ICCE Satellite Event
Oslo, 22 June 2017

NORMAN Information Exchange: Suspect Lists and Mass Spectra

Emma Schymanski[1], Tobias Schulze[2], Reza Aalizadeh[3], Antony Williams[4], Natalia Glowacka[5], Lubos Cirka[5], Nikiforos Alygizakis[5], Ildiko Ipolyi[5], Jaroslav Slobodnik[5], Nikolaos Thomaidis[3], Juliane Hollender[1] … and more

[1]Eawag, Switzerland, [2]UFZ, Germany, [3]University of Athens, Greece, [4]United States Environmental Protection Agency, [5]Environmental Institute, Slovak Republic

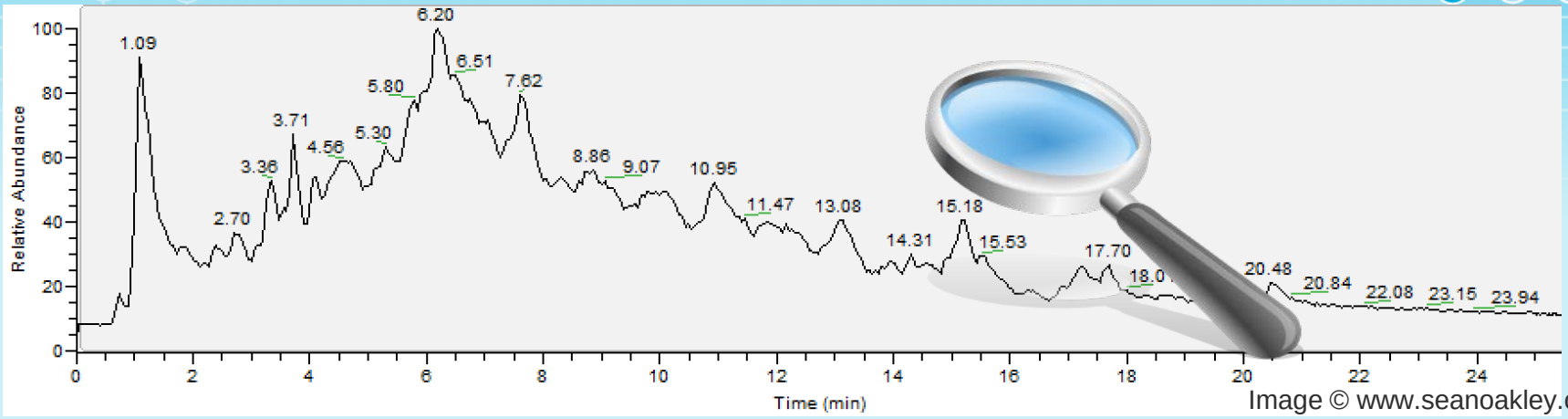


Image © www.seanoakley.com/

Target, Suspect and Non-Target Screening

KNOWNS	SUSPECTS	SUSPECTS	No Prior Knowledge
TARGET ANALYSIS	SPECTRUM SEARCH	SUSPECT SCREENING	NON-TARGET SCREENING
Targets found	Spectral match	Suspects found	Masses of interest
			(Molecular formula)
			DATABASE SEARCH / STRUCTURE GENERATION

Candidate selection (retention time, MS/MS, calculated properties)

Time, Effort & Number of Compounds….

MassBank: Japan, Europe, America ….

www.massbank.jp, www.massbank.eu, http://mona.fiehnlab.ucdavis.edu/

- MassBank started as a ***public repository*** in Japan, 2006
- No standard analytical method
  - Include many different data types (GC, LC, MS, MS/MS, HR, LR, AM…)
  - Contributor is responsible for data quality
- NORMAN network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances
  - Many different laboratories with different instruments & reference standards
  - “Emerging substances” and TPs: not yet widely known; not yet in databases
  - NORMAN joined MassBank in 2012 and founded MassBank.EU
- MassBank.JP and MassBank.EU are quite similar …
- MoNA (MassBank of North America) is the latest in the collection
  - Completely different database concept

Horai et al. 2010. JMS, 45(7) pp 703-714. DOI: 10.1002/jms.1777

MassBank – Crossing the World!

www.massbank.jp & www.massbank.eu

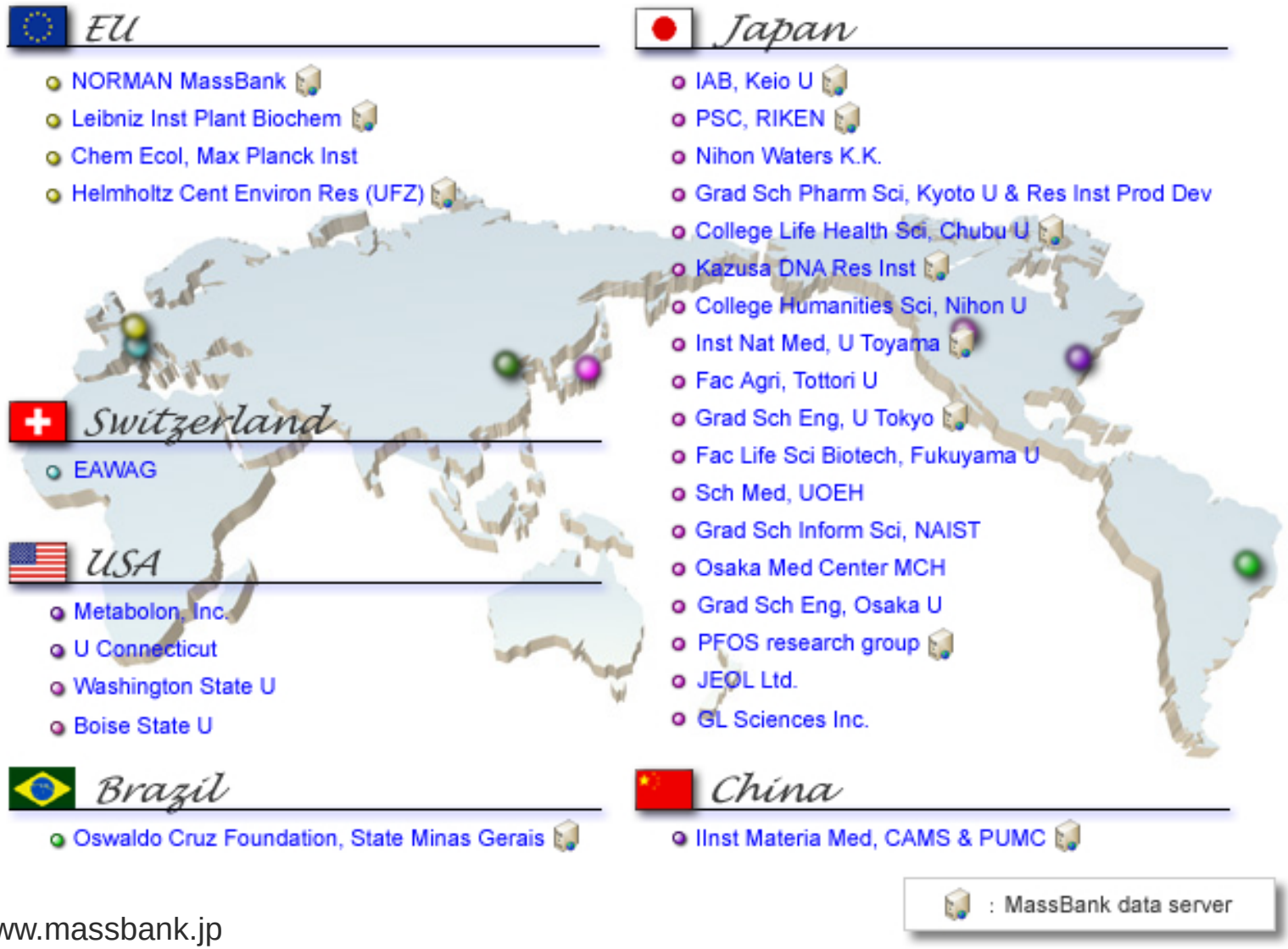
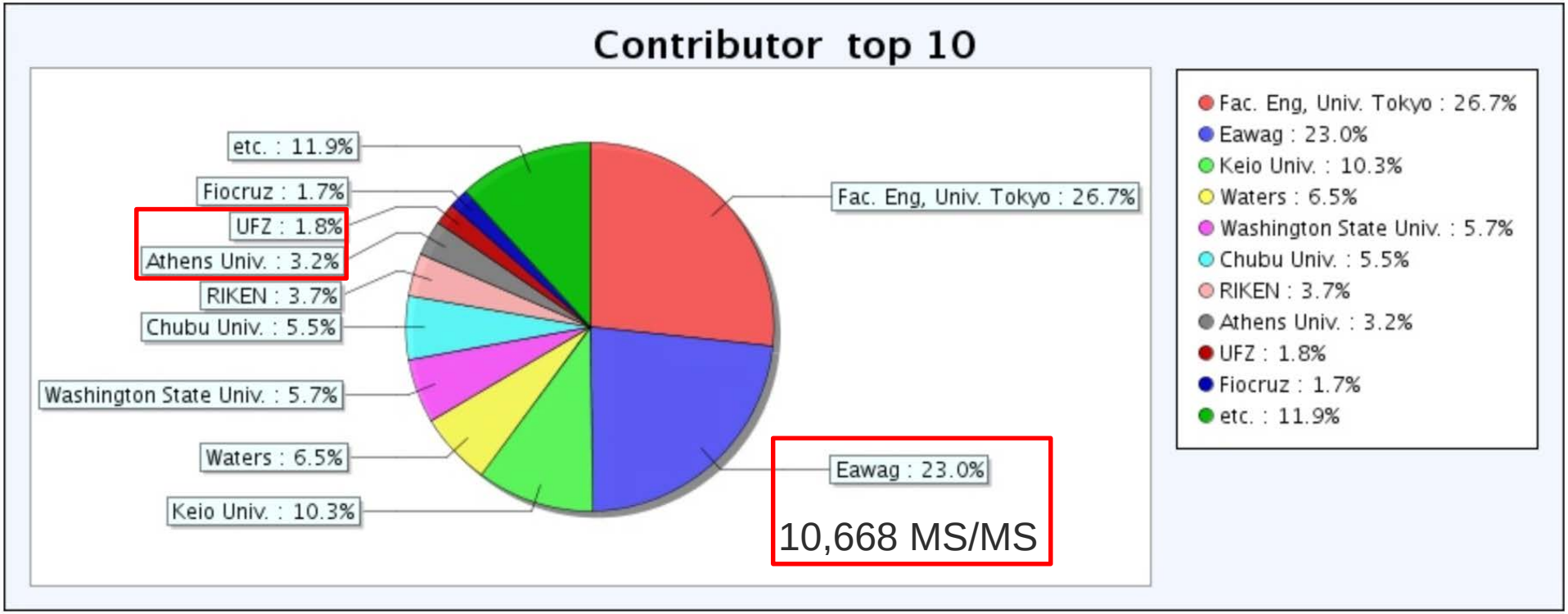


Image: www.massbank.jp

MassBank Now

www.massbank.jp & www.massbank.eu

MassBank now has **46,334 spectra*** from **32 contributing institutes**!



