http://www.orcid.org/0000-0002-2668-4821

The EPA Comptox Chemistry Dashboard: A Web-Based Data Integration Hub for Environmental Chemistry and Toxicology Data

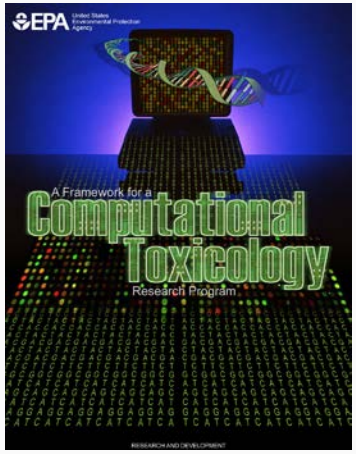
Antony Williams

U.S. Environmental Protection Agency, RTP, NC

February 13th 2018

Research Triangle Institute, RTP, NC

National Center for Computational Toxicology



- National Center for Computational Toxicology established in 2005 to integrate:
  - High-throughput and high-content technologies
  - Modern molecular biology
  - Data mining and statistical modeling
  - Computational biology and chemistry
- Currently staffed by ~60 employees as part of EPA's Office of Research and Development
- Home of ToxCast & ExpoCast research efforts
- Key partner in U.S. Tox21 federal consortium
- Multiple cross-division collaborations (e.g. NERL, OPP, OPPT)

The NCCT Team

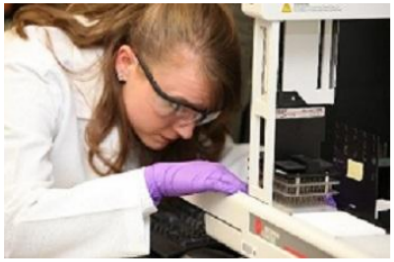
- Our team is broad in skills
  - Biologists, chemists and toxicologists
  - Bioinformaticians and cheminformaticians - modelers (QSAR, Deep Learning), text miners
  - Information technology team, software developers
  - A large IT support team for production applications (National Computing Center)

The need for data and derivative models and algorithms

- NCCT outputs: include a lot of data, models, algorithms and software applications
- We produce Open Data – we want people to interrogate it, learn from it, develop understanding

Toxicity Forecasting

Advancing the Next Generation of Chemical Evaluation



EPA needs rapid and efficient methods to prioritize, screen and evaluate thousands of chemicals. EPA's Toxicity Forecaster (ToxCast) generates data and predictive models on thousands of chemicals of interest to the EPA. ToxCast uses high-throughput screening methods and computational toxicology approaches to rank and prioritize chemicals. In fact, EPA's Endocrine Disruption Screening Program (EDSP) is working to use ToxCast to rank and prioritize chemicals.

- ToxCast has data on over 1,800 chemicals from a broad range of sources including industrial and consumer products, food additives, and potentially "green" chemicals that could be safer alternatives to existing chemicals.
- ToxCast screens chemicals in over 700 high-throughput assays that cover a range of high-

Downloadable Computational Toxicology Data

EPA's computational toxicology research efforts evaluate the potential health effects of thousands of chemicals. The process of evaluating potential health effects involves generating data that investigates the potential harm, or hazard of a chemical, the degree of exposure to chemicals as well as the unique chemical characteristics.

As part of EPA's commitment to share data, all of the computational toxicology data is publicly available for anyone to access and use.

High-throughput Screening Data

EPA researchers use rapid chemical screening (called high-throughput screening assays) to limit the number of laboratory animal tests while quickly and efficiently testing thousands of chemicals for potential health effects.

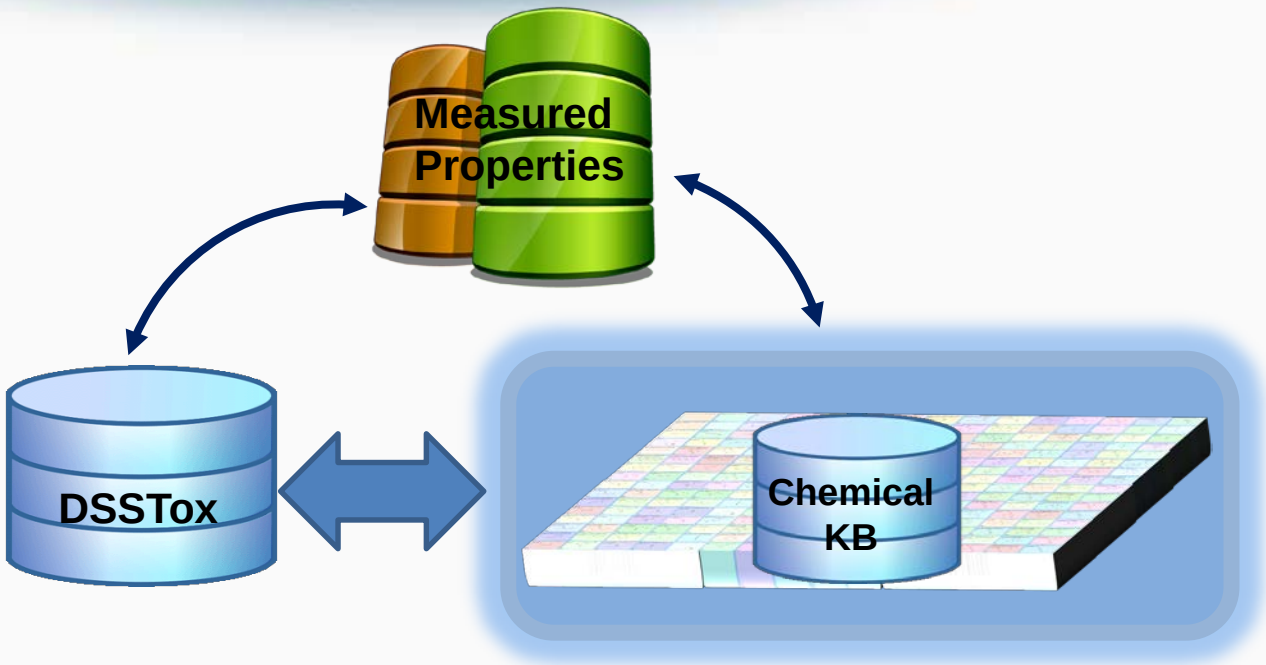
- ToxCast Data: High-throughput screening data on thousands of chemicals.

