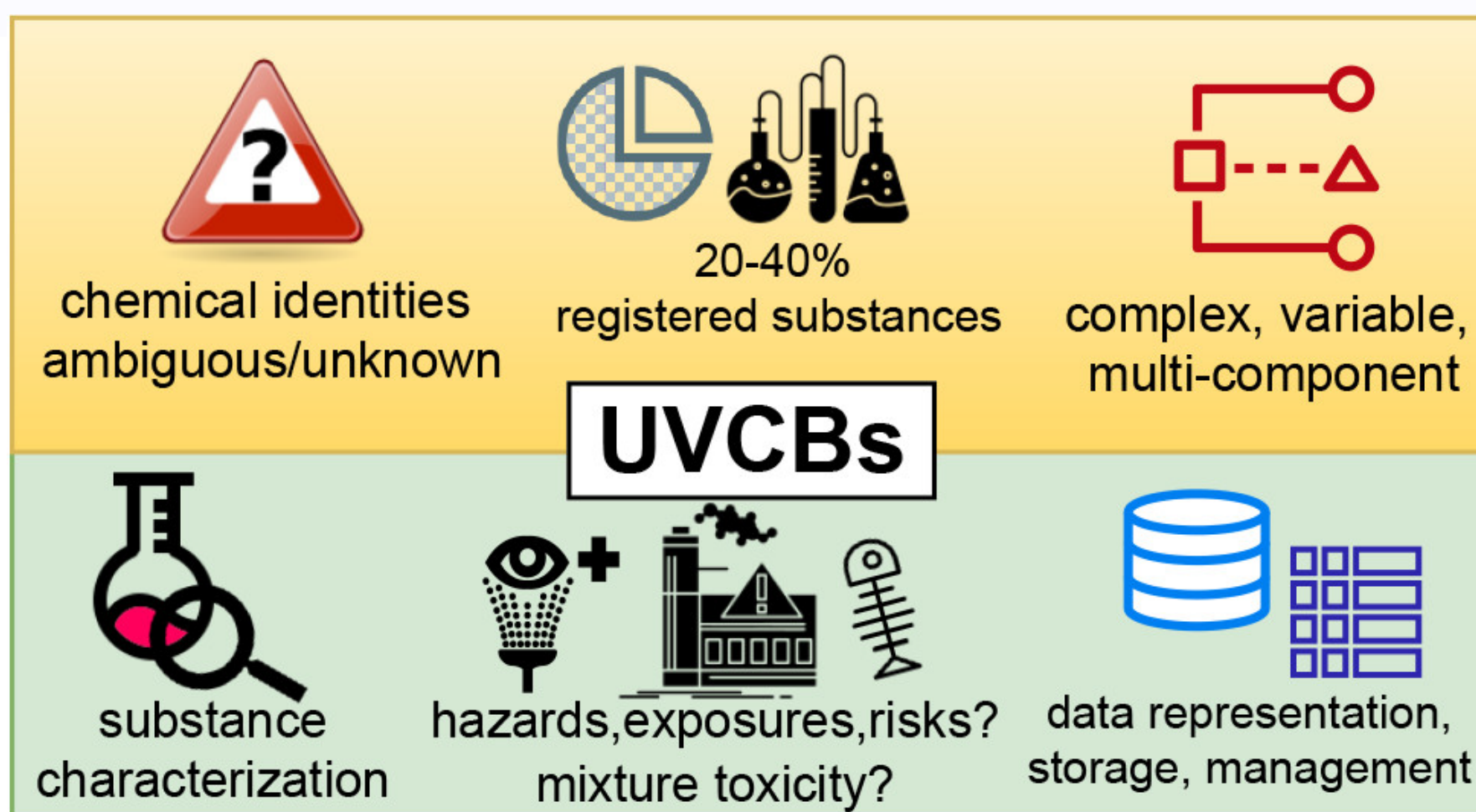


Integrating UVCBs and Related Data into Open Chemical Knowledgebases – PubChem and NORMAN-SLE Examples

Emma L. Schymanski¹, Anjana Elapavalore¹, Qingliang Li², Paul A. Thiessen², Leonid Zaslavsky², Jian Zhang², Evan E. Bolton². Contact: emma.schymanski@uni.lu
¹LCSB, University of Luxembourg. ²NCBI/NLM, National Institutes of Health, USA.