Contributions from European NORMAN member institutes

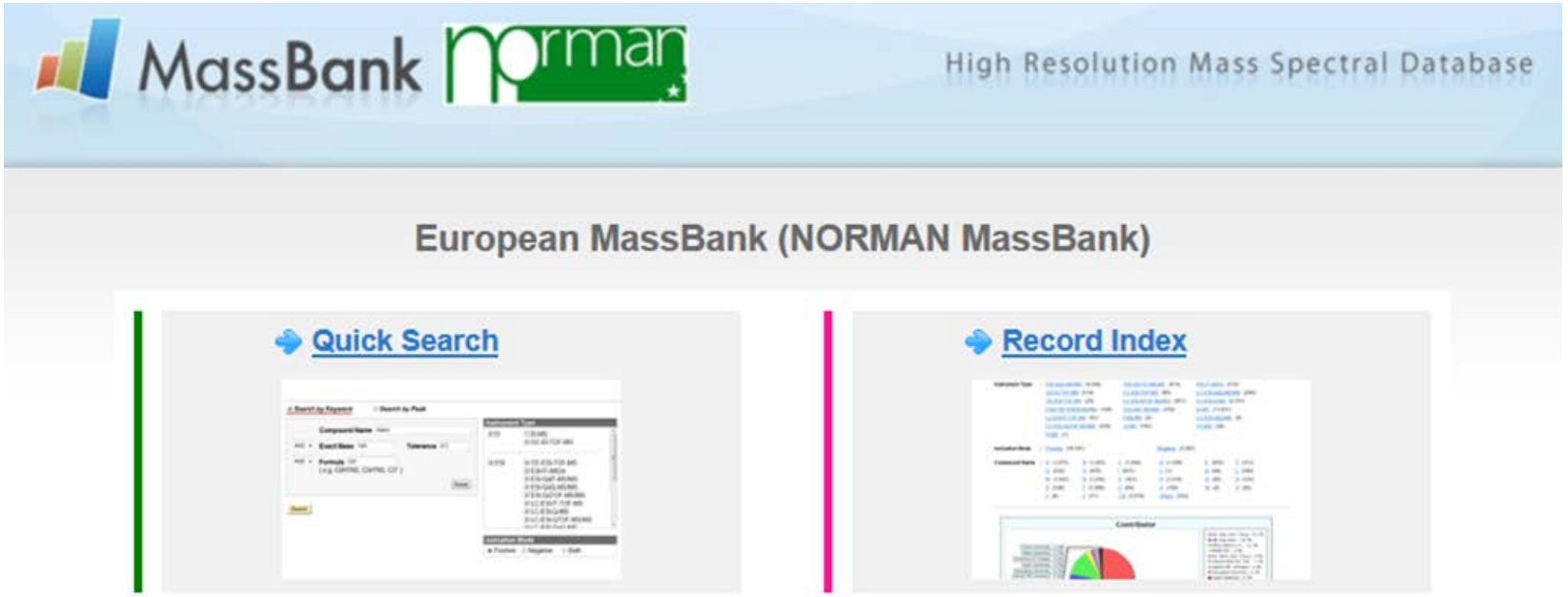
*Spectra numbers from http://mona.fiehnlab.ucdavis.edu/downloads

Image: www.massbank.jp

European MassBank

http://massbank.eu/MassBank

- MassBank.EU was founded late 2012, hosted at UFZ, Leipzig, Germany
  - 16,017 MS/MS spectra; 1,232 substances from NORMAN members
  - ***Tentative/unknown/literature*** spectra on massbank.eu (not massbank.jp)



European MassBank

http://massbank.eu/MassBank

- MassBank.EU was founded late 2012, hosted at UFZ, Leipzig, Germany
  - 16,017 MS/MS spectra; 1,232 substances from NORMAN members
  - ***Tentative/unknown/literature*** spectra on massbank.eu (not massbank.jp)

Athens Univ. (1,492)
Eawag (10,668)
European MassBank Server (NORMAN MassBank) (0)
Fukuyama Univ. (340)
JEOL Ltd. (45)
Kyoto Univ. (184)
MSSJ (34)
NAIST (671)
Osaka Univ. (449)
Tottori Univ. (16)
UOEH (35)
Univ. Toyama (253)

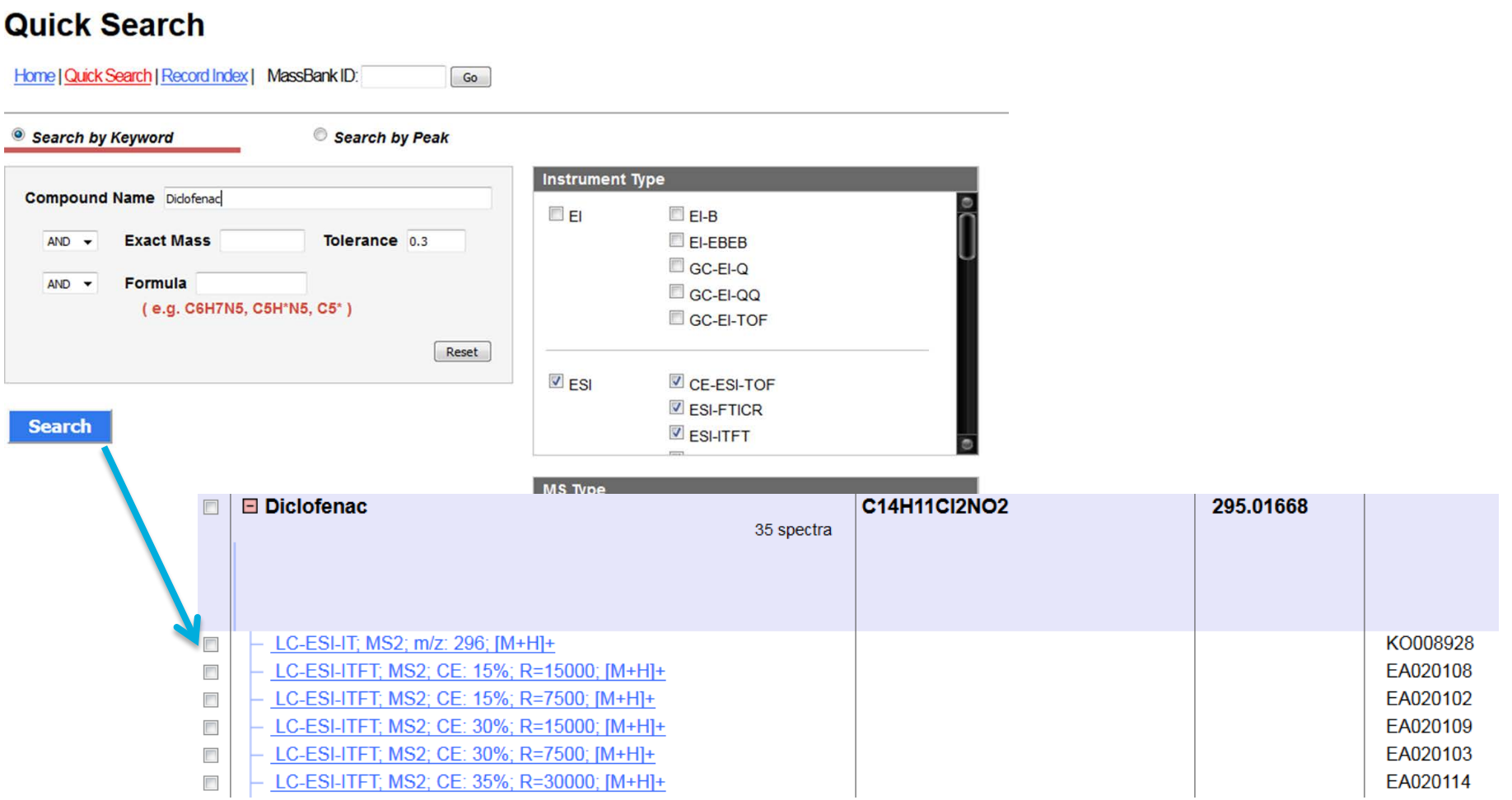
Boise State Univ. (4)
Eawag Additional Specs (620)
Fac. Eng, Univ. Tokyo (12,379)
GL Sciences Inc. (174)
Kazusa (273)
Literature Specs (39)
MetaboLights (58)
Nihon Univ. (488)
PFOS research group (413)
UFZ (2,758)
UPAO (12)
Washington State Univ. (2,626)

Chubu Univ. (2,563)
Env Anal Chem, U Tuebingen (116)
Fiocruz (800)
IPB Halle (528)
Keio Univ. (10,124)
MPI for Chemical Ecology (691)
Metabolon (149)
Osaka MCHRI (20)
RIKEN (1,718)
UFZ Additional Specs (107)
Univ. Connecticut (510)
Waters (2,992)

Image: www.massbank.eu

European MassBank

Basic search capabilities…

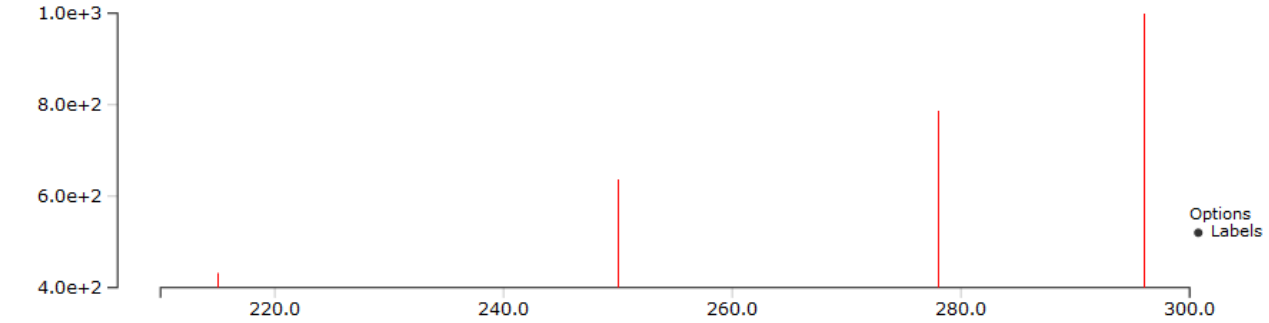


European MassBank

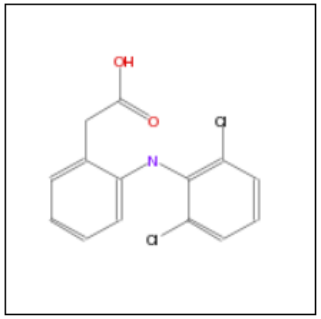
Example Mass Spectrum

Diclofenac; LC-ESI-ITFT; MS2; CE: 15%; R=15000; [M+H]+

Mass Spectrum



Chemical Structure




```
ACCESSION: EA020108
RECORD_TITLE: Diclofenac; LC-ESI-ITFT; MS2; CE: 15%; R=15000; [M+H]+
DATE: 2014.01.14
AUTHORS: Stravs M, Schymanski E, Singer H, Department of Environmental Chemistry, Eawag
LICENSE: CC BY
COPYRIGHT: Copyright (C) 2012 Eawag, Duebendorf, Switzerland
COMMENT: CONFIDENCE standard compound
COMMENT: EAWAG_UCHEM_ID 201
```

```
CH$NAME: Diclofenac
CH$NAME:  2-[2-(2,6-dichloroanilino)phenyl]acetic acid
CH$COMPOUND_CLASS: N/A; Environmental Standard
CH$FORMULA: C14H11Cl2N1O2
CH$EXACT_MASS: 295.0167
CH$SMILES: c1c(c(ccc1)Nc1c(cccc1Cl)Cl)CC(=O)O
CH$IUPAC: InChI=1S/C14H11Cl2NO2/c15-10-5-3-6-11(16)14(10)17-12-7-2-1-4-9(12)8-13(18)19/h1-7,17H,8H2,(H,18,19)
CH$LINK: CAS 15307-86-5
CH$LINK: CHEBI 47381
CH$LINK: KEGG C01690
CH$LINK: PUBCHEM CID:3033
CH$LINK: INCHIKEY DCOPUUMXTXDBNB-UHFFFAOYSA-N
CH$LINK: CHEMSPIDER 2925
```

Creating High Quality MS/MS Spectra

RMassBank

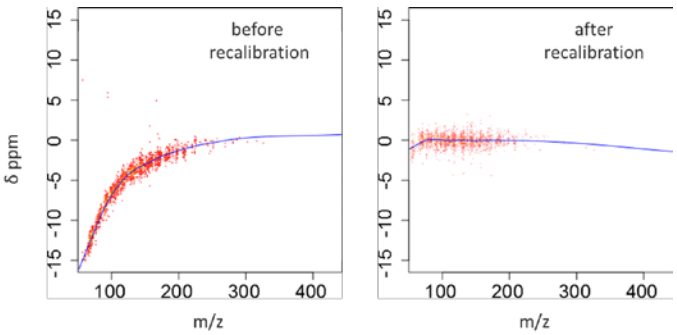
LC-MS/MS raw data

compound list: SMILES, name

online resources: CTS, CACTUS

Automatic MS and MS/MS Recalibration and Clean-up
Remove interfering peaks

Spectral Annotation with
- Experimental Details
- Compound Information



MassBank records

structure files

norman

MassBank.eu

https://github.com/MassBank/RMassBank/
http://bioconductor.org/packages/RMassBank/

Stravs, Schymanski, Singer and Hollender, 2013,
Journal of Mass Spectrometry, 48, 89–99. DOI: 10.1002/jms.3131

16,004 (61 %*) MS/MS spectra
1,269 (18 %*) substances
*% of ***all open*** LC-MS/MS data

MetFrag: *In silico* non-target identification

Status: 2016: MetFrag2.3 – Plus MS/MS Libraries!