Rapid Exposure and Dose Data

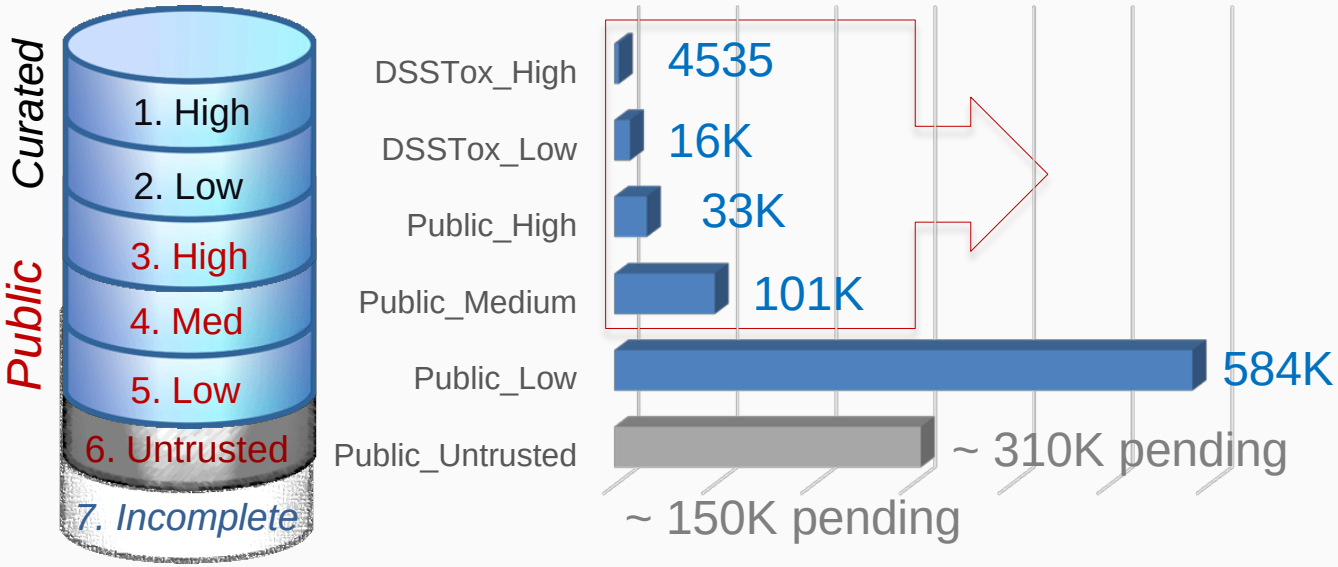
EPA researchers develop and use rapid exposure estimates to predict potential exposure for thousands of chemicals.

- High-throughput toxicokinetics data: It is important to link the external dose of a chemical to an internal blood or tissue concentration, this process is called toxicokinetics. EPA researchers measure the critical factors that determine the distribution

Underpinning with chemicals



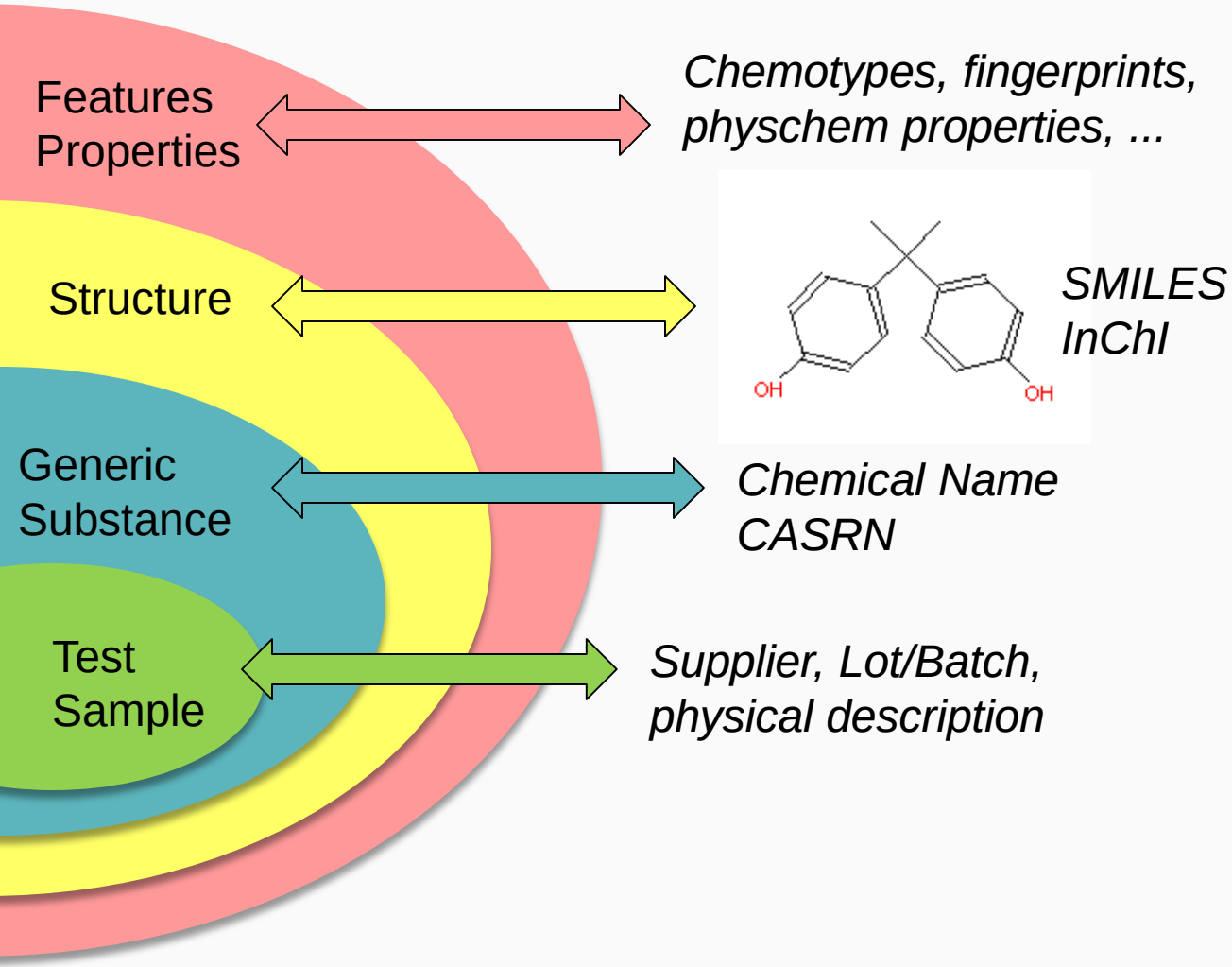
Approximately 15 Years of Data… Growing with daily curation



