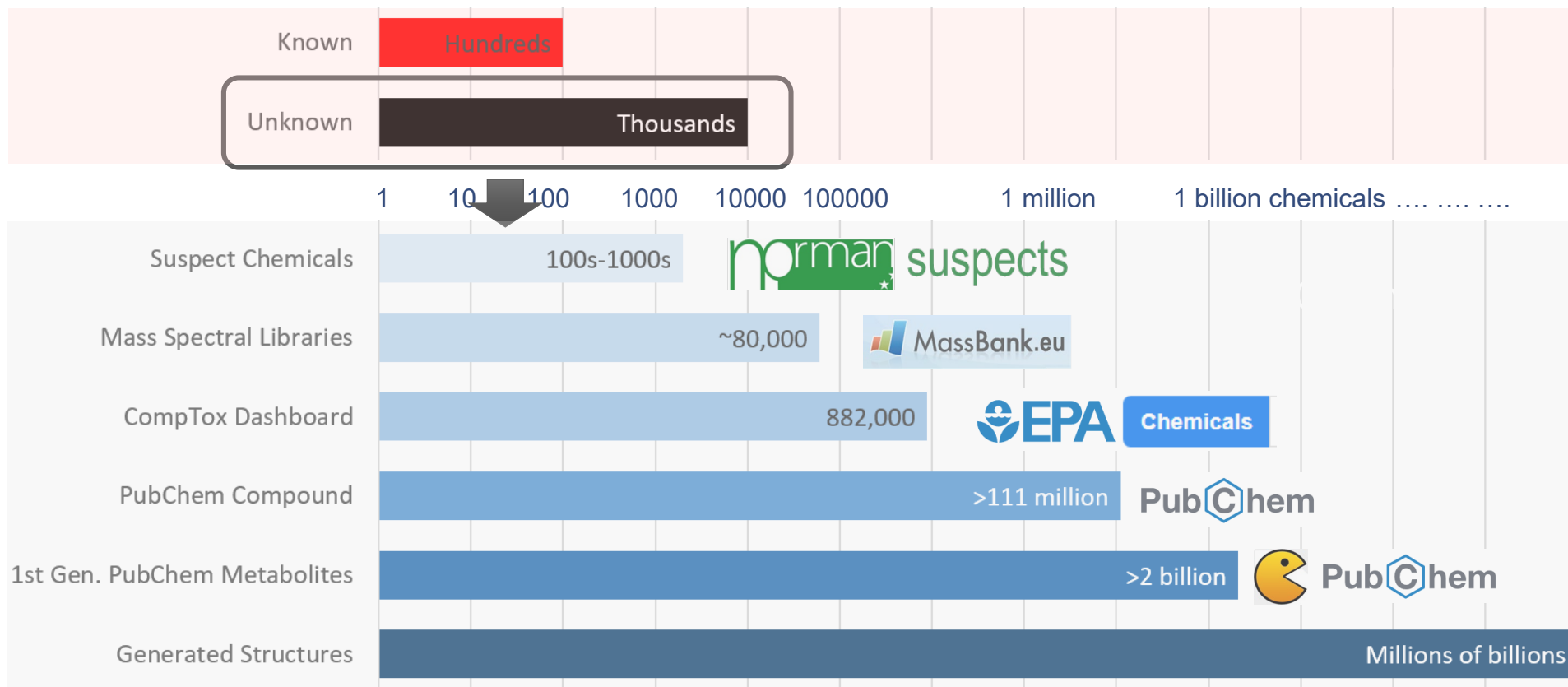
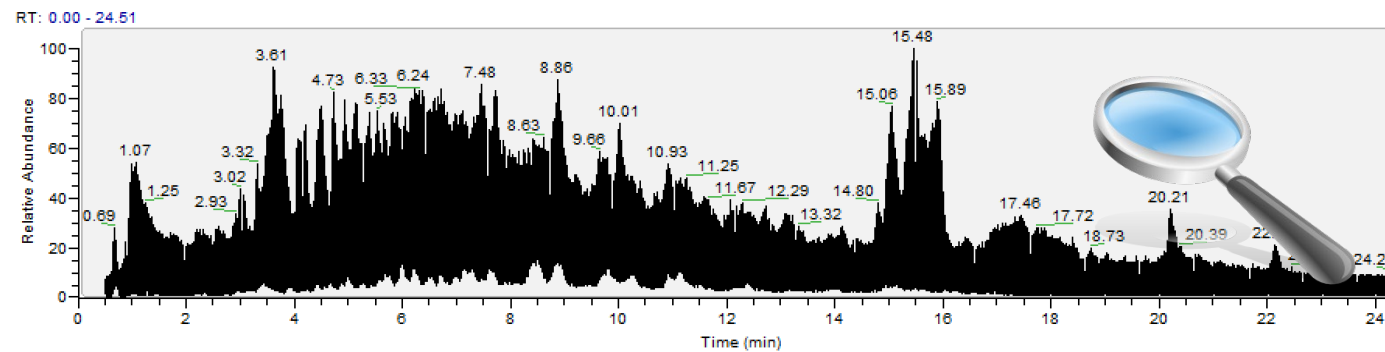


Mod. from Escher et al.
DOI: [10.1126/science.aay6636](https://doi.org/10.1126/science.aay6636)