References
External Refs
Data Sources
RSC Count
PubMed Count

ChemSpider
Search and share chemistry

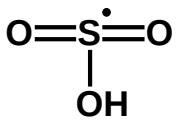
or

PubChem OPEN CHEMISTRY DATABASE

Literature Patents

mz [M-H]⁻
213.9637
± 5 ppm

Elements: C,N,S
5 ppm
0.001 Da



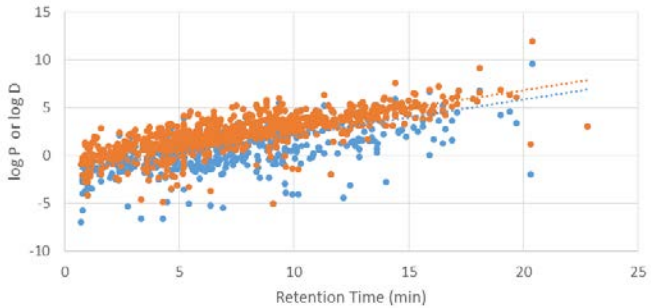
RT: 4.58 min
355 InChI/RTs

Suspect List(s)
InChIKeys

MS/MS

134.0054	339689.4
150.0001	77271.2
213.9607	632466.8

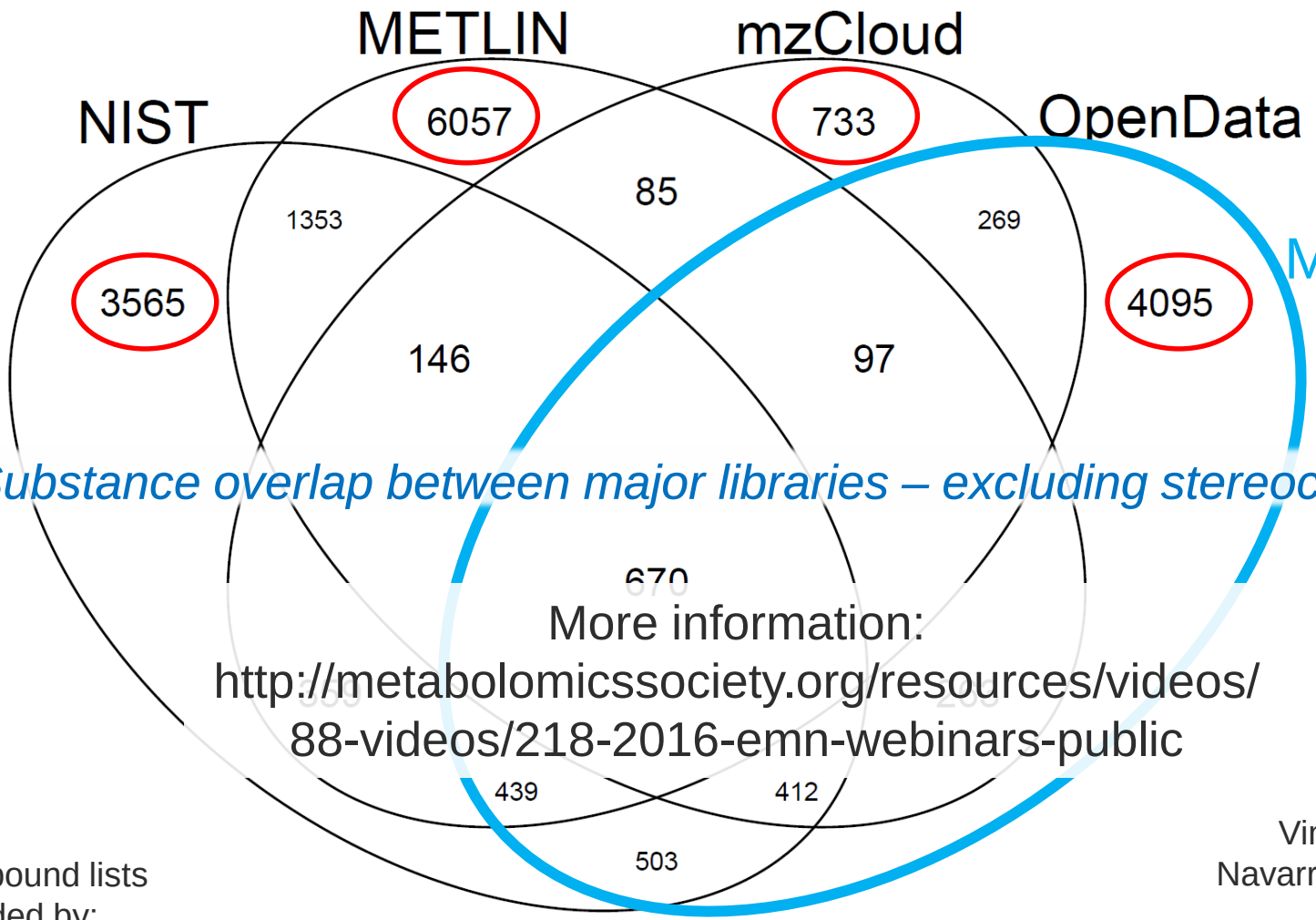
MoNA
MassBank of North America
MassBank.eu



Ruttkies, Schymanski, Wolf, Hollender, Neumann,
J. Chem. Inf., 2016, http://jcheminf.com/content/8/1/3

Enhancing Access to Mass Spectral Information

Most libraries still have many unique entries – *with different features*



OpenData = HMDB, GNPS, MassBank, ReSpect

Substance overlap between major libraries – excluding stereochemistry

More information:
http://metabolomicssociety.org/resources/videos/88-videos/218-2016-emn-webinars-public

Compound lists provided by:
S. Stein, R. Mistrik, Agilent

Vinaixa, **Schymanski**, Navarro, Neumann, Salek, Yanes, *2015, TrAC,* DOI: 10.1016/j.trac.2015.09.005.

SPLASH – Communicate between libraries

http://splash.fiehnlab.ucdavis.edu/

SPectraL hASH – an identifier for mass spectra

splash10 - 0002 - 0900000000 - b112e4e059e1ecf98c5f
[version] - [top10] - [histogram] - [hash of full spectrum]

http://mona.fiehnlab.ucdavis.edu/#/spectra/splash/splash10-0002-0900000000-b112e4e059e1ecf98c5f

https://www.google.ch/search?q=splash10-0002-0900000000-b112e4e059e1ecf98c5f

MassBank Record: EA278005

PK$SPLASH: splash10-0uxr-0973000000-87d07ddd2ed24b9598d7
PK$ANNOTATION: m/z tentative_formula formula_count mass error(ppm)
58.0651 C3H8N+ 1 58.0651 0.25
69.0335 C4H5O+ 1 69.0335 -0.45

Wohlgemuth *et al.* 2016,
Nature Biotechnology, 34 (11), 1099-1101
http://splash.fiehnlab.ucdavis.edu/

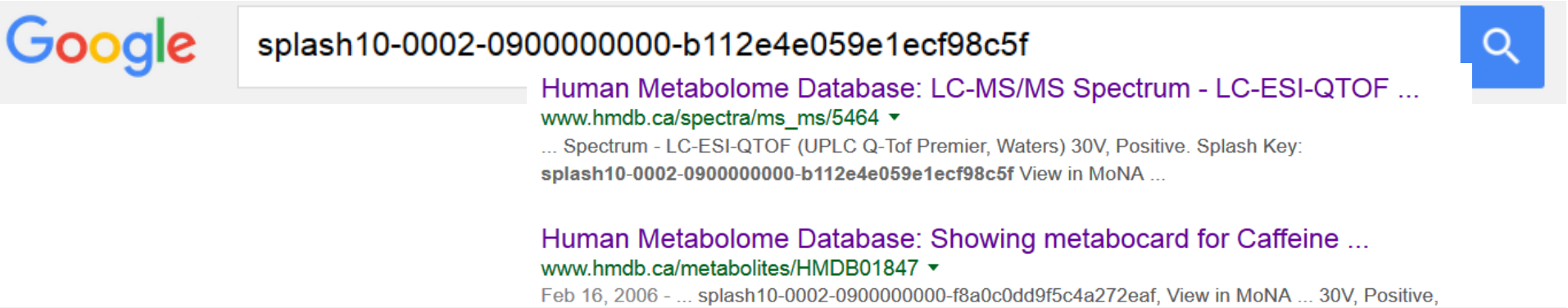
Image: www.massbank.eu

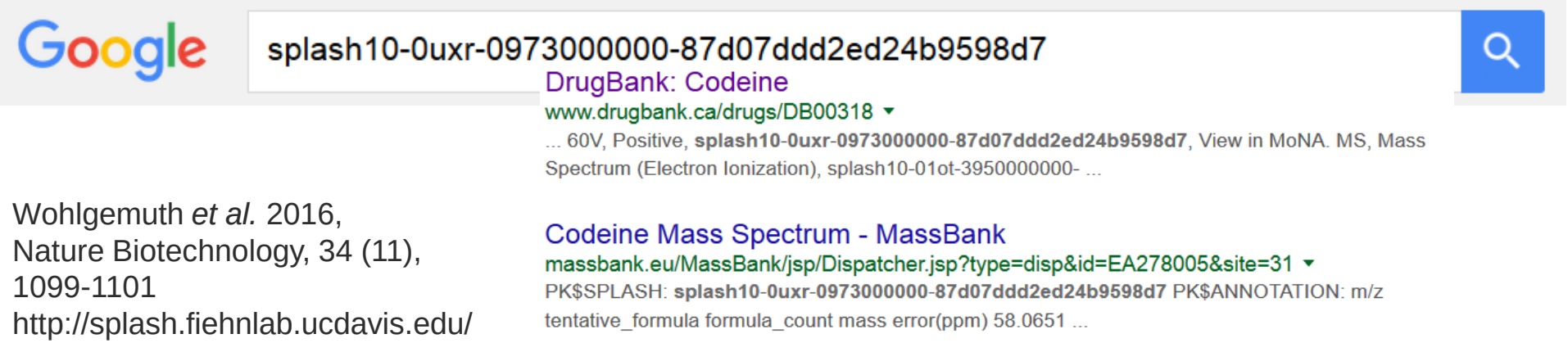
SPLASH – Communicate between libraries

splash10 - 0002 - 0900000000 - b112e4e059e1ecf98c5f
[version] - [top10] - [histogram] - [hash of full spectrum]