QC Levels

DSSTox_High:	Hand curated and validated
DSSTox_Low:	Hand curated and confirmed using multiple public sources
Public_High:	Extracted from EPA SRS and confirmed to have no conflicts in ChemID and PubChem
Public_Medium:	Extracted from ChemID and confirmed to have no conflicts in PubChem
Public_Low:	Extracted from ACToR or PubChem
Public_Untrusted:	Postulated, but found to have conflicts in public sources

Chemical representation levels supporting data integration



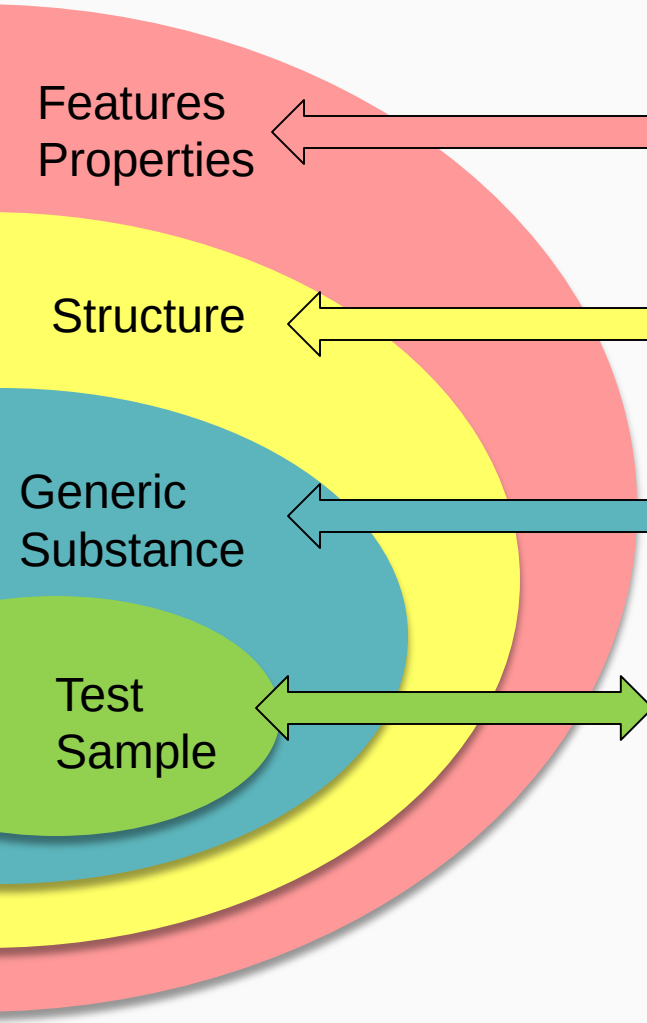
Chemotypes, fingerprints, physchem properties, ...

SMILES
InChI

Chemical Name
CASRN

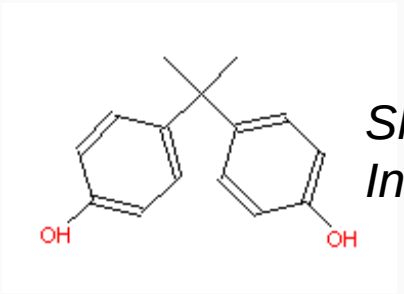
Supplier, Lot/Batch, physical description

Chemical representation levels supporting data integration



Chemotypes, fingerprints, physchem properties, ...