UVCBs are ...



PubChem is ...

- A large open chemical database
- >115 million compounds
- >300 million substances
- >930 data sources
- ...and much, much more

<https://pubchem.ncbi.nlm.nih.gov/>

NORMAN is ...

- A worldwide network on emerging substances of concern, based in Europe
- UVCBs a priority topic since Dec. 2022
- Action via NORMAN Suspect List Exchange
- ...first data @ DOI: [10.5281/zenodo.7414795](https://doi.org/10.5281/zenodo.7414795)

<https://www.norman-network.com/>

Lai *et al.* (2022) DOI: [10.1021/acs.est.2c00321](https://doi.org/10.1021/acs.est.2c00321); Kim *et al.* (2023) DOI: [10.1093/nar/gkac956](https://doi.org/10.1093/nar/gkac956); Mohammed Taha *et al.* (2022) DOI: [10.1186/s12302-022-00680-6](https://doi.org/10.1186/s12302-022-00680-6)

Examples of HESI UVCBs (and associated data) in PubChem

Perfluoro compounds, C5-18 (Compound)

9 Toxicity

9.1 Toxicological Information

9.1.1 Adverse Effects

Occupational hepatotoxin - Secondary hepatotoxins: the potential for toxic effect in the occupational setting is based on cases of poisoning by human ingestion or animal experimentation.

Haz-Map, Information on Hazardous Chemicals and Occupational Diseases

Source	Haz-Map, Information on Hazardous Chemicals and Occupational Diseases
Record Name	Perfluoro compounds, C5-18
URL	https://haz-map.com/Agents/6594

Quaternary ammonium compounds, benzyl-... (Compound)

6.1.2 Hazard Classes and Categories

Acute Tox. 4 (96.62%)
Acute Tox. 4 (17.57%)
Skin Corr. 1B (90.54%)
Eye Dam. 1 (27.82%)
Aquatic Acute 1 (81.31%)
Aquatic Chronic 1 (22.64%)

European Chemicals Agency (ECHA)

Acute toxicity - category 4
Acute toxicity - category 4
Skin corrosion - category 1B

Hazardous Chemical Information System (HCIS), Safe Work Australia

View these records online!



Amines, C12-14-tert-alkyl (Compound)

4.2 U.S. Production

Aggregated Product Volume

2019: 1,000,000 lb - <20,000,000 lb
2018: 1,000,000 lb - <20,000,000 lb
2017: 1,000,000 lb - <20,000,000 lb
2016: 1,000,000 lb - <20,000,000 lb

<https://www.epa.gov/chemical-data-reporting>

EPA Chemicals under the TSCA

4.3 General Manufacturing Information

Industry Processing Sectors

All Other Basic Organic Chemical Manufacturing

EPA Chemicals under the TSCA

<https://pubchem.ncbi.nlm.nih.gov/compound/Perfluoro%20compounds,%20C5-18#section=Toxicity>

All data is also accessible programmatically

Example UVCBs – PubChem / NORMAN-SLE Efforts

Carboxylic acids, di-, C4-11

Related Records

UVCB Carboxylic acids, di-, C4-11

UVCB Component	Type	Component Structure	Component CID	Reference	Source	Comment
Hexanedioic acid	Component		196	DOI:10.5281/zenodo.7883657	DOI:10.5281/zenodo.7414795	Shortest chain length with annotation content
Nonanedioic acid	Component		2266			
Octanedioic acid	Component		10457	DOI:10.5281/zenodo.7883657		