http://mona.fiehnlab.ucdavis.edu/#/spectra/splash/splash10-0002-0900000000-b112e4e059e1ecf98c5f

https://www.google.ch/search?q=splash10-0002-0900000000-b112e4e059e1ecf98c5f

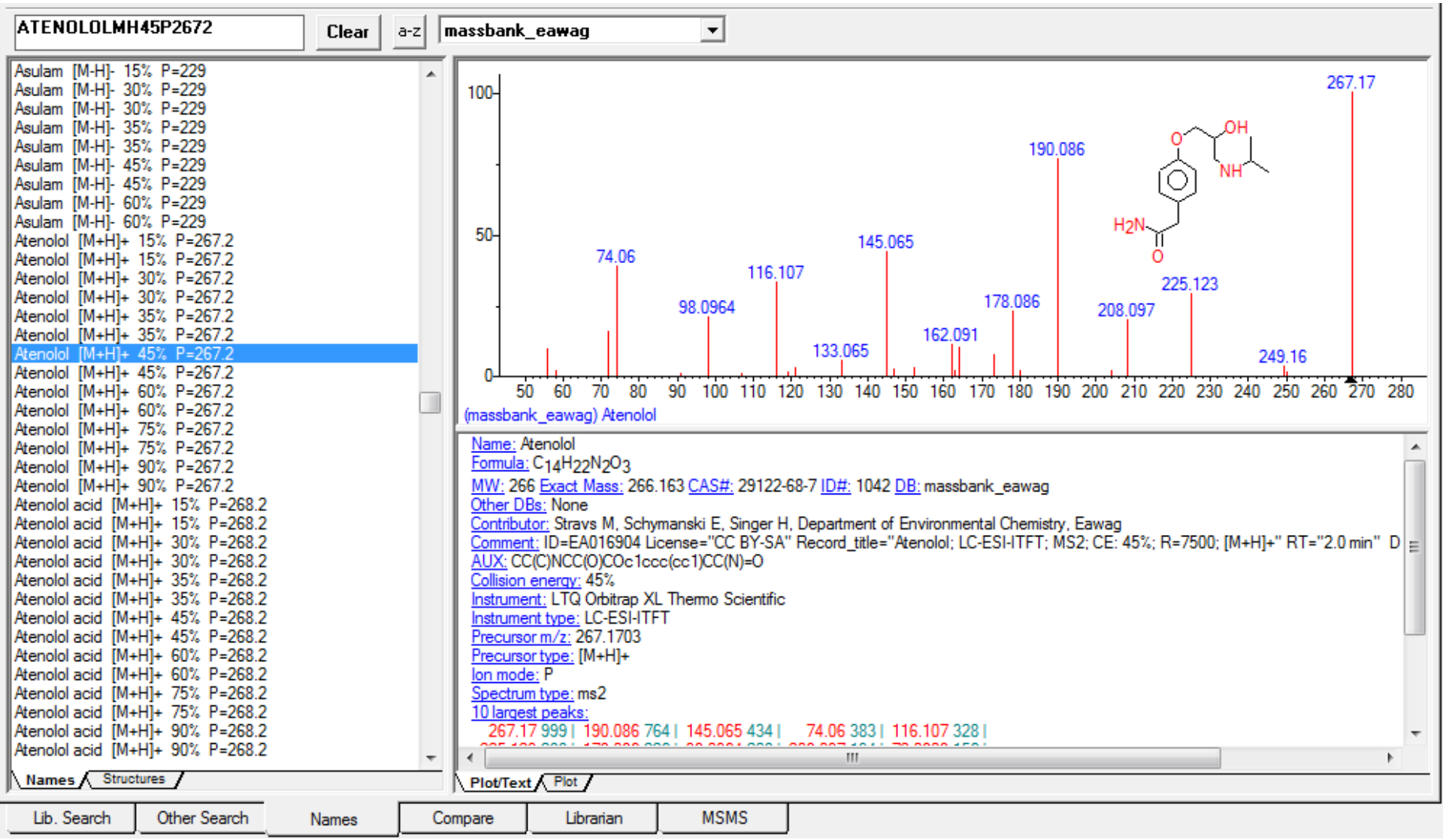




Wohlgemuth *et al.* 2016, Nature Biotechnology, 34 (11), 1099-1101
http://splash.fiehnlab.ucdavis.edu/

MassBank: Integration in the NIST library

MassBank records as separate databases



Target, Suspect and Non-Target Screening

KNOWNS	SUSPECTS	SUSPECTS	No Prior Knowledge
TARGET ANALYSIS	SPECTRUM SEARCH	SUSPECT SCREENING	NON-TARGET SCREENING
Targets found	Spectral match	Suspects found	Masses of interest
			(Molecular formula)
			DATABASE SEARCH / STRUCTURE GENERATION

Candidate selection (retention time, MS/MS, calculated properties)

Time, Effort & Number of Compounds….

2015: Suspect and Non-target Screening Across Europe

Collaborative Trial Suspect Screening Lists

19 institutes …

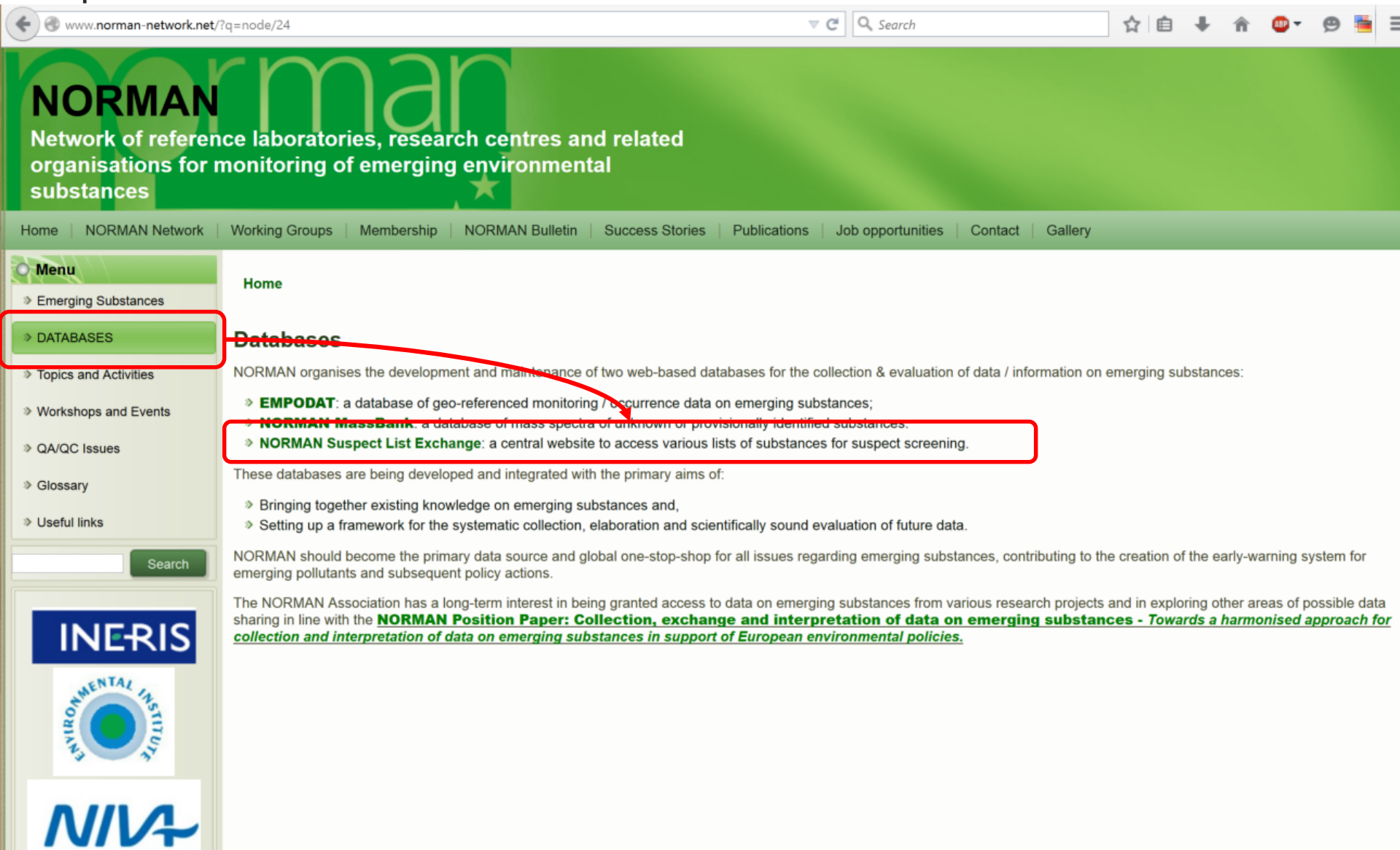
More data sources and "lists" than participants!

Database/Library Name	State as used during the trial: Total Compounds	State as used during the trial: Compounds with Spectra	Current State: Compounds at March 2015
ChemSpider [35]	32 million		32 million
DAIOS [49,50]	1,404	>1,000[a]	1,404
PubChem [48]	63,105,228		68,479,719
STOFF-IDENT [38]	7,864[b]		7,864
MassBank MS/MS [51-53]		3,350	3,350
mzCloud [54]		1,956	2,510
NIST EI-MS [11,55]		212,961[c]	242,477
NIST MS/MS [11,55]		4,628	8,171
Wiley Registry of Mass Spectral Data (EI-MS) [56]		289,000 [12]	638,000
Agilent Broecker, Herre & Pragst Toxicology/Forensics[f] [57,58]	8,998[c]	3,497	8,998
Agilent Pesticide Library LC/Q-TOF MS/MS[f] [59]	1,664	~700[c]	1,664
Agilent Pesticide Library GC/Q-TOF EI-MS[f]	750	750	750
Agilent METLIN Synthetic Substance Library[g]	64,092[c]	~10,000[c]	64,092
Agilent METLIN Scripps Online Database[f,g] [60,61]	83,135	12,171[c]	240,566
Agilent Veterinary Drug Library[f]	1,684	770	1,684
Bruker ToxScreener (incl. Pesticide Screener)[g] [62]		704[ad]	1753
Sciex / AB Sciex LC/MS/MS Meta Library[g] [63]		2,381[c]	2,381
Thermo Environmental Food Safety (EFS)[g] with retention time (RT)[g]		447[p]; 278[n]; 454[dp]; 90[dn]	732
Thermo toxicology[g]		618[p]; 36[n]	654
Waters database with RT[g]		730[de]	730
In-house Libraries without spectra (two participants)	2,000; 1,600 [17]		2,000; 1,600
In-house Libraries with spectra (two participants)		526[d]; 63[d]	526; 63
In-house Libraries with spectra for some substances	2,200[d]	835[ad]	2,200
	7,815	1500[ap]; 500[an]	7,815
	3,000	350[d]	3,000
Surfactant List [3]	394		394

Schymanski *et al*. 2015, ABC, DOI: 10.1007/s00216-015-8681-7

NORMAN Network Suspect List Exchange

…part of the NORMAN Databases Collection

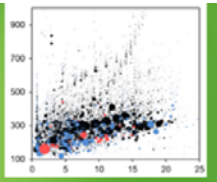


NORMAN Network Suspect List Exchange

http://www.norman-network.com/?q=node/236

NORMAN

Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances



Home | NORMAN Network | Working Groups | Membership | NORMAN Bulletin | Success Stories | Publications | Job opportunities | Contact | Gallery | NORMAN GA meetings

Menu

- Emerging Substances
- DATABASES
- Topics and Activities
- Workshops and Events
- QA/QC Issues
- Glossary

User login

Username *

Password *

Request new password

Log in

NORMAN Suspect List Exchange

As part of a series of workshops in September 2014, NORMAN members expressed the need to exchange various lists of substances to improve their suspect screening efforts. An initiative of the 2015 Joint Programme of Activities involved establishing this website as a central access point for NORMAN members (and others) to find suspect lists relevant for their environmental monitoring question. All suspect lists currently available are compiled in the table below and are being progressively integrated into the US EPA CompTox Chemistry Dashboard (**website**, **downloads**). The "Link to full list" column below contains an excel or comma-separated file (csv) with all available information, e.g. as provided as supporting information for the publication, while the third column provides a list of the structures as InChIKeys only, which allows suspect searching using MetFrag or other workflows. The fourth column contains references for the data: please cite these references if you use the respective datasets.