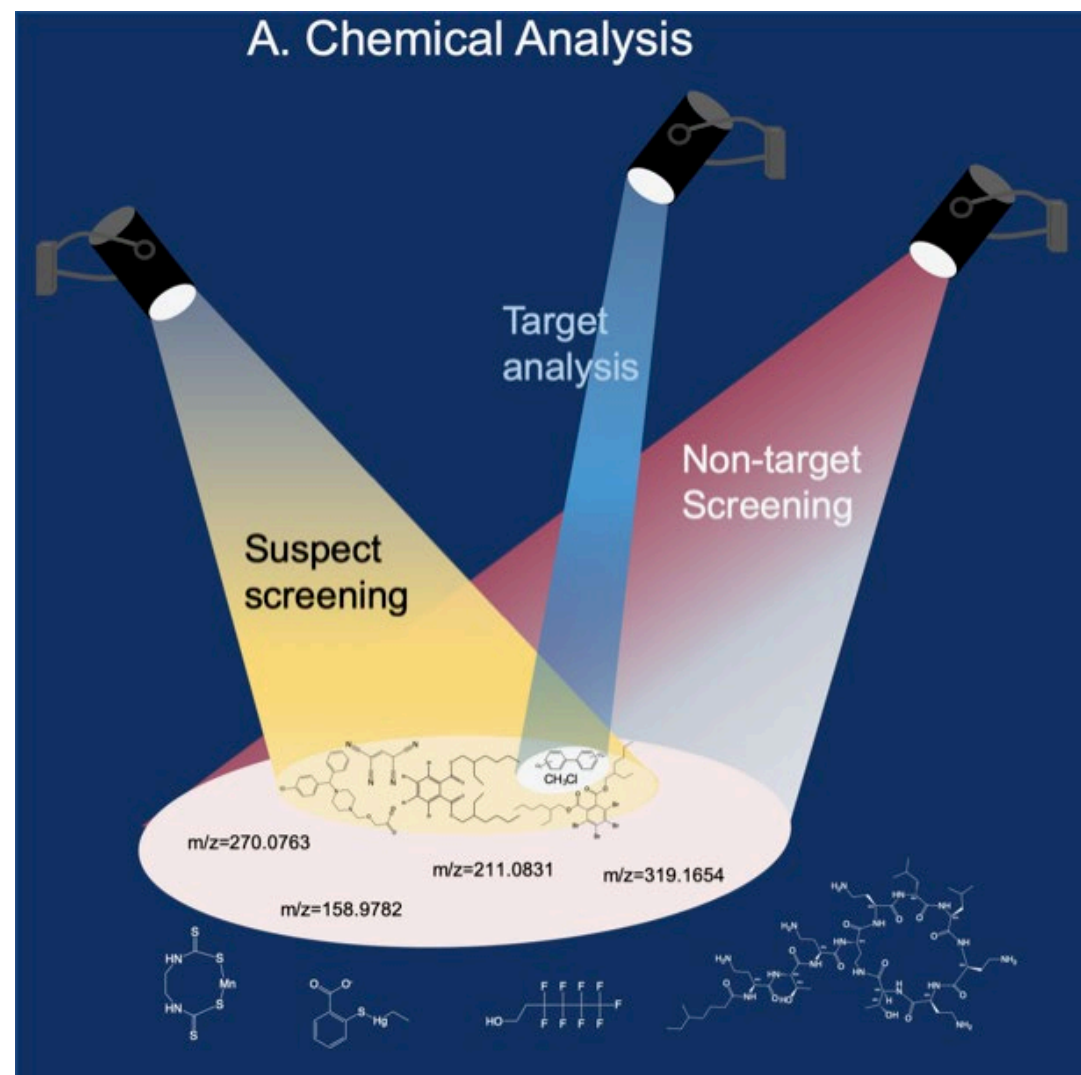
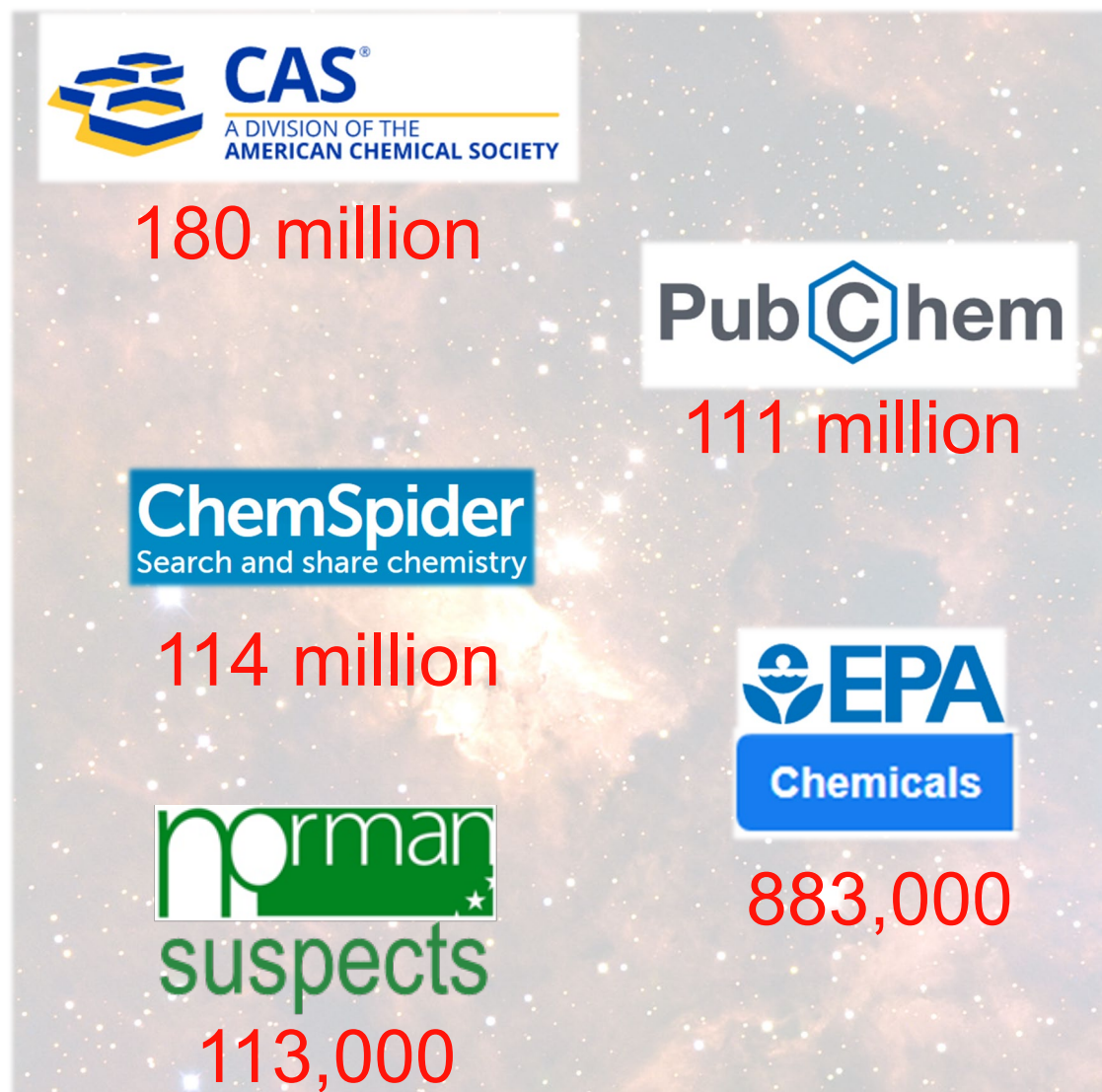
Environmental Cheminformatics & High Resolution Mass Spec