Chemical analogs, Read-across, SAR modeling



SMILES
InChI

Structure searching & modeling

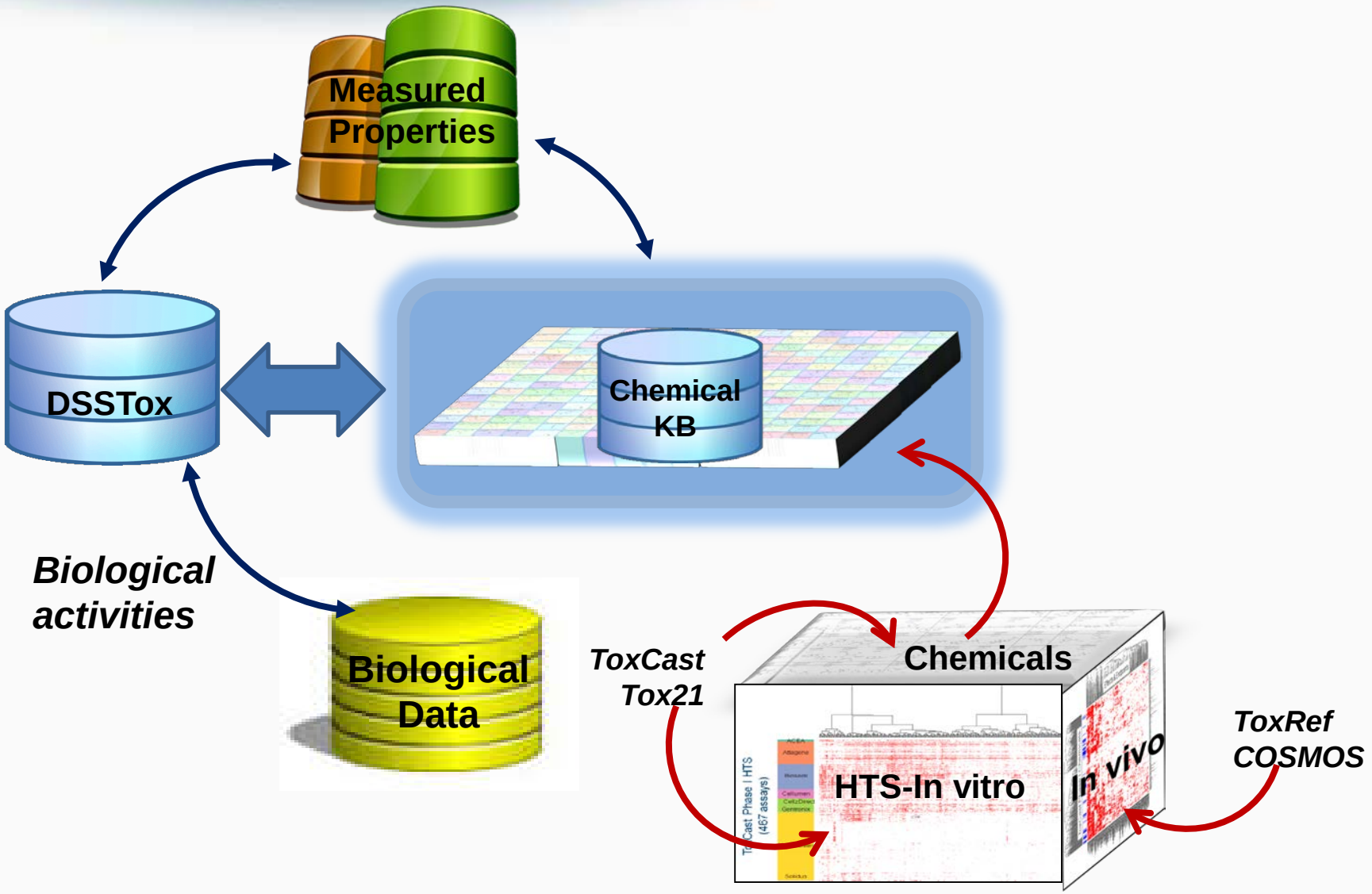
Chemical Name
CASRN

Public toxicity datasets

Supplier, Lot/Batch, physical description

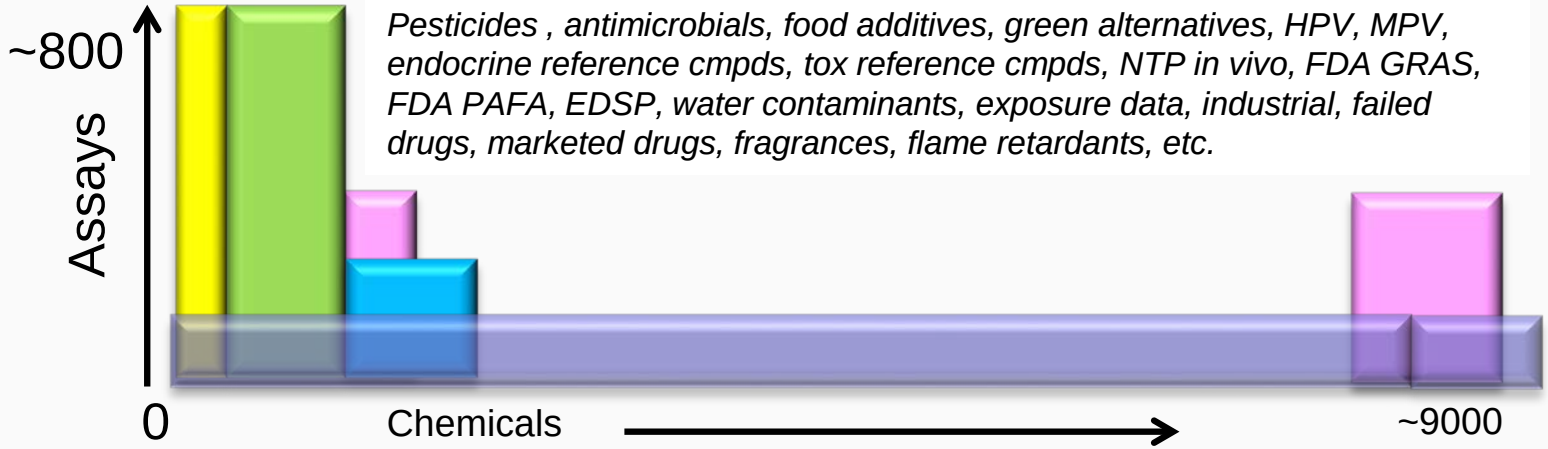
Experimental Endpoint Data

Integrating *in vitro* and *in vivo* data

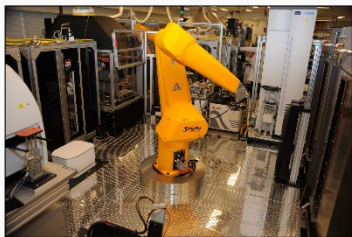


TEN YEARS of Assay Measurements: ToxCast & Tox21

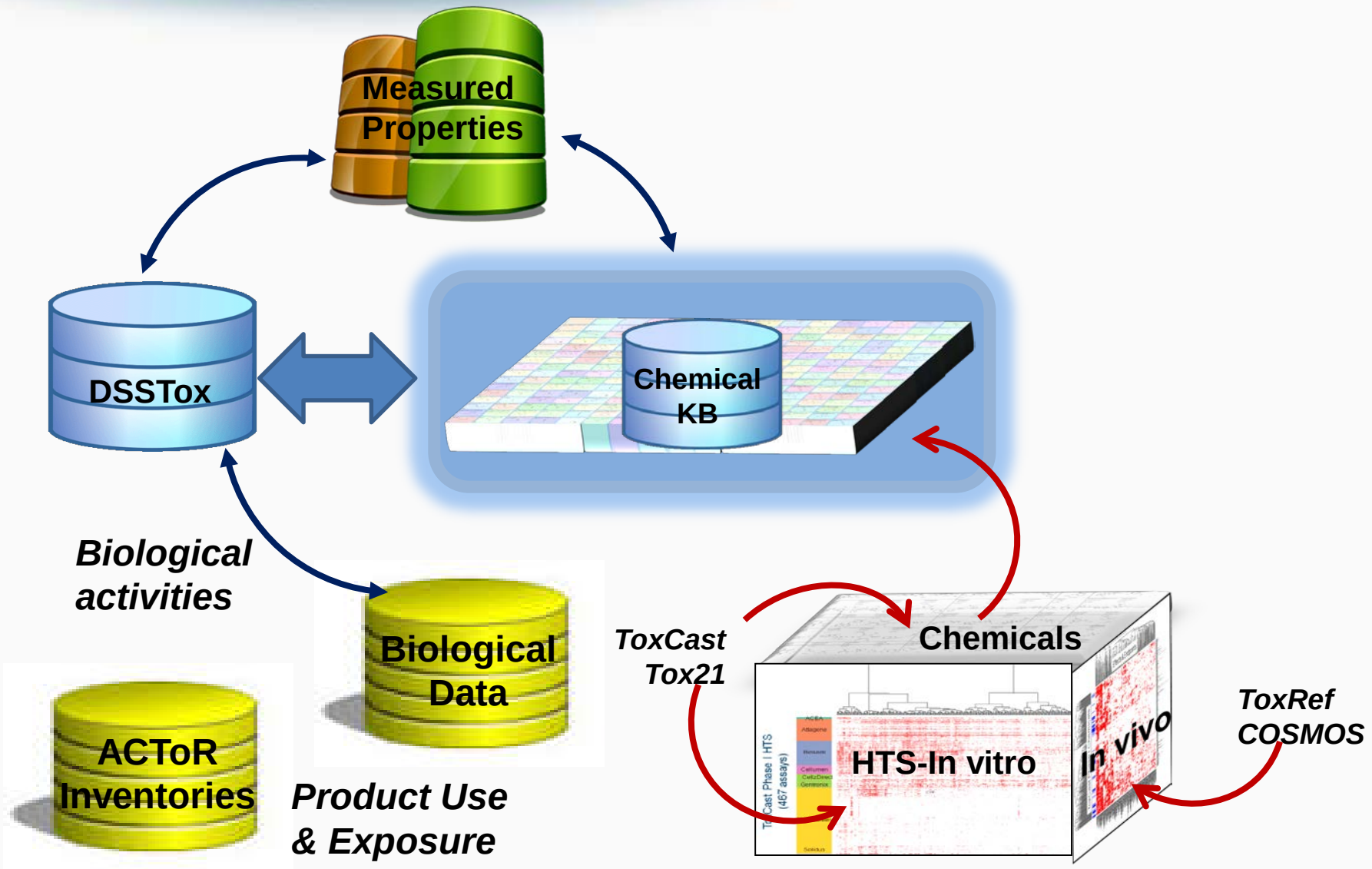
Set	Chemicals	Assays	Endpoints	Completion
ToxCast Phase I	293	~600	~700	2011
ToxCast Phase II	767	~600	~700	03/2013
ToxCast E1K	800	~50	~120	03/2013
ToxCast Phase III	~900	~300	~300	In progress
Tox21	~9000	~80	~150	In progress