Coordination: Emma Schymanski, Eawag; Curation/RTI/toxicity: Reza Aalizadeh & Nikos Thomaidis, Uni. Athens; CompTox: Antony Williams, US EPA; Webmaster: Natalia Glowacka, Environmental Institute; IT: Lubos Cirka, Environmental Institute; Contributors: see below.

If you have any feedback or a list that you would like included, please contact suspects@normandata.eu.

Interactive merged list of all suspect substances (update in progress)

Full Lists **InChIKeys** **References**

Name and Description	Link to full list	Link to InChIKey	References
Merged NORMAN Suspect List "SusDat"	NORMAN_SusDat_MergedSuspects24052017.xlsx	NORMAN_SusDat_MSready	This is the merged list of all suspect lists containing structures. See **here** for an interactive version. Compiled by Reza Aalizadeh, University of Athens, now including RTI and toxicity values.
NORMAN Compounds in MassBank	MassBankEU_Compounds_11042017.csv	MassBankEU …42017.txt	**www.massbank.eu** Stravs *et al.* 2012. DOI: 10.1002/jms.3131
HSWT/LfU STOFF-IDENT database of water-relevant substances	STOFF-IDENT_content_ed_17052016.xlsx STOFF-IDENT_Content_28102016.xlsx STOFF-IDENT_Content_28102016.csv	STO …eys.txt	The database enables the search for exact masses from target or unknown lists and the automatic use of a Retention Time Index. See: **http://bb-x-stoffident.hswt.de** - free access after registration
NORMAN Collaborative Trial Targets and	Targ_Sus_NT-wID_LC_final_31102016.xlsx Targ_Sus_NT-wID_LC_final_31102016.csv Targ_Sus_NT-wID_GC_final_31102016.xlsx	Targ_Sus_…_GC_final_InChIKeys_31102016.txt Targ_Sus_N…-wID_LC_final_InChIKeys_31102016.txt	Schymanski *et al.* 2015. DOI: **10.1007/s00216-015-8681-7**

MetFrag 2.3

NORMAN Suspect List Exchange (2016)

Contributions so far…

MassBank.eu

National and Kapodistrian University of Athens

HELMHOLTZ CENTRE FOR ENVIRONMENTAL RESEARCH – UFZ

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

eawag aquatic research

PFAS Suspect List of fluorinated substances

?TOFF !DENT

Antibiotic Suspect List (ITN MSCA ANSWER)

normaNEWS

Strategies to Characterize Polar Organic Contamination in Wastewater: Exploring the Capability of High Resolution Mass Spectrometry

Emma L. Schymanski,[†] Heinz P. Singer,[†] Philipp Longrée,[†] Martin Loos,[†,§] Matthias Ruff,[†] Michael A. Stravs,[†,§] Cristina Ripollés Vidal,[‡] and Juliane Hollender[†,§,*]

Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis

Emma L. Schymanski[1] · Heinz P. Singer[1] · Jaroslav Slobodnik[2] · Ildiko M. Ipolyi[2] · Peter Oswald[2] · Martin Krauss[3] · Tobias Schulze[3] · Peter Haglund[4] · Thomas Letzel[5] · Sylvia Grosse[5] · Nikolaos S. Thomaidis[6] · Anna Bletsou[6] · Christian Zwiener[7] · María Ibáñez[8] · Tania Portolés[8] · Ronald de Boer[9] · Malcolm J. Reid[10] · Matthias Onghena[11] · Uwe Kunkel[12] · Wolfgang Schulz[13] · Amélie Guillon[14] · Naïke Noyon[14] · Gaëla Leroy[15] · Philippe Bados[16] · Sara Bogialli[17] · Draženka Stipaničev[18] · Pawel Rostkowski[19] · Juliane Hollender[1,20]

Critical evaluation of a simple retention time predictor based on LogKow as a complementary tool in the identification of emerging contaminants in water

Richard Bade, Lubertus Bijlsma, Juan V. Sancho, Felix Hernández *

Data-driven prioritization of chemicals for various water types using suspect screening LC-HRMS

Rosa M.A. Sjerps [a, *], Dennis Vughs [a], Jan A. van Leerdam [a], Thomas L. ter Laak [a, b], Annemarie P. van Wezel [a, c]

Extended Suspect and Non-Target Strategies to Characterize Emerging Polar Organic Contaminants in Raw Wastewater with LC-HRMS/MS

Pablo Gago-Ferrero,[†] Emma L. Schymanski,[‡] Anna A. Bletsou,[†] Reza Aalizadeh,[†] Juliane Hollender,[‡,§] and Nikolaos S. Thomaidis*[,†]

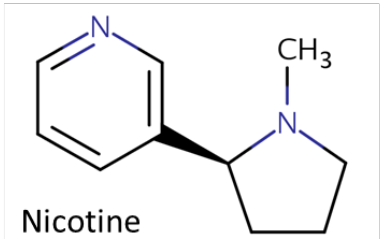
NORMAN Suspect List Exchange (NEW in 2017)

Pharmaceutical List with Consumption Data	SwissPharma_TableS2.csv	SwissPharma_TableS2_InChIKeys.txt	Singer *et al.* 2016. DOI: **10.1021/acs.est.5b03332**
Swiss Insecticides, Fungicides and TPs	SwissPesticides_TableS1.csv	SwissPesticides_TableS1_InChIKeys.txt	Moschet *et al.* 2013. DOI: **10.1021/ac4021598**
NormaNEWS for retrospective screening of new emerging contaminants	NormaNEWS_V4_26042017.csv	NormaNEWS_V4_InChIKeys.txt	**NormaNEWS** list provided by Nikiforos Alygizakis, Saer Samanipour and Kevin Thomas
Combined Inventory of Ingredients Employed in Cosmetic Products (2000) and Revised Inventory (2006)	Merged_CosmeticProducts_04052017.csv 3,333 Cosmetic products	Merged_CosmeticProducts_04052017_InChIKeys.txt	The scientific committee on cosmetic products and non-food products Intended for consumers - **SCCNFP/0389/00 Final** and Commission **Decision 2006/257/EC** amending the Decision 96/335/EC. Provided by Peter von der Ohe, UBA, curated by Reza Aalizadeh, University of Athens
PFAS Highly fluorinated substances list: KEMI	PFAS_Market_Keml_EPA_1Feb2017.xlsx ~2,600 PFAS	Curation in progress: coming soon	Appendix 2 from Swedish Chemicals Agency **KEMI Report 7/15**. Provided by Stellan Fischer, KEMI
NORMAN Priority List 2015	NORMAN_PriorityList_2016.csv Further curation in progress...	NORMAN_PriorityList_2016_InChIKeys.txt	Priority substances from NORMAN WG-1 (Prioritisation), provided by Valeria Dulio
French Monitoring List	French_List_08052017.csv Further curation in progress	FrenchList_UniqueInChIKeys_08052017.txt	Provided by Valeria Dulio, curated by Reza Aalizadeh, University of Athens
KEMI Market List	KEMI_MarketList_12052017_MSready.xlsx 24,883 Substances (Expo, Hazard Scores)	KEMI_MarketList_12052017_MSready_InChIKeys.txt	Provided by Stellan Fischer, KEMI including Hazard and Exposure scores, documented **here**. Curated by Reza Aalizadeh, University of Athens.
TSCA Surfactants	Coming soon...		Provided by Lee Ferguson, sourced from James Little

The Chemical Identity Challenge

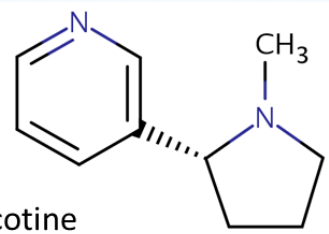
Schymanski & Williams, 2017, ES&T
DOI: 10.1021/acs.est.7b01908

LEGEND: Name, SMILES
DTXSID | InChIKey 1st Block
CAS | **Monoiso. Mass** | logP | **Sources**
Data on: **Toxicity | Exposure | Bioassays**



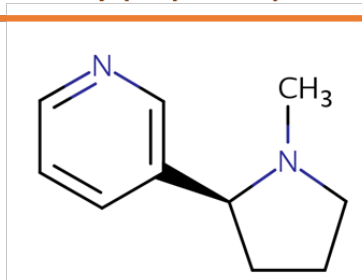
Nicotine

CN1CCC[C@H]1C1=CN=CC=C1
DTXSID1020930| SNICXCGAKADSCV
54-11-5 | **162.1157**| 0.929| **72**
Tox: **yes**| Expo: **yes**| Bioassay: **yes**



D-Nicotine

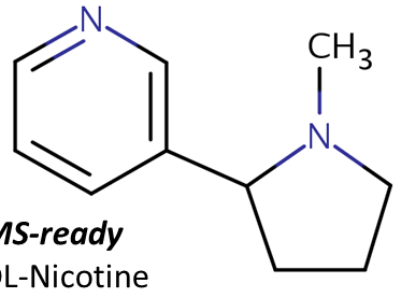
CN1CCC[C@@H]1C1=CN=CC=C1
DTXSID004635| SNICXCGAKADSCV
25162-00-9 | **162.1157**| 0.929| **20**
Tox: **no**| Expo: **yes**| Bioassay: **yes**



HCl

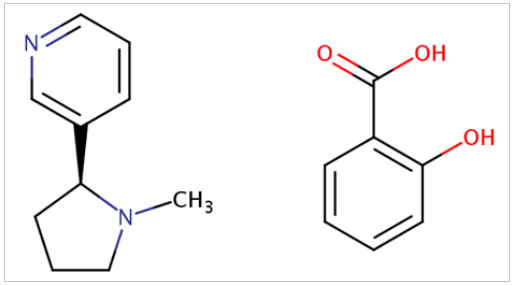
Nicotine hydrochloride

Cl.CN1CCC[C@H]1C1=CN=CC=C1
DTXSID602093| HDJBTCAJIMNXEW
2820-51-1 | **198.0924** | 0.929| **9**
Tox: **no**| Expo: **yes**| Bioassay: **yes**



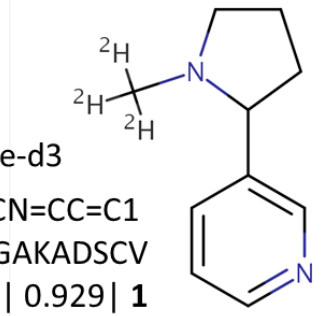
MS-ready
DL-Nicotine

CN1CCCC1C1=CN=CC=C1
DTXSID3048154 | SNICXCGAKADSCV
22083-74-5 | **162.1157**| 0.953| 9
Tox: **yes**| Expo: **no**| Bioassay: **yes**