OngLai: An Algorithm to Classify Homologous Series

PubChemLite EXPOSOMICS BIOHACKATHON EUROPE

Polychlorinated Biphenyls (Compound)

4 Related Records

UVCB Polychlorinated biphenyls

UVCB Component	Type	Component Structure	Source	Reference
2',3,4,4',5-Pentachlorobiphenyl	Component		DOI:10.5281/zenodo.7414795	https://comptox.epa.gov/dashboard/chemical-lists/PCBCHEMICALS
2',3,4-Trichlorobiphenyl	Component		DOI:10.5281/zenodo.7414795	https://comptox.epa.gov/dashboard/chemical-lists/PCBCHEMICALS

Value Adding UVCB Data

Environmental Cheminformatics > PubChem Docs > Repository

added camphor tree example
Emma Schymanski authored 1 minute ago

pubchem-docs / taxonomy / Camphor_tree

Camphor_tree_metabolites.Rmd

title: "Retrieving_Camphor tree_Chemical" author: "Emma SCHYMANSKI" date: "15/04/2023" output: pdf_document urlcolor: blue

knitr::opts_chunk\$set(echo = TRUE)

Retrieving Camphor tree Chemicals from PubChem

Emma SCHYMANSKI 15/04/2023

Automated extraction & formatting for workflows (e.g. RMarkdown)

HRMS data -> Features -> Processing & Prioritization -> Annotation

Data acquisition & pre-treatment

Chromatographic peak detection & cross-analysis grouping

Rule-based filtering, componentization & prioritization

Chemical formula & compound annotation

Software | Open Access | Published: 06 January 2021

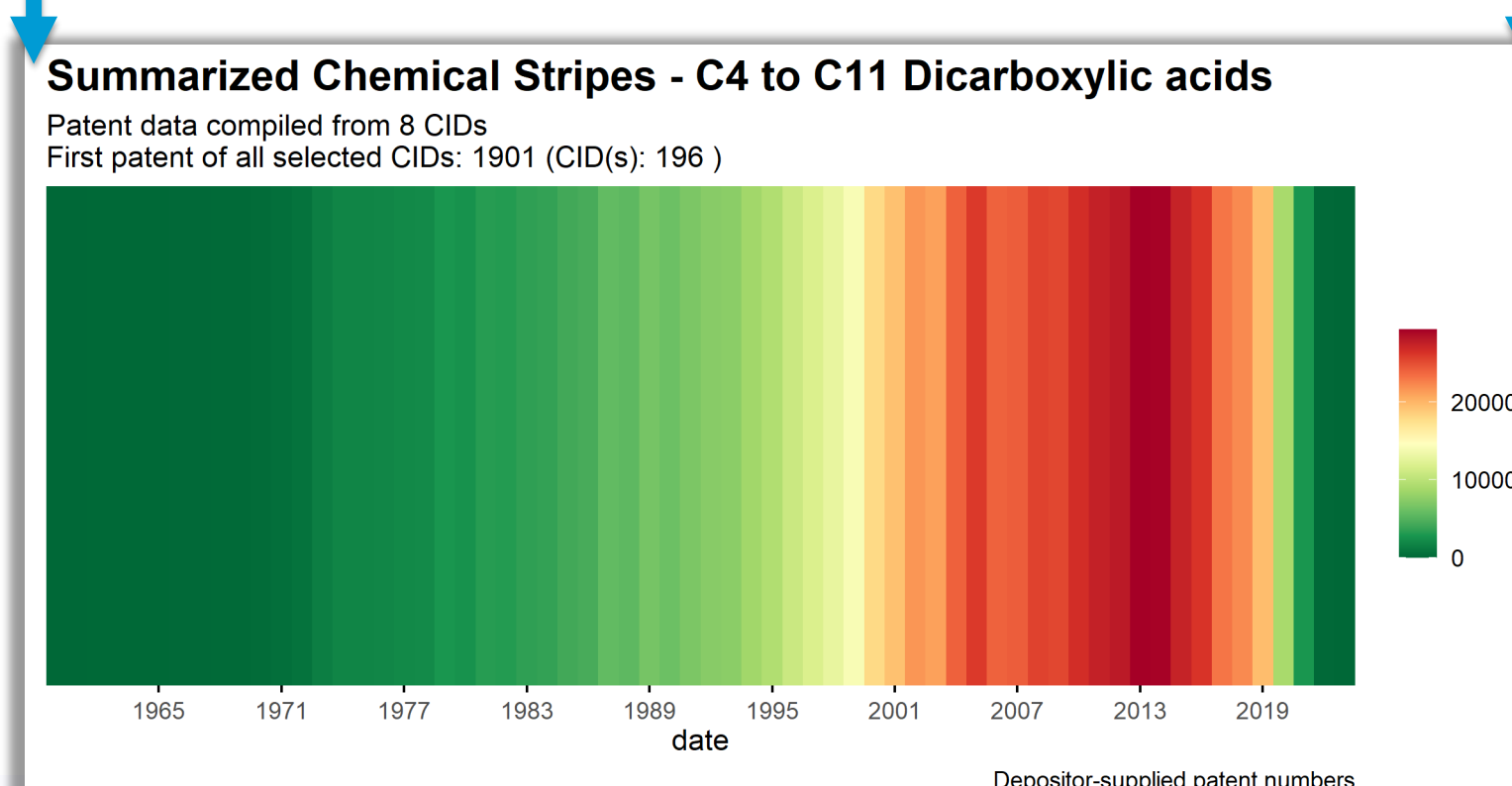
patRoom: open source software platform for environmental mass spectrometry based non-target screening