Pesticides , antimicrobials, food additives, green alternatives, HPV, MPV, endocrine reference cmpds, tox reference cmpds, NTP in vivo, FDA GRAS, FDA PAFA, EDSP, water contaminants, exposure data, industrial, failed drugs, marketed drugs, fragrances, flame retardants, etc.



Adding Product Use and Exposure



High Throughput Measurement to Identify **Exposure**

EPA United States Environmental Protection Agency

EPA US Environmental Protection Agency

Español | 中文: 繁體版 | 中文: 简体版 | Tiếng Việt | 한국어

Learn the Issues | Science & Technology | Laws & Regulations | About EPA

Search EPA.gov

Related Topics: Safer Chemicals Research

Contact Us Share

Rapid Chemical Exposure and Dose Research

EPA is responsible for ensuring the safety of thousands of chemicals. Quantitative exposure data are available for only a small fraction of registered chemicals. This type of exposure data is needed to thoroughly evaluate chemicals for potential risks to humans, wildlife and ecosystems. EPA is developing innovative methods to develop exposure estimates for thousands of chemicals to better protect human health and the environment. These innovative methods are called rapid exposure and dose assessments.

Rapid Exposure Predictions

Rapid, also called high-throughput, exposure predictions or ExpoCast provide rapid exposure estimates for thousands of chemicals. ExpoCast quickly and efficiently looks at multiple routes of exposure to provide exposure estimates. ExpoCast uses and enhances two well-known exposure models to estimate chemical exposure.