High resolution
mass spectrometry

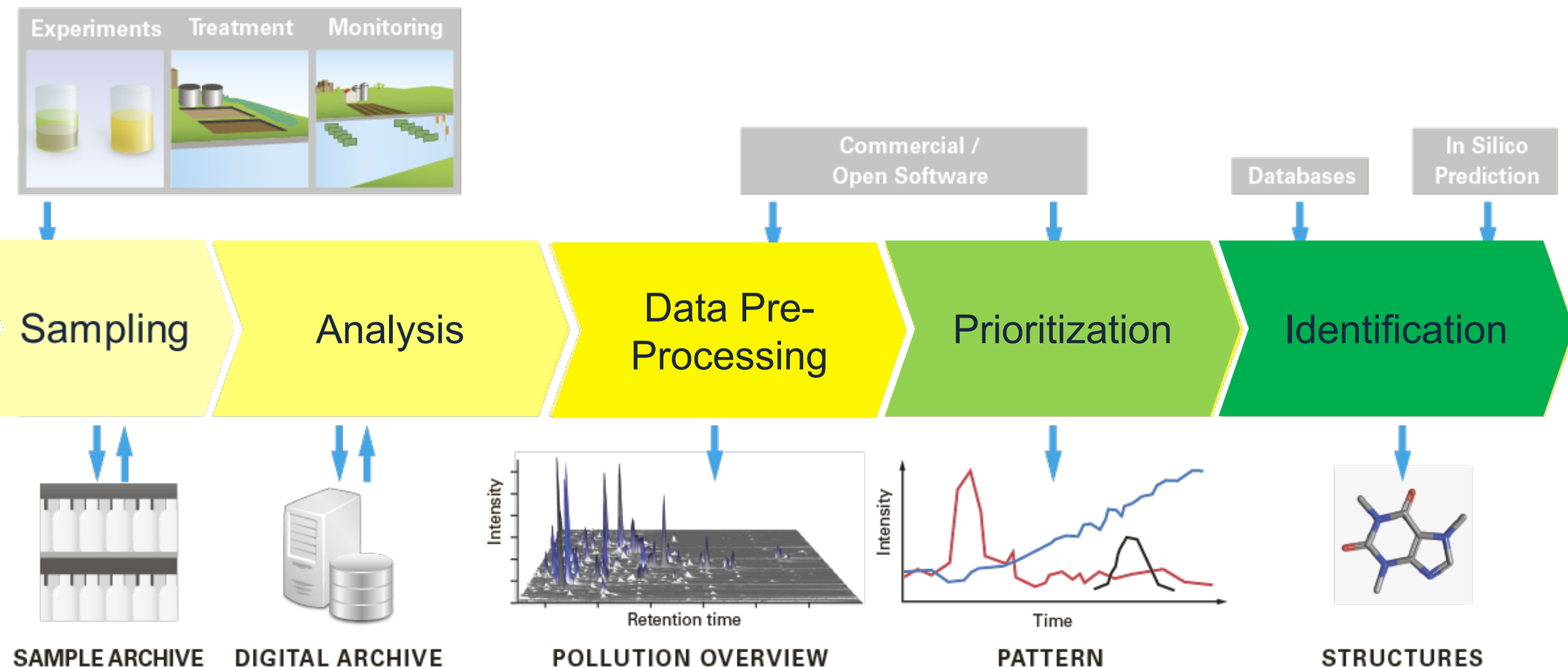
AND connecting
chemical knowledge



The Problem: Which chemicals are relevant? How to find them?



Non-target High Resolution Mass Spectrometry (NT-HRMS)