Benzoic acid, 2-hydroxy-, compd. with 3-[(2S)-1-methyl-2-pyrrolidinyl]pyridine (1:1)

OC(=O)C1=C(O)C=CC=C1.CN1CCC[C@H]1C1=CN=CC=C1
DTXSID5075319| AIBWPBUAKCMKNS
29790-52-1| **300.1474**| 0.929| **6**
Tox: **no**| Expo: **yes**| Bioassay: **no**



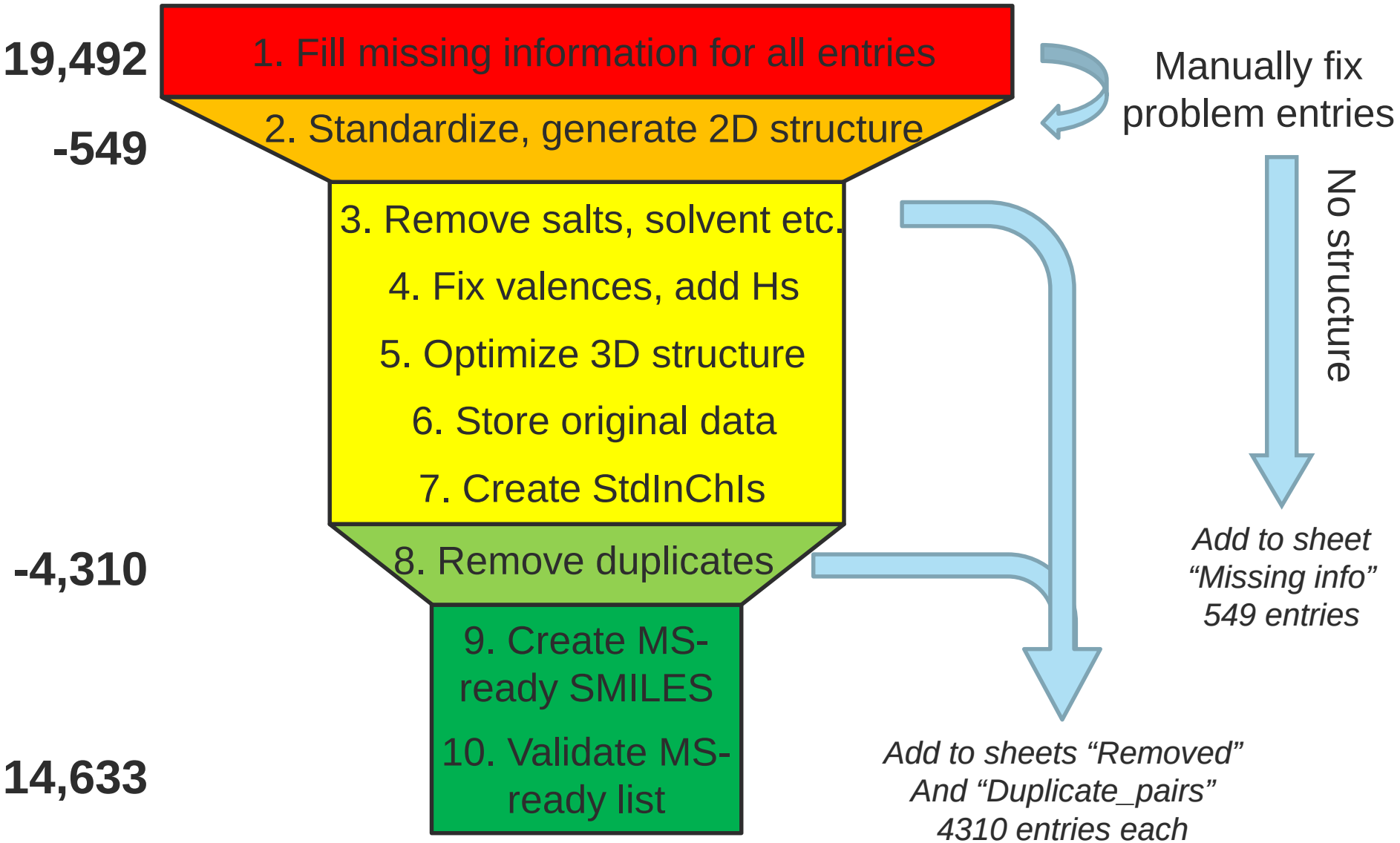
DL-Nicotine-d3

[2H]C([2H])([2H])N1CCCC1C1=CN=CC=C1
DTXSID80442666| SNICXCGAKADSCV
69980-24-1| **165.1345**| 0.929| **1**
Tox: **no**| Expo: **no**| Bioassay: **no**

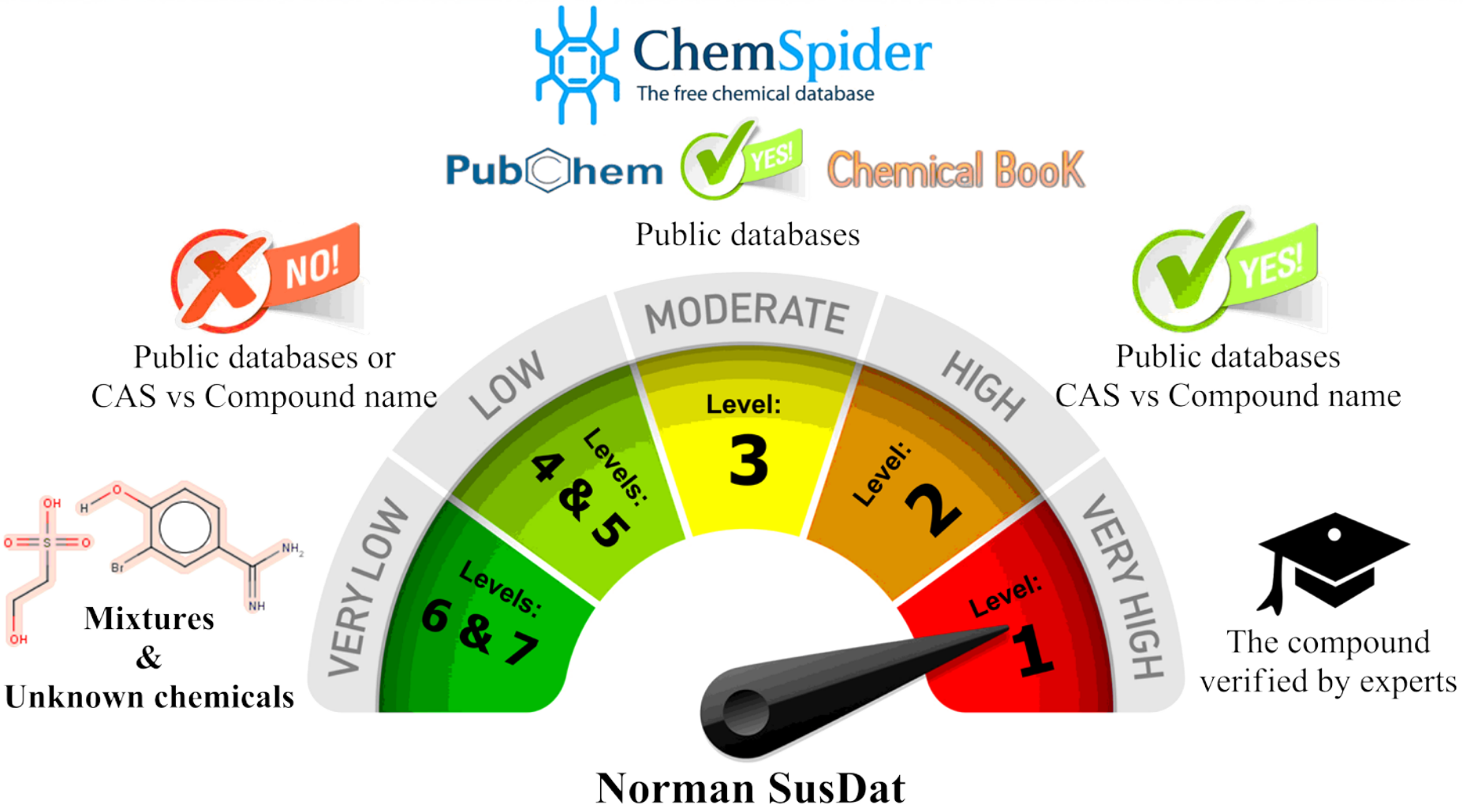
Curation and Merging Workflow

Aalizadeh *et al*. in prep.
Fourches et al 2010, 2016

15(+) lists => one



Validation “Level”



Aalizadeh *et al*. in prep.; modified validation level concept from the CompTox Chemistry Dashboard

NORMAN-SusDat – the “merged” data table

Name, Identifiers, Validation level, Source, MSMS, RTI, Toxicity, logKow

Mol_ID	Name	CAS_RN	ValidationLevel	SMILES	StdInChI	StdInChIKey	Optimized_SMILES	Optimized_StdInChI
SA1	Sulfaclozine	CAS_RN: 102-65-8	Level 4	c1cc(ccc1N)S(=	InChI=1S/C10H9ClN4O	QKLPUVXBJHRF	c1cc(ccc1N)S(=O)(=O)N	InChI=1S/C10H9ClN4
SA2	Sulfachlorpyridazine	CAS_RN: 80-32-0	Level 2	c1cc(ccc1N)S(=	InChI=1S/C10H9ClN4O	XOXHILFPRYWF	c1cc(ccc1N)S(=O)(=O)N	InChI=1S/C10H9ClN4

Mol_ID	MS_Ready_SMILES	MS_Ready_StdInChI	MS_Ready_StdInChIKey	Source	PubChem_CID	ChemSpiderID		
SA7							)(=O)(=O)N	InChI=1S/C7H10N4O
SA10	SA1	c1cc(ccc1N)S(=O)(=O)N	InChI=1S/C10H9ClN4(	QKLPUVXBJHRFQZ-UHFFFAOY	UOA	66890	60252	NS(=O)(=O)	InChI=1S/C11H12N4
SA11	SA2	c1cc(ccc1N)S(=O)(=O)N	InChI=1S/C10H9ClN4(	XOXHILFPRYWFOD-UHFFFAOY	UOA	6634	6382	S(=O)(=O)c	InChI=1S/C9H10N4O

Mol_ID	Monoiso_Mass	[M+H]+	[M-H]-	Pred_RTI_Positive_ESI	Uncertainty_RTI_pos	Pred_RTI_Negative_ESI	Uncertainty_RTI_neg
SA2618	134.1096	135.1174	133.1017	651.14	Covered by Model	602.41	Covered by Model
SA2619	174.1620	175.1698	173.1542	653.00	Covered by Model	507.67	Experimental proof is needed

Mol_ID	Pimephales_promelas_toxicity	LC50_96_hr_ug/L	Uncertainty_Pimephales_promel	logKow_EPISuite	Exp_logKow_EPISuite
SA2618	4.826	2001.23	Covered by Model	4.01	4.38
SA2619	4.451	6159.47	Covered by Model	4.43	NA
SA2620	2.708	184000.79	Covered by Model	1.87	1.77
SA2621	2.857	177844.92	Covered by Model	0.52	0.92
SA2622	5.820	383.64	Covered by Model	5.3	4.2
SA2623	2.395	595909.45	Covered by Model	-0.97	NA
SA2624	7.720	7.86	Covered by Model	4.87	3.49
SA2625	4.912	3002.80	Covered by Model	2.69	NA
SA2627	3.527	70059.88	Covered by Model	0.76	1.31
SA2628	7.138	36.75	Experimental proof is needed	14.31	NA
SA2629	4.873	1824.36	Covered by Model	4.61	4.38
SA2630	7.729	8.49	Experimental proof is needed	12.23	NA
SA2631	5.490	961.81	Covered by Model	3.74	NA
SA2632	4.648	6653.84	Covered by Model	2.49	2.45
SA2633	4.756	8846.76	outside of Chemical space	8.78	NA
SA2636	1.928	18301421.78	outside of Chemical space	-3.37	NA
SA2637	2.628	414360.84	Experimental proof is needed	-1.99	NA

NORMAN-SusDat – the “merged” data table

SCREEN SMART – OR BIG – OR BOTH?