Journal of Cheminformatics 13, Article number: 1 (2021) | [Cite this article](#)

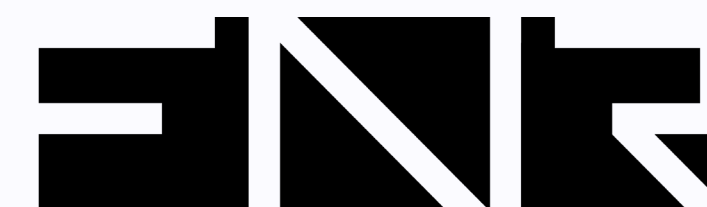
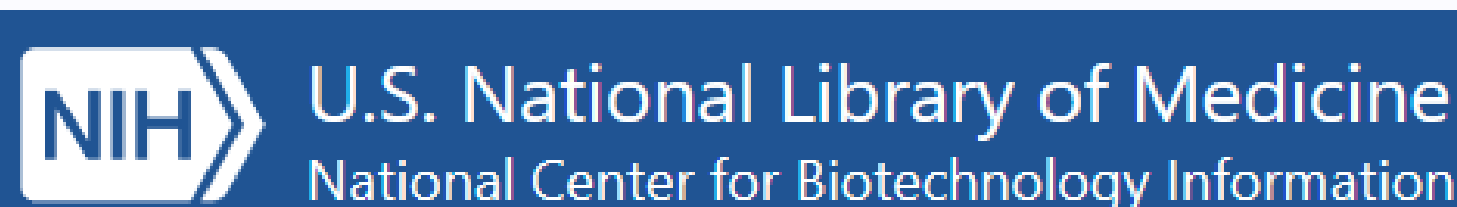
Rick Helmus, Thomas L. ter Laak, Annemarie P. van Wezel, Pim de Voogt & Emma L. Schymanski

Integration in to HRMS workflows

Trend visualisation (literature, patents) with Chemical Stripes



This work was supported in part by the National Center for Biotechnology Information (NCBI) of the National Library of Medicine (NLM), National Institutes of Health (NIH)



Luxembourg National Research Fund