Pictured Above: Farfield Exposure Examples

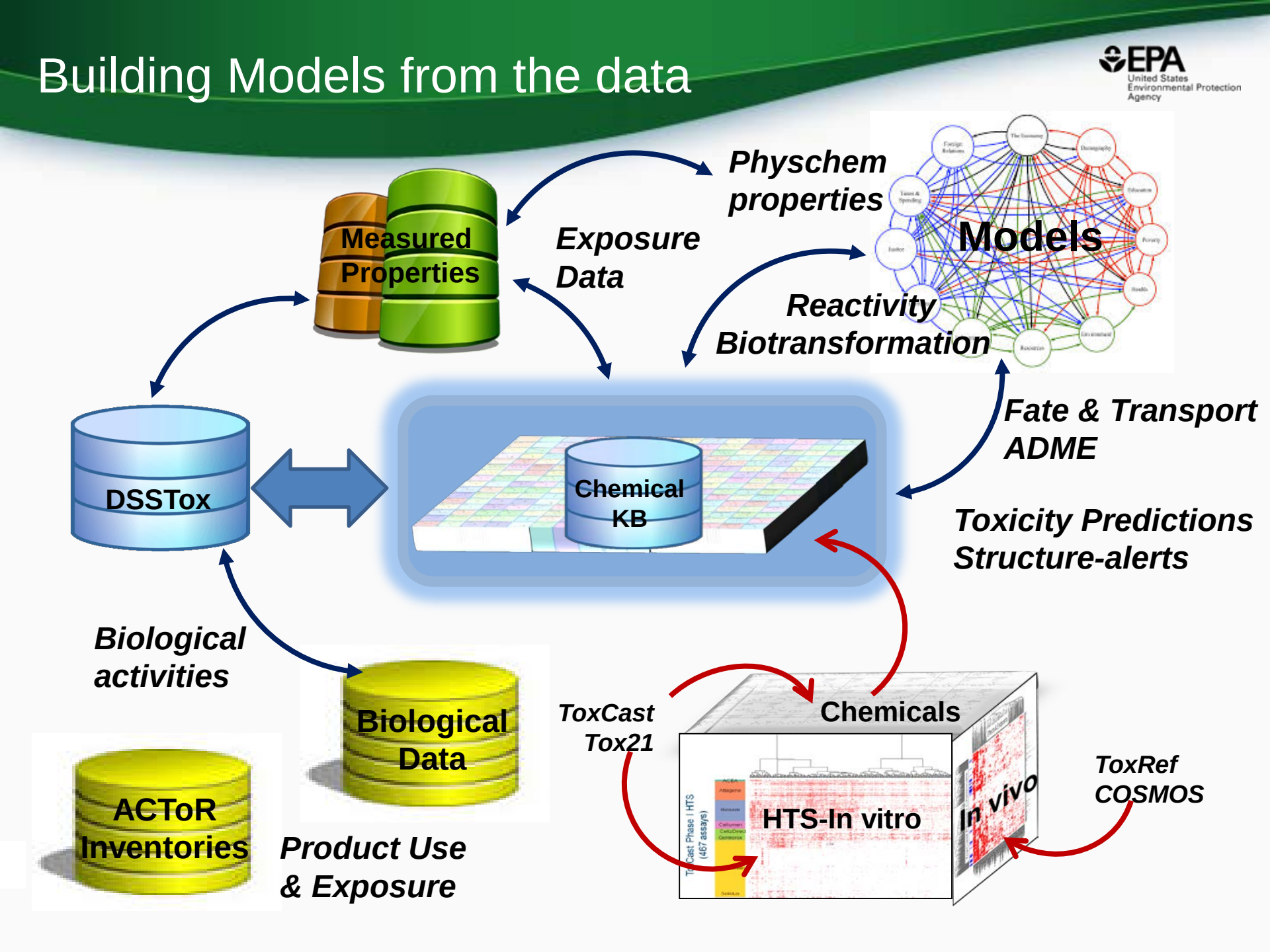
› Farfield Exposure Models

› Nearfield Exposure Models

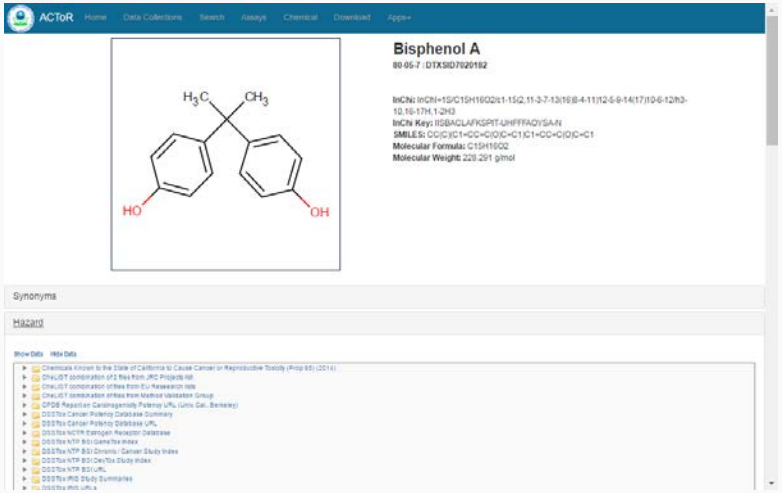
Evaluating High-throughput Exposure Predictions

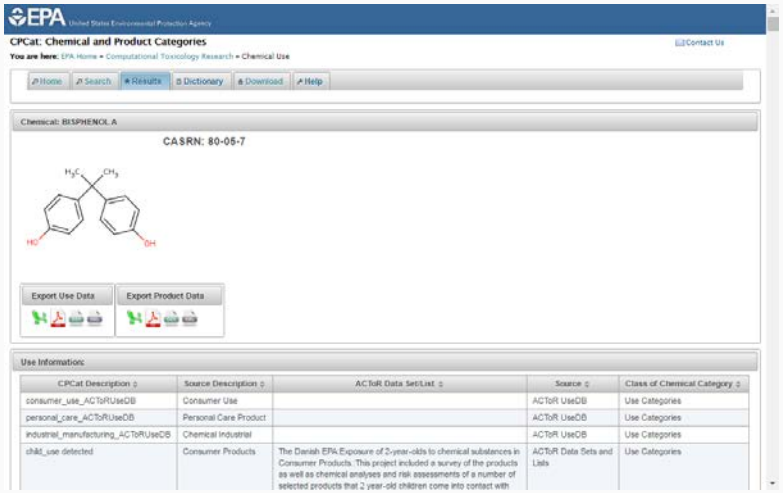
EPA is currently evaluating the effectiveness of high-throughput exposure models