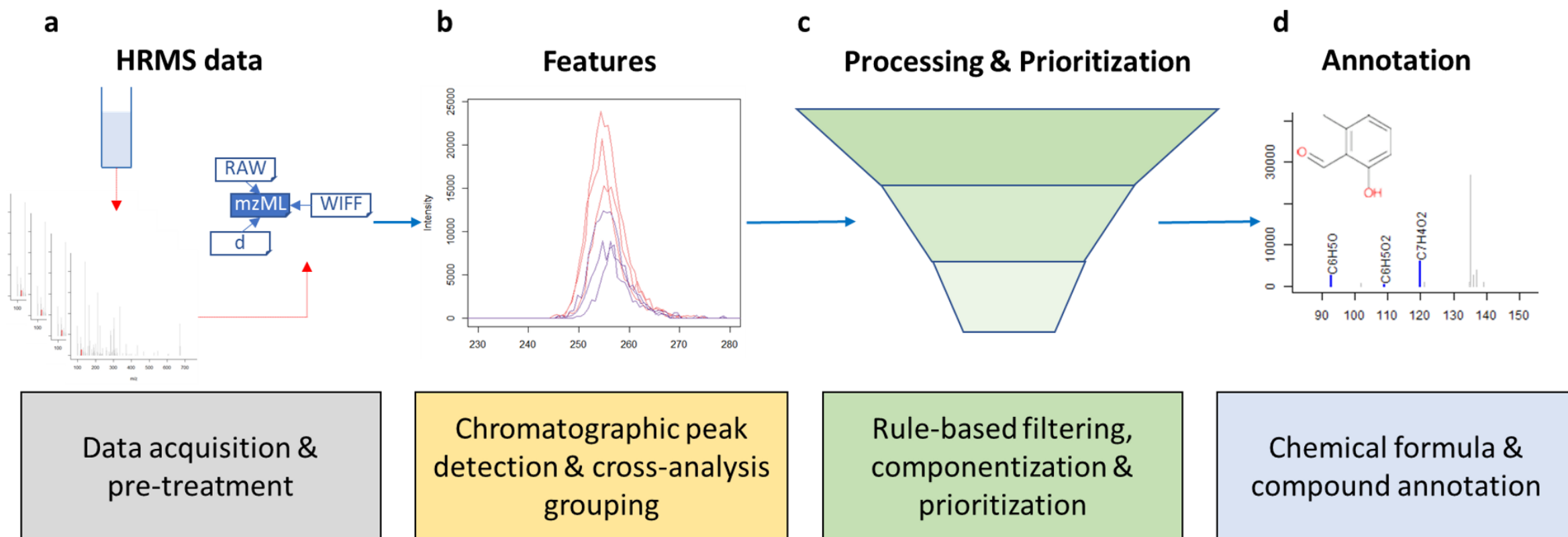


Open Source Workflows for NT-HRMS: patRoon

<https://rickhelmus.github.io/patRoon/>



Rick
Helmus



UNIVERSITY OF AMSTERDAM

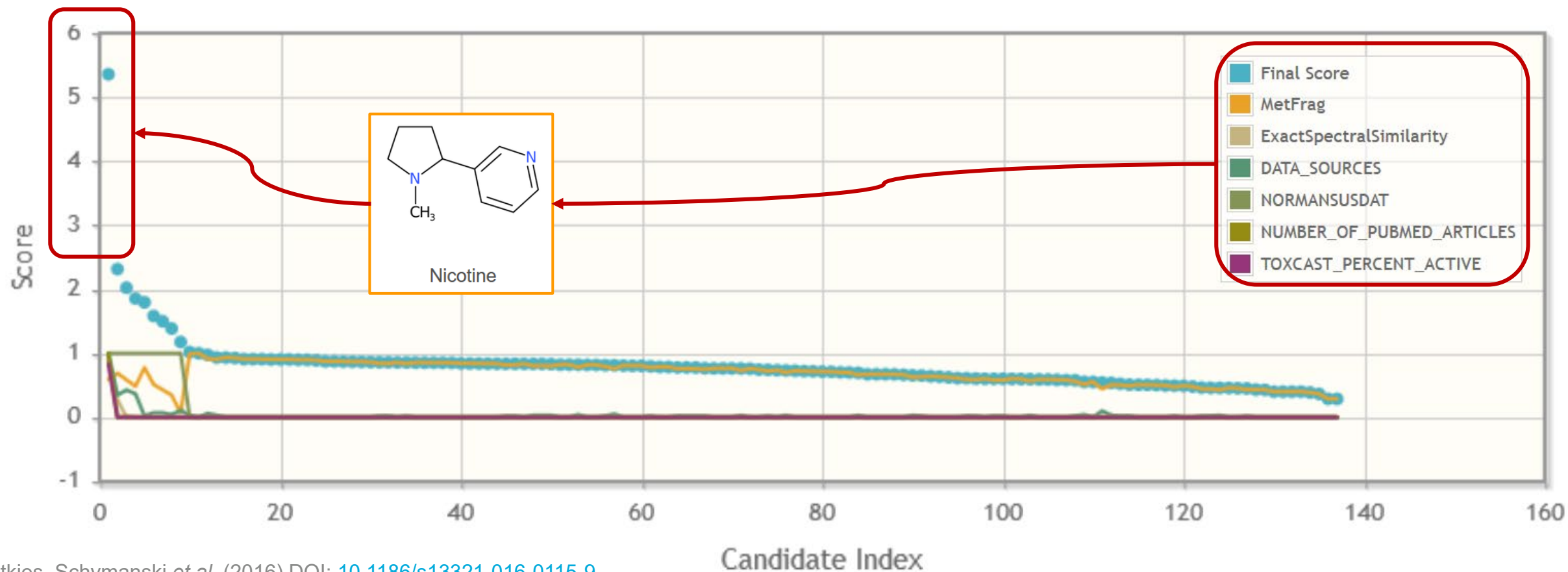


MetFrag: Connecting Knowledge for Chemical Identification



<https://msbi.ipb-halle.de/MetFrag/>

<http://ipb-halle.github.io/MetFrag/>



Problem: “Chemical Space” is too big!