All suspect lists available in one table:

- http://www.norman-network.com/datatable/
- Quick search options on every field, e.g. name, mass, …

NORMAN-SusDat: NORMAN Suspect List Exchange Merged Data Table

Reset search results

Show 100 entries

Mol_ID	Name	CAS_RN	ValidationLevel	SMILES
SA1	Sulfaclozine	CAS_RN: 102-65-8	Level 1	c1cc(ccc1N)S(=O)(=O)Nc2cncc(n2)Cl
SA10	Sulfamerazine	CAS_RN: 127-79-7	Level 2	Cc1ccnc(n1)NS(=O)(=O)c2ccc(cc2)N
SA100	Bromhexine	CAS_RN: 3572-43-8	Level 2	CN(Cc1cc(cc(c1N)Br)Br)C2CCCCC2
SA1000	Sotalol	CAS_RN: 3930-20-9	Level 2	CC(C)NCC(C1=CC=C(C=C1)NS(=O)(=O)C)O
SA10000	nicomorphine	CAS_RN: 639-48-5	Level 4	CN1CC[C@@]23[C@H]4Oc5c2c(C[C@@H]1[C@@H]3C
SA10001	2-(3-Pyridyl)-1H-benzimidazole	CAS_RN: 1137-67-3	Level 4	c1ccc2[nH]c(nc2c1)-c1cccnc1
SA10002	2-chlorobenzylamine	CAS_RN: 89-97-4	Level 2	NCc1ccccc1Cl

MERGING SEVERAL LISTS IS NOT TRIVIAL!
WORK IN PROGRESS!!!

The CompTox Chemistry Dashboard

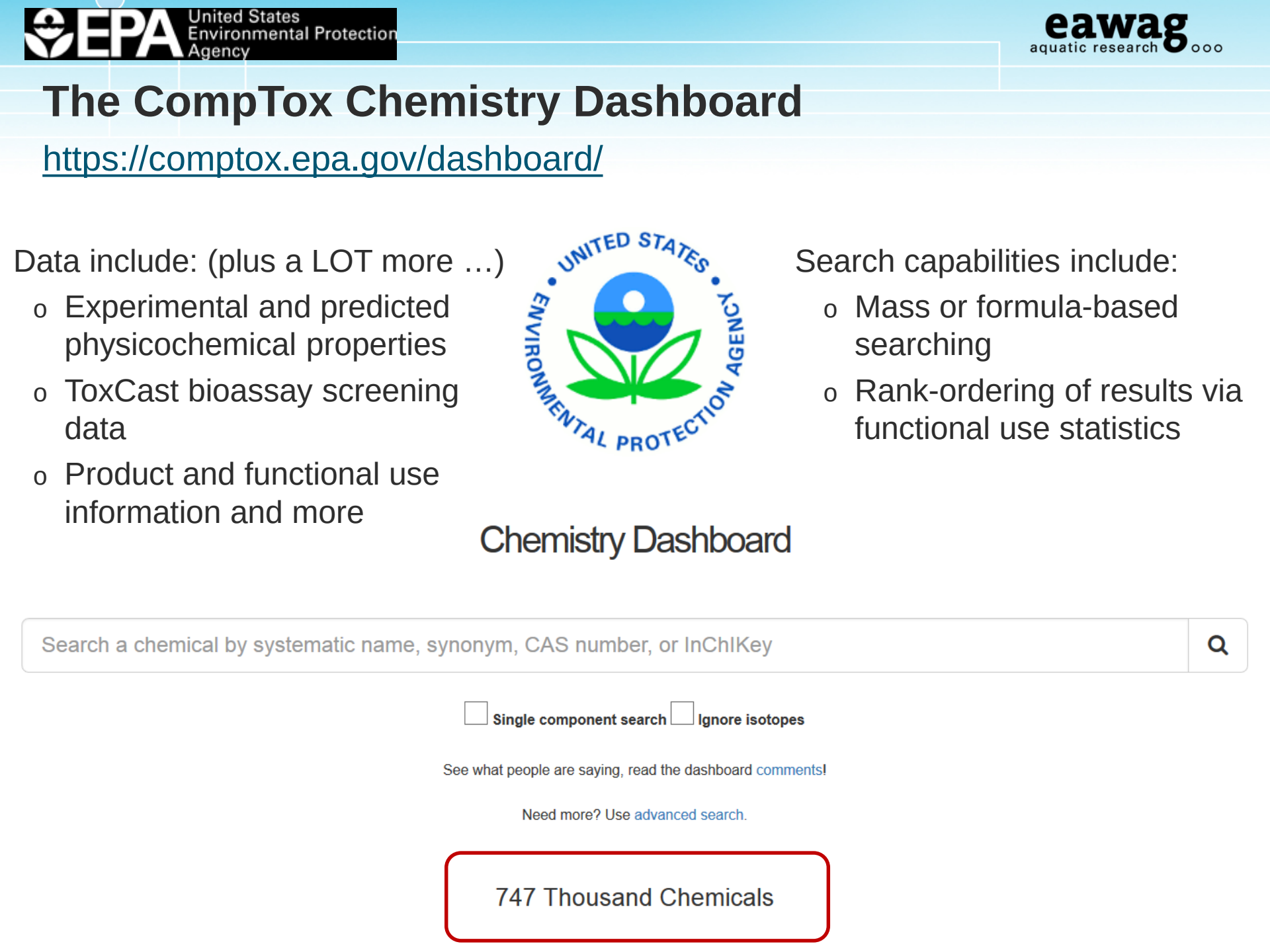
https://comptox.epa.gov/dashboard/

Data include: (plus a LOT more …)

- Experimental and predicted physicochemical properties
- ToxCast bioassay screening data
- Product and functional use information and more

Search capabilities include:

- Mass or formula-based searching
- Rank-ordering of results via functional use statistics



The Dashboard in brief – Example PFOS

https://comptox.epa.gov/dashboard/

PFOS

1763-23-1 | DTXSID3031864

Searched by Approved Name: Found 1 result for 'PFOS'.

Chemical Properties

- Summary
- LogP: Octanol-Water
- Water Solubility
- Density
- Melting Point
- Boiling Point
- Surface Tension
- Vapor Pressure
- LogKoa: Octanol-Air
- Henry's Law
- Index of Refraction
- Molar Refractivity
- pKa Acidic Apparent

Wikipedia

Perfluorooctanesulfonic acid (conjugate base perfluorooctanesulfonate) (PFOS) is an anthropogenic fluorosurfactant and global pollutant. PFOS was the key ingredient in Scotchgard, a fabric protector made by 3M, and numerous stain repellents. It was added to

Download as: TSV | Excel | SDF

Property	Average		Median		Range		Unit
	Experimental	Predicted	Experimental	Predicted	Experimental	Predicted	
LogP: Octanol-Water	-	4.44 (4)	-	4.44	-	2.32 to 6.28	-
Water Solubility	-	2.41e-03 (4)	-	2.41e-03	-	6.25e-09 to 9.12e-03	mol/L
Density	-	1.84 (1)	-	1.84	-	-	g/cm^3
Melting Point	-	65.5 (3)	-	65.5	-	51.9 to 73.5	°C
Boiling Point	145 (1)	237 (3)	145	237	145	218 to 262	°C
Surface Tension	-	19.6 (1)	-	19.6	-	-	dyn/cm
Vapor Pressure	-	7.87e-03 (2)	-	7.87e-03	-	7.36e-04 to 1.50e-02	mmHg
LogKoa: Octanol-Air	-	4.75 (1)	-	4.75	-	-	-
Henry's Law	-	2.27e-10 (1)	-	2.27e-10	-	-	atm-m3/mole
Index of Refraction	-	1.30 (1)	-	1.30	-	-	-
Molar Refractivity	-	51.5 (1)	-	51.5	-	-	cm^3
pKa Acidic Apparent	-	-3.27 (1)	-	-3.27	-	-	-
Molar Volume	-	272 (1)	-	272	-	-	cm^3
Polarizability	-	20.4 (1)	-	20.4	-	-	Å^3

The Dashboard in brief – Example PFOS

https://comptox.epa.gov/dashboard/

Exposure Limit

Regulatory Toxicity Val...

Effect Level

Download as: TSV Excel

PFOS

1763-23-1 | DTX

Searched b

Exposure Limit

Grouping ID	Priority	Type	Subtype	Value	Units	Study Type	Exposure Route	Study Duration	Species	Media	Details	Source
174454	5	water q...	ground...	0.3	ug/L	-	-	-	-	ground...	Minnes...	ACToR
174455	2	water q...	ground...	860	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174456	2	water q...	ground...	610	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174457	2	water q...	ground...	110	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174458	2	water q...	ground...	80.0	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174459	2	water q...	ground...	5.80e-02	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174460	2	water q...	ground...	0.02	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174461	2	water q...	ground...	5.80e-04	mg/L	-	-	-	-	ground...	Texas ...	ACToR
174462	2	water q...	ground...	2.00e-04	mg/L	-	-	-	-	ground...	Texas ...	ACToR

Related Compounds (Beta)

Presence in Lists

Record Information

Chemical Properties | Env. Fate/Transport | Synonyms | External Links | Toxicity Values (Beta) | Exposure | Bioassays | Similar Molecules (Beta) | Literature | Comments

Collaboration on Chemical Curation of Lists

Pharmaceutical List with Consumption Data	SwissPharma_TableS2.csv	SwissPharma_TableS2_InChIKeys.txt	Singer *et al.* 2016. DOI: **10.1021/acs.est.5b03332**
Swiss Insecticides, Fungicides and TPs	SwissPesticides_TableS1.csv	SwissPesticides_TableS1_InChIKeys.txt	Moschet *et al.* 2013. DOI: **10.1021/ac4021598**
NormaNEWS for retrospective screening of new emerging contaminants	NormaNEWS_V4_26042017.csv	NormaNEWS_V4_InChIKeys.txt	**NormaNEWS** list provided by Nikiforos Alygizakis, Saer Samanipour and Kevin Thomas
Combined Inventory of Ingredients Employed in Cosmetic Products (2000) and Revised Inventory (2006)	Merged_CosmeticProducts_04052017.csv	Merged_CosmeticProducts_04052017_InChIKeys.txt	The scientific committee on cosmetic products and non-food products Intended for consumers - **SCCNFP/0389/00 Final** and Commission **Decision 2006/257/EC** amending the Decision 96/335/EC. Provided by Peter von der Ohe, UBA, curated by Reza Aalizadeh, University of Athens
PFAS Highly fluorinated substances list: KEMI	PFAS_Market_Keml_EPA_1Feb2017.xlsx ~2,600 PFAS	Curation in progress: coming soon	Appendix 2 from Swedish Chemicals Agency **KEMI Report 7/15**. Provided by Stellan Fischer, KEMI
NORMAN Priority List 2015	NORMAN_PriorityList_2016.csv Further curation in progress...	NORMAN_PriorityList_2016_InChIKeys.txt	Priority substances from NORMAN WG-1 (Prioritisation), provided by Valeria Dulio
French Monitoring List	French_List_08052017.csv Further curation in progress...	FrenchList_UniqueInChIKeys_08052017.txt	Provided by Valeria Dulio, curated by Reza Aalizadeh, University of Athens
KEMI Market List	KEMI_MarketList_12052017_MSready.xlsx	KEMI_MarketList_12052017_MSready_InChIKeys.txt	Provided by Stellan Fischer, KEMI including Hazard and Exposure scores, documented **here**. Curated by Reza Aalizadeh, University of Athens.
TSCA Surfactants	Coming soon...		Provided by Lee Ferguson, sourced from James Little