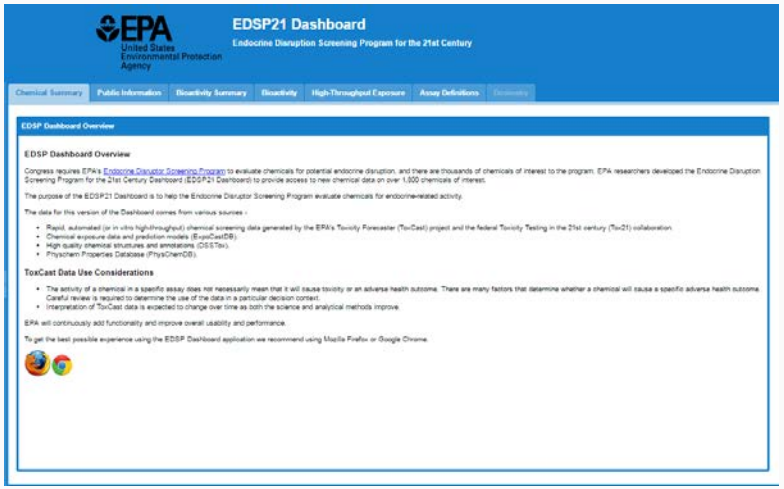
Building Models from the data



Multiple Dashboards Delivered

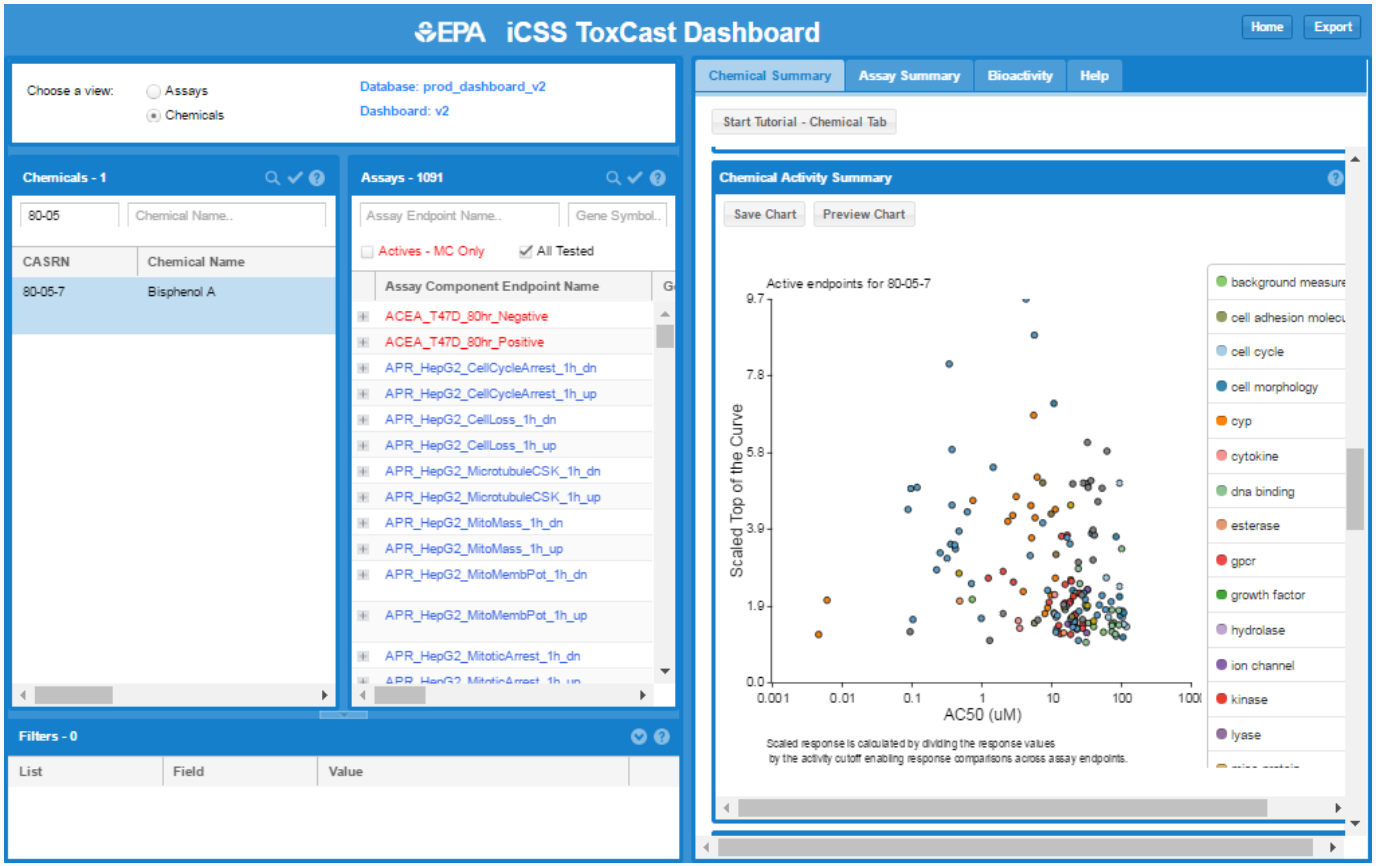






Toxcast Dashboard
https://actor.epa.gov/dashboard/

- Access and Interrogate chemical screening data from ToxCast and the Tox21 collaboration