180 million



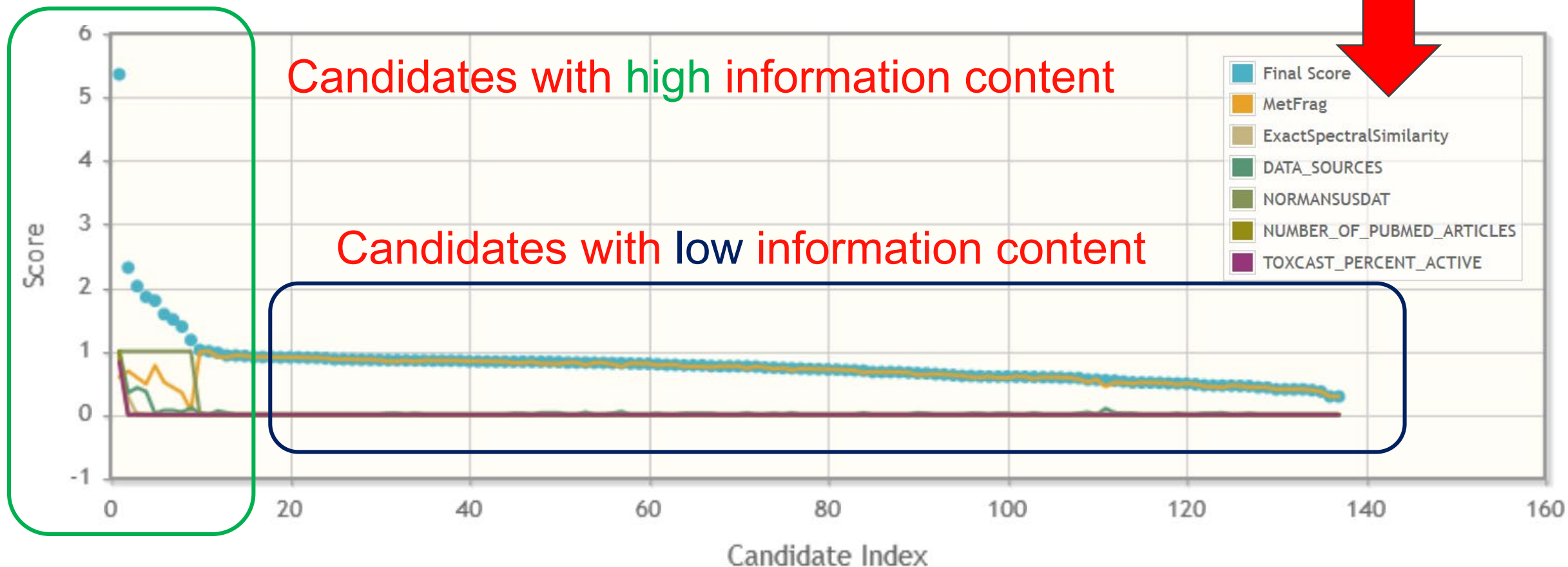
111 million



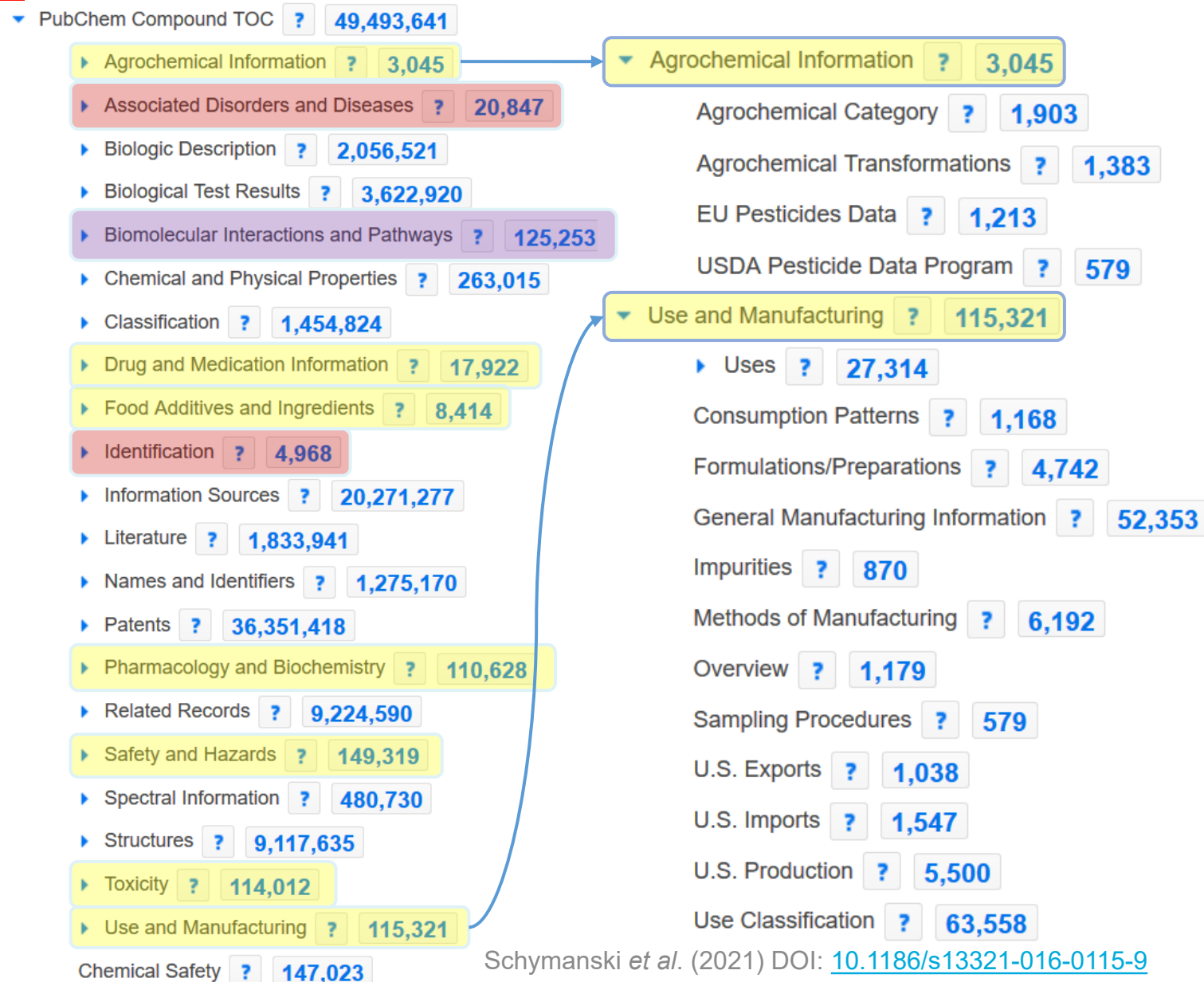
114 million



883,000



Solution: Pick out the most relevant pieces - e.g. PubChem



Schymanski et al. (2021) DOI: [10.1186/s13321-016-0115-9](https://doi.org/10.1186/s13321-016-0115-9)

PubChem Furathiocarb (Compound)

CONTENTS



7 Agrochemical Information



7.1 Agrochemical Category



Insecticides

▶ EU Pesticides Database

7.2 Agrochemical Transformations



Furathiocarb has known environmental transformation products that include [carbofuran](#).

S60 | SWISSPEST19 | Swiss Pesticides and Metabolites from Kiefer et al 2019 | DOI:10.5281/zenodo.3544759

▶ NORMAN Suspect List Exchange

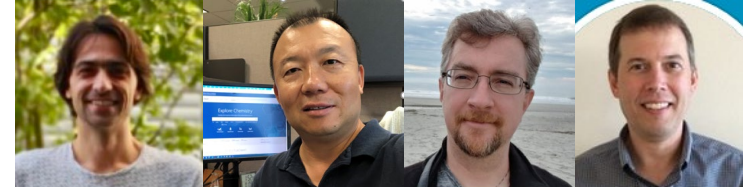
7.3 EU Pesticides Data



Active Substance	furathiocarb
Status	Not Approved [Reg. (EC) No 1107/2009]
Categories	Insecticides

<https://pubchem.ncbi.nlm.nih.gov/compound/Furathiocarb#section=Agrochemical-Information>

Introducing PubChemLite



Portable, dynamic and highly relevant!

▼ PubChem Compound TOC ? 49,493,641

▶ Agrochemical Information ? 3,045

▶ Associated Disorders and Diseases ? 20,847

▶ Biologic Description ? 2,056,521

▶ Biological Test Results ? 3,622,920

▶ Biomolecular Interactions and Pathways ? 125,253

▶ Chemical and Physical Properties ? 263,015

▶ Classification ? 1,454,824

▶ Drug and Medication Information ? 17,922

▶ Food Additives and Ingredients ? 8,414

▶ Identification ? 4,968

▶ Information Sources ? 20,271,277

▶ Literature ? 1,833,941

▶ Names and Identifiers ? 1,275,170

▶ Patents ? 36,351,418

▶ Pharmacology and Biochemistry ? 110,628

▶ Related Records ? 9,224,590

▶ Safety and Hazards ? 149,319

▶ Spectral Information ? 480,730

▶ Structures ? 9,117,635

▶ Toxicity ? 114,012

▶ Use and Manufacturing ? 115,321

Chemical Safety ? 147,023

PubChemLite
EXPOSOMICS
~370,000 entries “small”

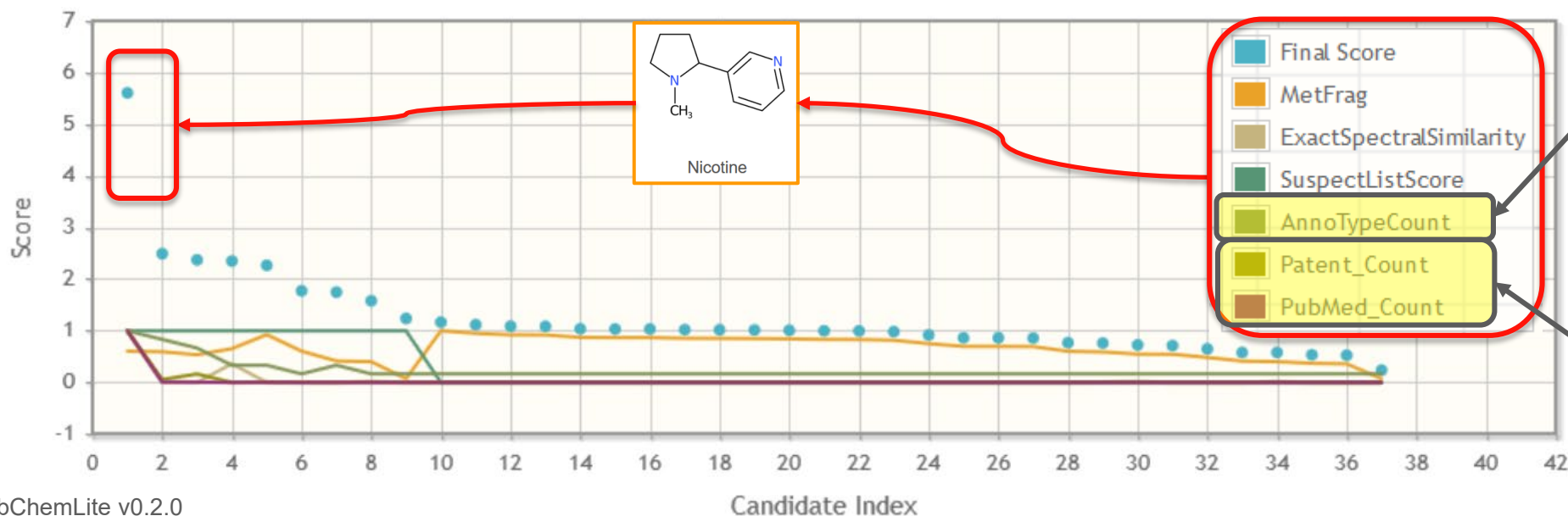
Schymanski *et al.* (2021) DOI: [10.1186/s13321-016-0115-9](https://doi.org/10.1186/s13321-016-0115-9)