KEMI PFAS List

PFAS Highly fluorinated substances list: KEMI	**PFAS_Market_KemI_EPA_1Feb2017.xlsx** ~2,600 PFAS	Curation in progress: coming soon	Appendix 2 from Swedish Chemicals Agency **KEMI Report 7/15**. Provided by Stellan Fischer, KEMI

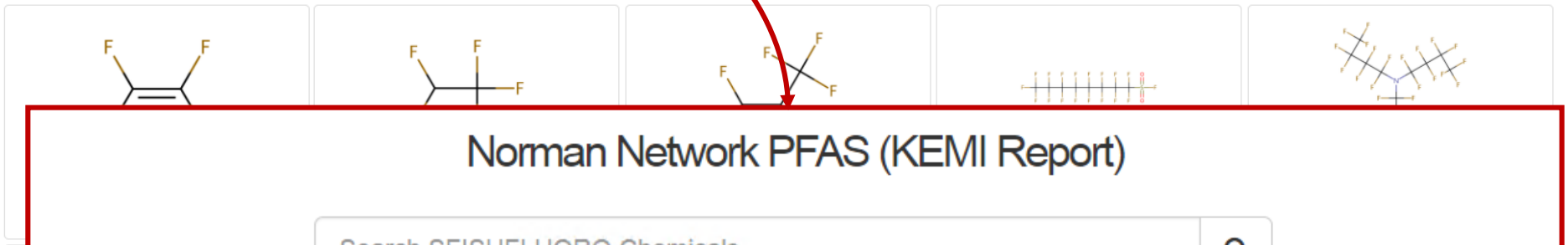
Search SFISHFLUORO Chemicals

List Details

Description: This list of perfluorinated substances originated from Appendix 2 from Swedish Chemicals Agency Report 7/15 (available at http://www.kemi.se/en/global/rapporter/2015/report-7-15-occurrence-and-use-of-highly-fluorinated-substances-and-alternatives.pdf) on the occurrence and use of highly fluorinated substances and alternatives (2015). The current KEMI PFAS list includes substances beyond the original report and was provided by Stellan Fischer.

Number of Chemicals: 2257

Sort Options | Select/Deselect All | **Download as:** TSV | Excel | SDF | View Selected



Norman Network PFAS (KEMI Report)

Search SFISHFLUORO Chemicals

List Details

Description: This list of perfluorinated substances originated from Appendix 2 from Swedish Chemicals Agency Report 7/15 on the occurrence and use of highly fluorinated substances and alternatives (2015). The current KEMI PFAS list includes substances beyond the original report and was provided by Stellan Fisher.

Number of Chemicals: 970

NormaNEWS

NormaNEWS for retrospective screening of new emerging contaminants	**NormaNEWS_V4_26042017.csv**	**NormaNEWS_V4_InChIKeys.txt**	**NormaNEWS** list provided by Nikiforos Alygizakis, Saer Samanipour and Kevin Thomas

NORMANews

Search NORMANEWS Chemicals

List Details

Description: The Norman Early Warning System (NormaNEWS: http://www.norman-network.com/?q=node/244) is a pilot network designed to investigate the spatial and temporal distribution of newly identified contaminants of emerging concern in environmental samples through performing retrospective suspect screening on HRMS data acquired using different instrumental platforms and data processing software. The NormaNEWS pilot study was performed through recruiting eight reference laboratories with available archived HRMS data with the goal of exploring the potential of an early warning network to rapidly establish the occurrence of newly-identified contaminants of emerging concern across Europe and beyond, through the use of retrospective suspect screening employing HRMS. The pilot study was referred to as the Norman Early Warning System, abbreviated to NormaNEWS.

Number of Chemicals: 131

NormaNEWS: Norman Early Warning System

Search NORMANEWS Chemicals

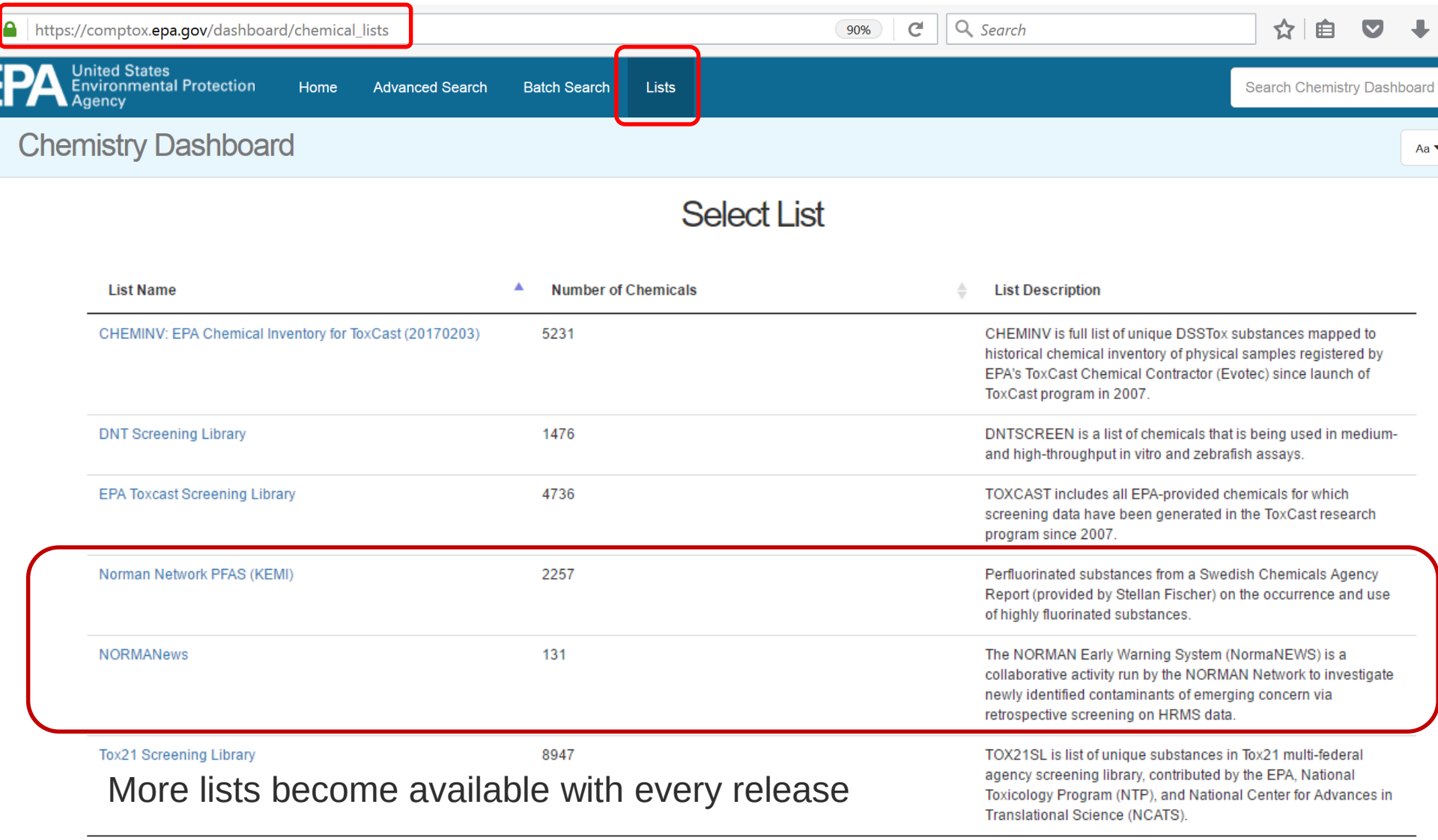
List Details

Description: The Norman Early Warning System (NormaNEWS) is a pilot network designed to investigate the spatial and temporal distribution of newly identified contaminants of emerging concern in environmental samples through performing retrospective suspect screening on HRMS data acquired using different instrumental platforms and data processing software. The NormaNEWS pilot study was performed through recruiting eight reference laboratories with available archived HRMS data with the goal of exploring the potential of an early warning network to rapidly establish the occurrence of newly-identified contaminants of emerging concern across Europe and beyond, through the use of retrospective suspect screening employing HRMS. The pilot study was referred to as the Norman Early Warning System, abbreviated to NormaNEWS.

Number of Chemicals: 131

List Functionality in the Dashboard

An overview of all the lists …



More lists become available with every release

The Dashboard in brief – Example PFOS

https://comptox.epa.gov/dashboard/

Wikipedia

Intrinsic Properties

Structural Identifiers

Related Compounds (Beta)

Presence in Lists

DNT Screening Library | CHEMINV: EPA Chemical Inventory for ToxCast (20170203) | EPA Toxcast Screening Library | Tox21 Screening Library | NORMANews

Norman Network PFAS (KEMI)

Record Information

or result from the degradation of precursors. PFOS levels that have been detected in wildlife... Read more

Intrinsic Properties

Structural Identifiers

Related Compounds (Beta)

Presence in Lists

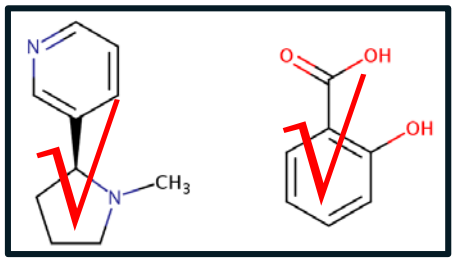
Record Information

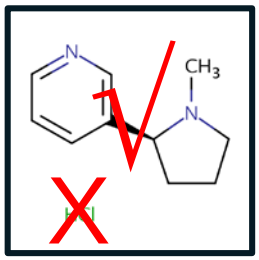
Chemical Properties | Env. Fate/Transport | Synonyms | External Links | Toxicity Values (Beta) | Exposure | Bioassays | Similar Molecules (Beta) | Literature | Comments

This is only the beginning … future challenges:

Huge progress in a short time – but much more to follow

- Mixture identification and curation





- Progressive curation – error detection and removal (early days!)
- Progressive registration of additional substances
  - Contributions of additional lists are welcome!
- Consolidation of the “MS-ready” concept – consistency between resources
- Treatment of UVCBs: **U**nknown or **V**ariable composition, **C**omplex reaction products or **B**iological materials
  - https://comptox.epa.gov/dashboard/dsstoxdb/results?utf8=✓&search=C10-12+chloroalkanes

Acknowledgements

National and Kapodistrian UNIVERSITY OF ATHENS
Reza Aalizadeh
Nikos Thomaidis

Stellan Fischer,

Jaroslav Slobodnik, Natalia Glowacka, Lubos Cirka, Ildiko Ipolyi, Nikiforos Alygizakis & more at EI




M. Stravs, E. Müller, T. Schulze, S. Neumann

“Suspect Exchange” task partners:

Tony Williams, Andrew McEachran, Jon Sobus, US EPA

C. Ruttkies, S. Wolf, S. Neumann

Questions?

NORMAN Resources:

www.massbank.eu

http://www.norman-network.com/datatable/

http://www.norman-network.com/?q=node/236

CompTox Chemistry Dashboard:

https://comptox.epa.gov/

Contact:

emma.schymanski@eawag.ch

EU Grant 603437