PubChemLite: Fewer and more relevant candidates – with context



Statistics

Candidate Score Distribution



PubChem Compound TOC	49,493,641
Agrochemical Information	3,045
Associated Disorders and Diseases	20,847
Biologic Description	2,056,521
Biological Test Results	3,622,920
Biomolecular Interactions and Pathways	125,253
Chemical and Physical Properties	263,015
Classification	1,454,824
Drug and Medication Information	17,922
Food Additives and Ingredients	8,414
Identification	4,968
Information Sources	20,271,277
Literature	1,833,941
Names and Identifiers	1,275,170
Patents	36,351,418
Pharmacology and Biochemistry	110,628
Related Records	9,224,590
Safety and Hazards	149,319
Spectral Information	480,730
Structures	9,117,635
Toxicity	114,012
Use and Manufacturing	115,321
Chemical Safety	147,023



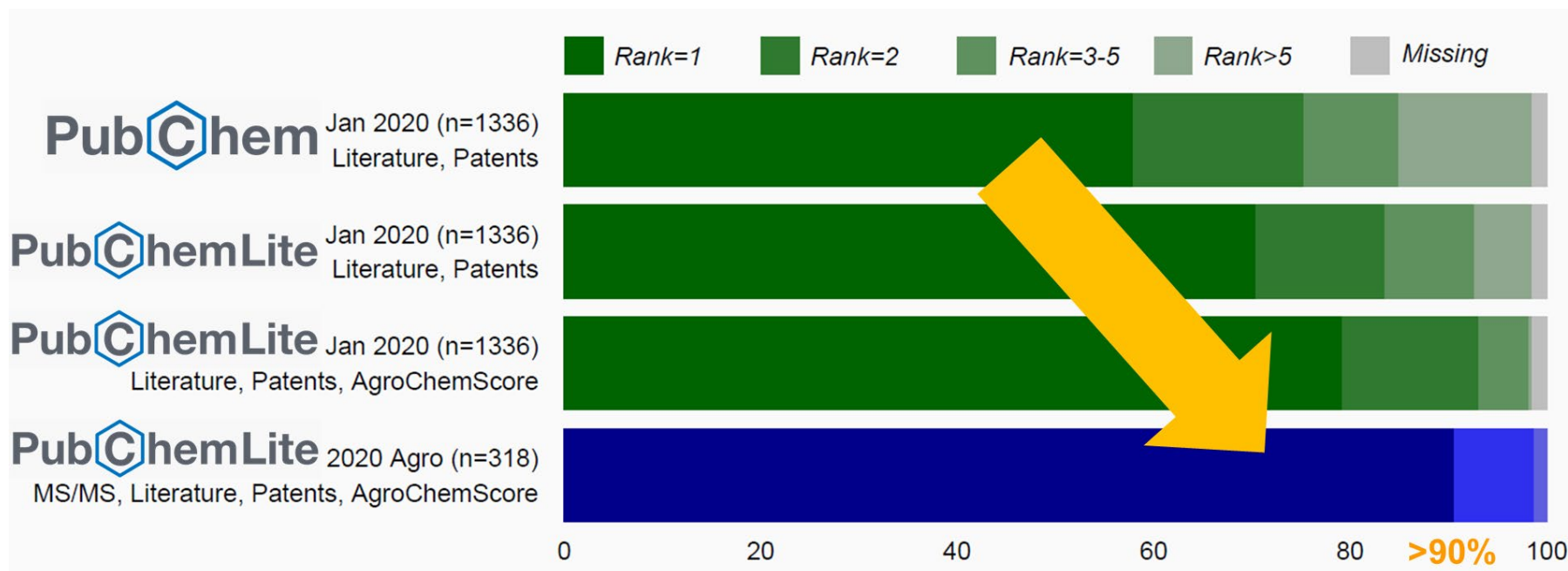
PubChemLite v0.2.0

Candidate Index

MetFrag+PubChemLite+Formula+MoNA+SusDat+Pat+Refs+Anno + <https://massbank.eu/MassBank/RecordDisplay?id=EQ300804>



Influence of the Annotation Content in PubChemLite



>90% of (well known) entries are Top 1 rank

>97% are Top 2 ... 100 % Top 3 ...

***Non-target screening for “known unknowns”
now yields quick answers!!***

Putting PubChemLite into Practice?



- Downloadable & fully integrated in patRoön, MetFragWeb & MetFrag CL!

<https://rickhelmus.github.io/patRoön/>

zenodo

October 31, 2020

PubChemLite for Exposomics

Bolton, Evan; Schymanski, Emma; Kondic, Todor; Thiessen, Paul; Zhang, Jeff

PubChemLite is a subset of PubChem (<https://pubchem.ncbi.nlm.nih.gov/>) selected from major categories of the Table of Contents page at the PubChem Classification Browser

MetFrag

There is now just one associated Disorders &

from 10 categories: SafetyInfo, ToxicityInfo,

structure from the most s, disconnected

Patent and PubMed ID "int" term counts how category) counts the C entry.

Dataset Open Access

Edit

New version

Communities

LCSB Environmental Cheminformatics Group

1,409 views

1,515 downloads

See more details...

patRoön 1.2.0

Reference Tutorial Handbook Changelog

Installation

patRoön itself can be installed as any other R package, however, depending on which algorithms you want to use in your workflow, some extra steps may be required to install all the necessary tools. Please see the [installation section in the handbook](#) for more information.

Getting started

For a very quick start:

```
library(patRoön)
newProject()
```

The `newProject()` function will pop-up a dialog screen (requires [R Studio!](#)) which will allow you to quickly select the analyses and common workflow options to subsequently generate a template R processing script.

However, for a better guide to get started it is recommended to read the [tutorial](#). Afterwards the [handbook](#) is a recommended read if you want to know more about advanced usage of patRoön. Finally, the [reference](#) outlines all the details of the patRoön package.

<https://github.com/rickhelmus/patRoön/issues>

License

GPL-3

Citation

Citing patRoön

Developers

Rick Helmus

Author, maintainer

All authors...

Dev status

PASSED

build passing

codecov 82%

image size 2.17 GB

DOI [10.5281/zenodo.3548653](https://doi.org/10.5281/zenodo.3548653)

<https://ipb-halle.github.io/MetFrag/>

MetFragWeb

Learn more

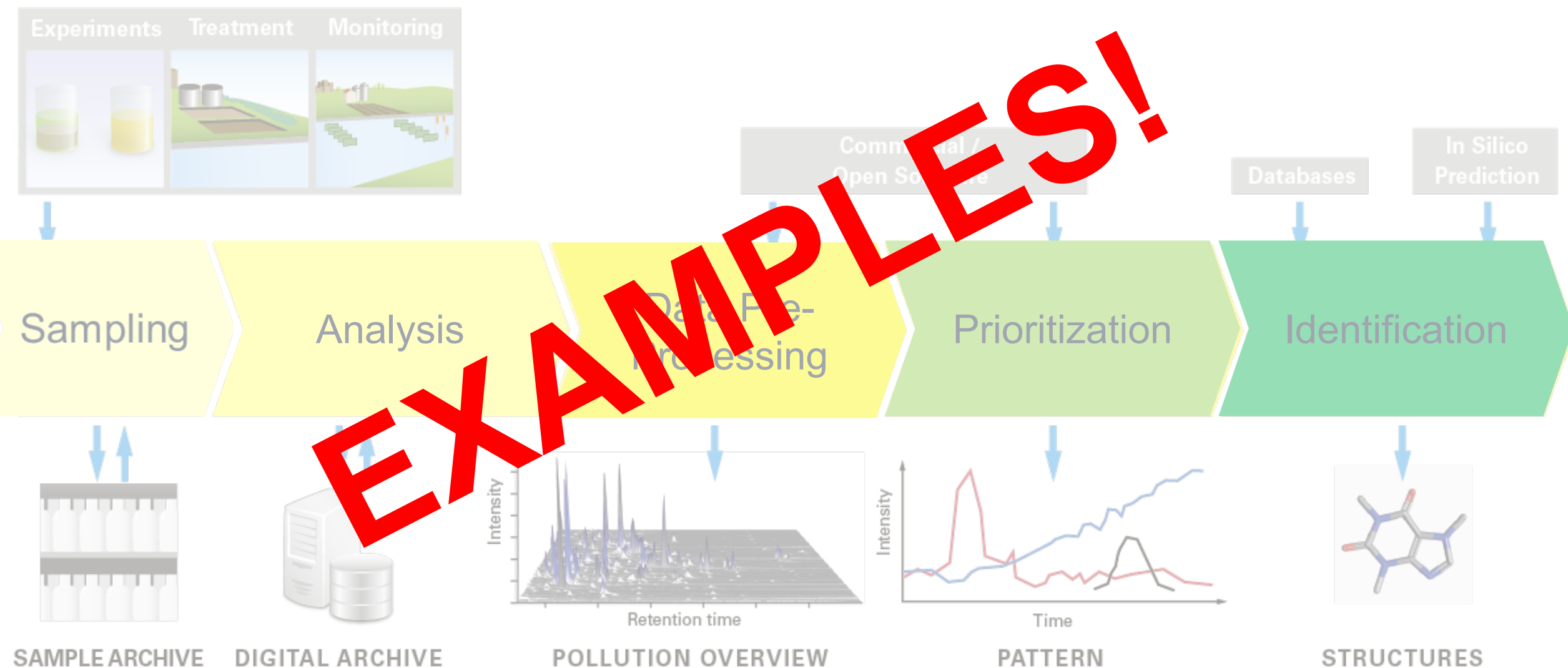
Use MetFrag directly from your browser which makes it pretty easy for beginners

MetFrag CL

Learn more

The new MetFrag commandline tool providing additional scoring terms for scoring candidates with MS/MS spectra

Non-target High Resolution Mass Spectrometry (NT-HRMS)

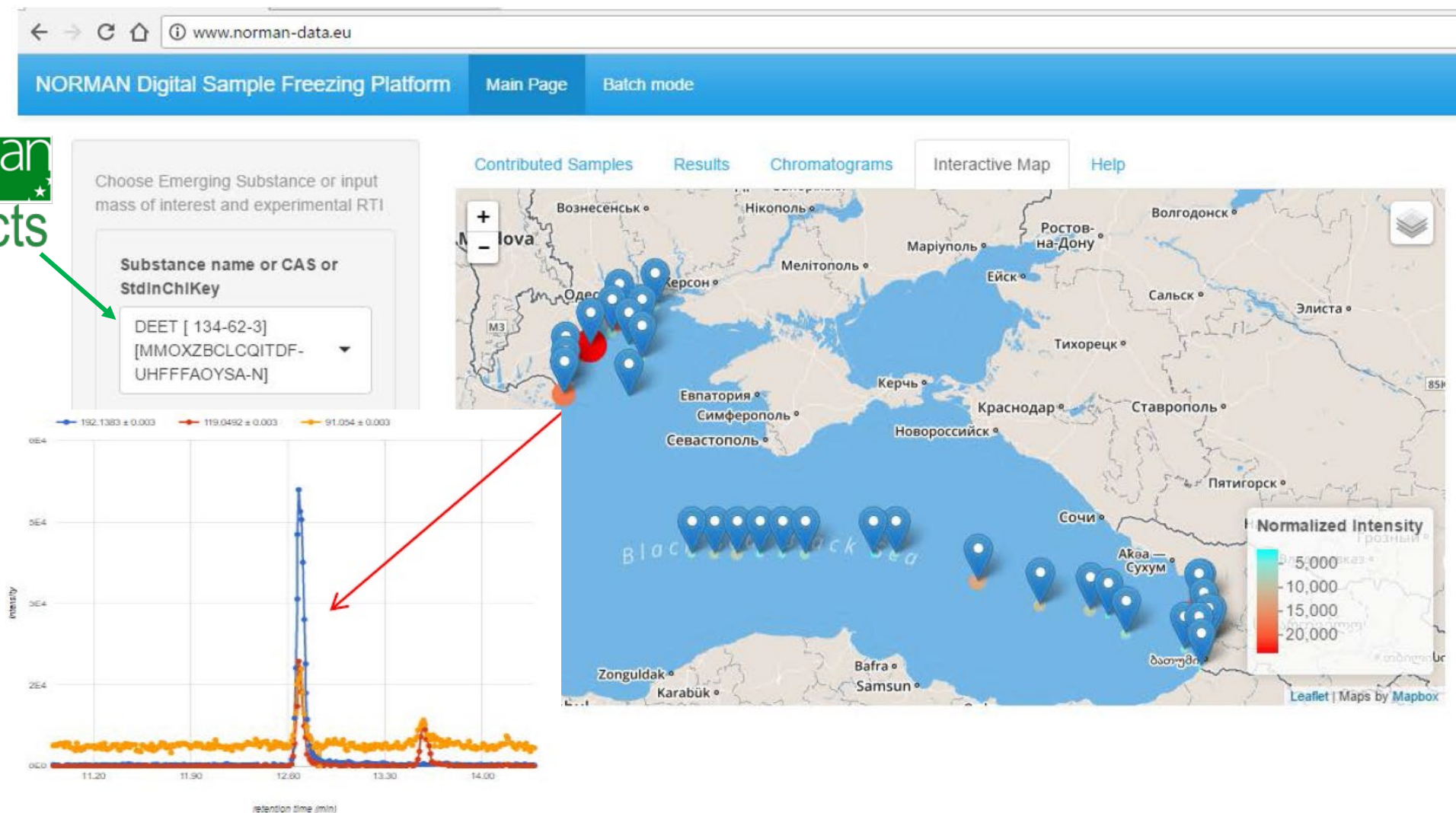


Examples: NORMAN Digital Sample Freezing Platform

“Live” retrospective screening of known and unknown chemicals in European samples

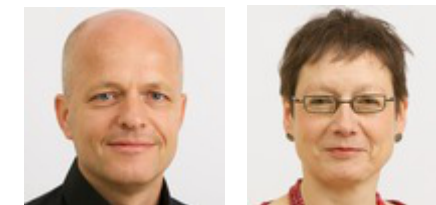


norman
suspects

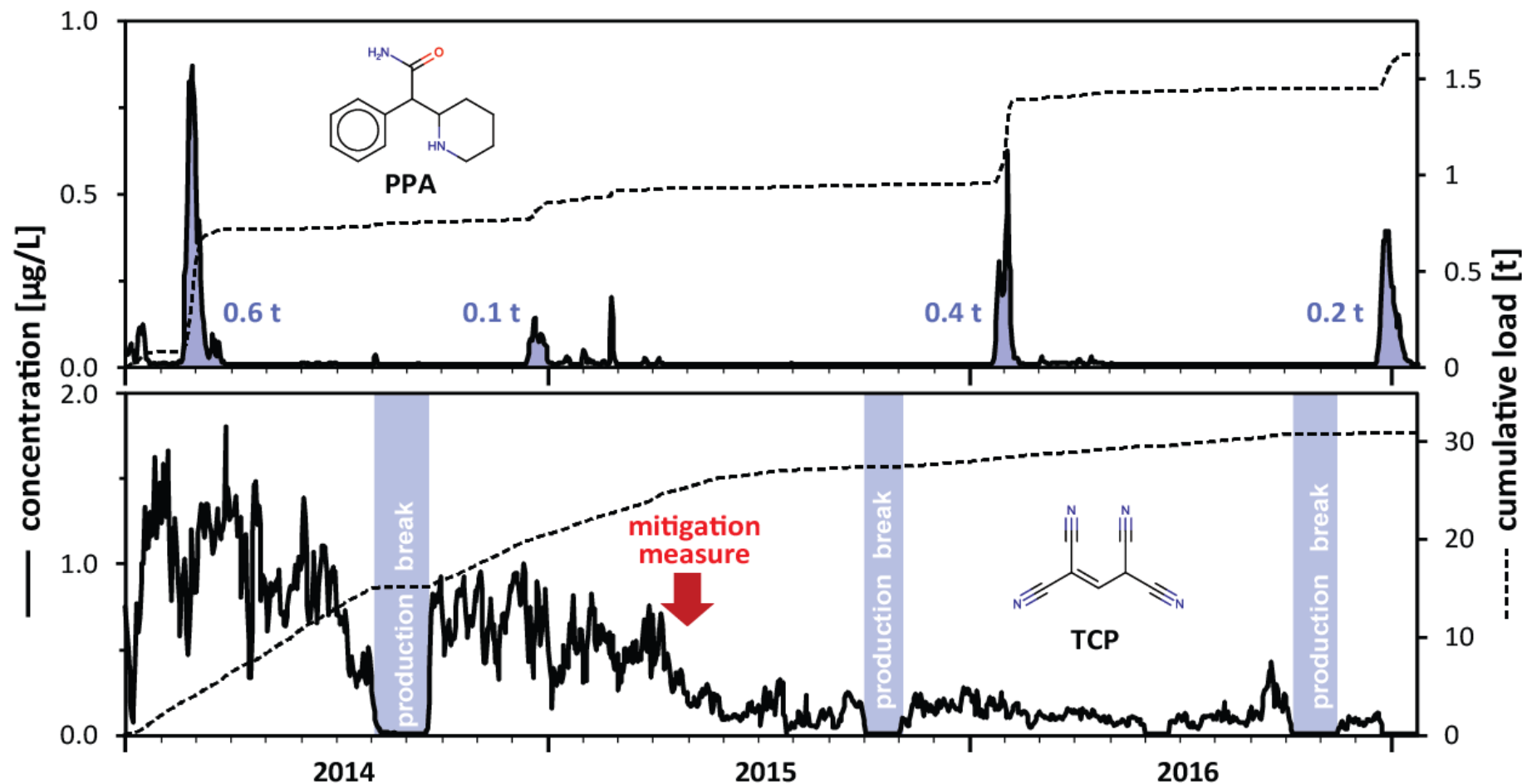


Examples: Real Time Monitoring of the Rhine River

Previously unknown chemicals detected due to “stand-out” patterns



eawag
aquatic research



Thank you!

Email: emma.schymanski@uni.lu

Twitter: [@ESchymanski](https://twitter.com/ESchymanski)

Slides @ DOI: [10.5281/zenodo.5576637](https://doi.org/10.5281/zenodo.5576637)

[https://www.uni.lu/lcsb/research/
environmental_cheminformatics/](https://www.uni.lu/lcsb/research/environmental_cheminformatics/)

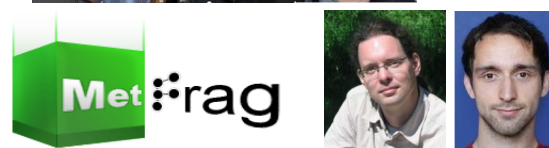
PubChemLite
EXPOSOMICS

ECI@LCSB



PubChem

... and
team



NIH U.S. National Library of Medicine
National Center for Biotechnology Information



Luxembourg National
Research Fund