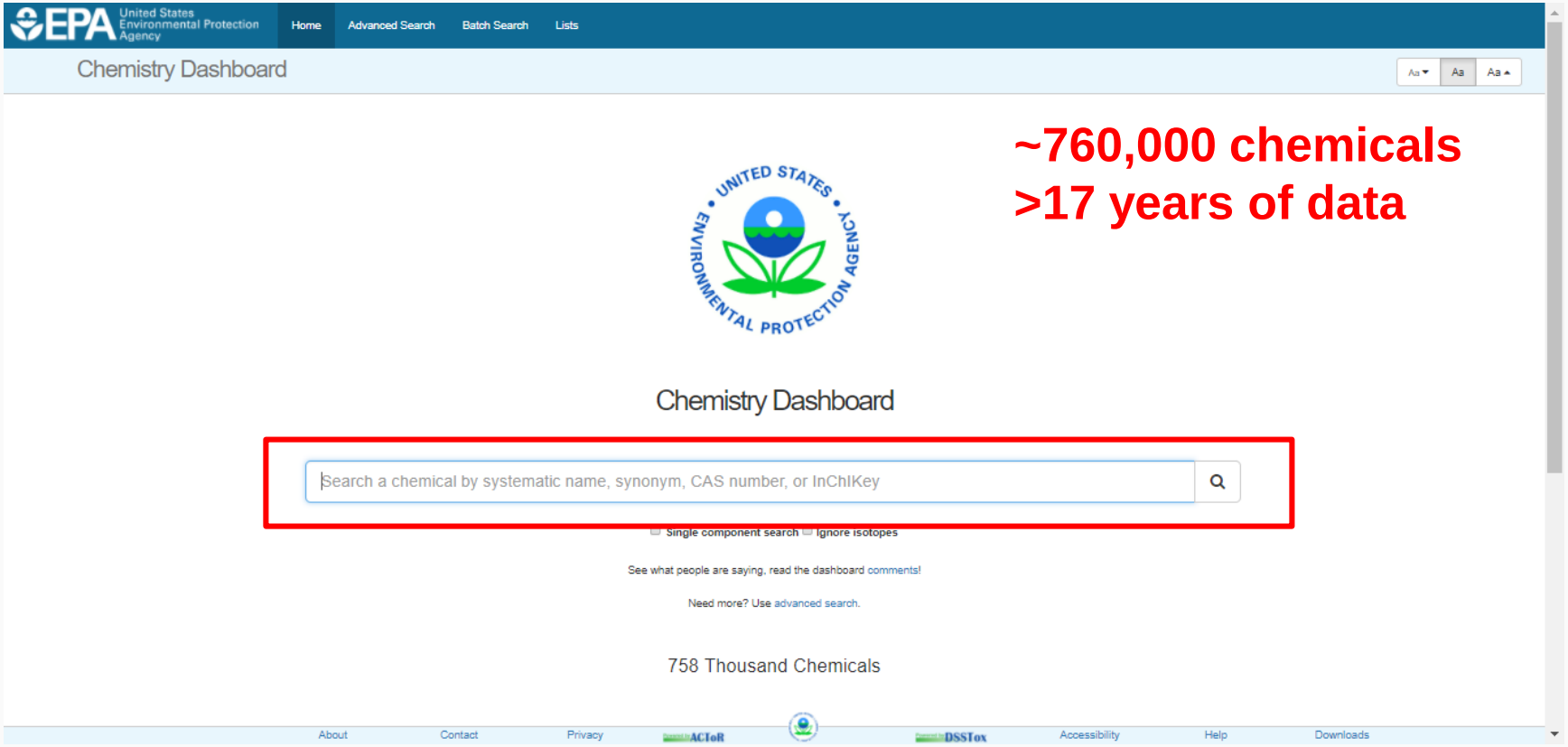


What we have learned…

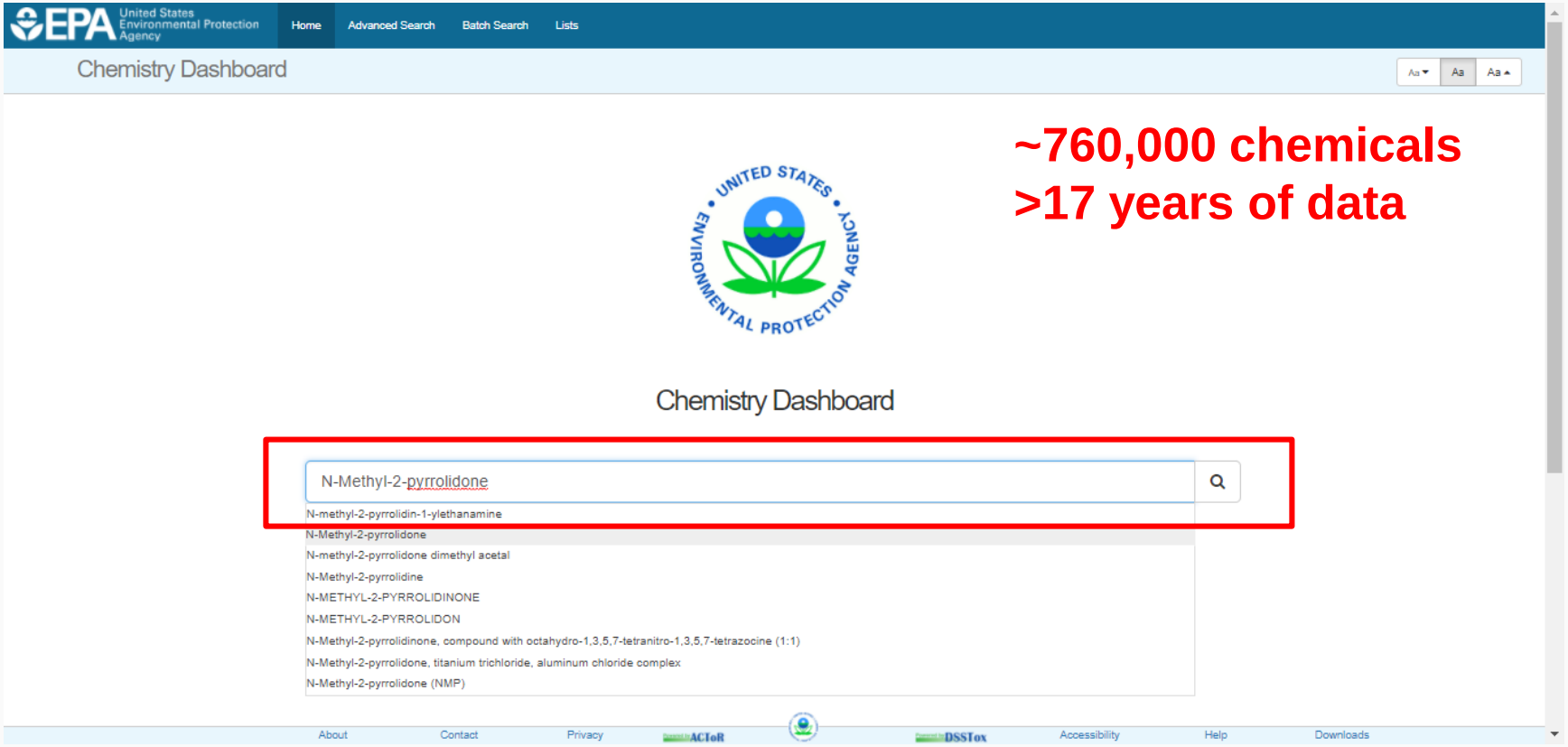
- Data curation, standardization and versioning is **essential**
- **Prototype** application development suffices **for research** projects
- **Production apps** need managed processes
- ODOSOS (Open Data, Open Source and Open Standards) endows many benefits

- We are building: CompTox Chemistry Dashboard

Comptox Chemistry Dashboard
https://comptox.epa.gov



Comptox Chemistry Dashboard
https://comptox.epa.gov

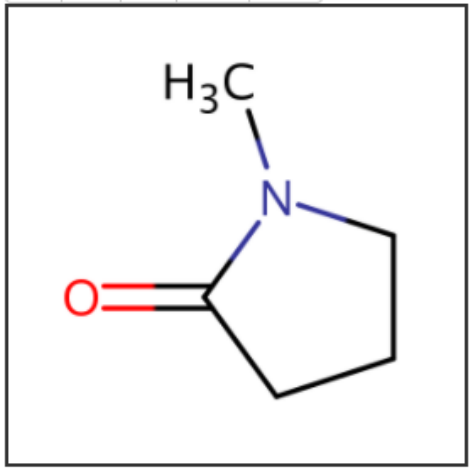


Chemical Page

N-Methyl-2-pyrrolidone

872-50-4 | DTXSID6020856

Searched by Synonym from Valid Source: Found 1 result for 'N-METHYLPYRROLIDONE'.



- Wikipedia
- Intrinsic Properties
  - **Molecular Formula:** C5H9NO — Find All Chemicals
  - **Average Mass:** 99.133 g/mol
  - **Monoisotopic Mass:** 99.068414 g/mol
- Structural Identifiers
- Related Compounds
- Presence in Lists
- Record Information

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

Chemical Properties

EPA United States Environmental Protection Agency

- Summary
- LogP: Octanol-Water
- Water Solubility
- Density
- Flash Point
- Melting Point
- Boiling Point
- Surface Tension
- Thermal Conductivity
- Vapor Pressure
- Viscosity
- LogKoa: Octanol-Air

Download as: TSV | Excel | SDF

Property	Average		Median		Range		Unit
	Experimental	Predicted	Experimental	Predicted	Experimental	Predicted	
LogP: Octanol-Water	-0.380 (1)	-0.329 (5)	-0.380	-0.329	-0.380	-0.494 to -0.110	-
Water Solubility	10.1 (1)	6.68 (4)	10.1	6.68	10.1	1.48 to 12.8	mol/L
Density	-	1.02 (2)	-	1.02	-	1.01 to 1.03	g/cm^3
Flash Point	-	75.7 (2)	-	75.7	-	65.2 to 86.1	°C
Melting Point	-23.8 (8)	1.32e-01 (4)	-24.0	1.32e-01	-24.0 to -23.0	-10.2 to 25.9	°C
Boiling Point	203 (6)	199 (5)	204	199	202 to 204	191 to 202	°C
Surface Tension	-	33.8 (1)	-	33.8	-	-	dyn/cm
Thermal Conductivity	-	158 (1)	-	158	-	-	mW/(m*K)
Vapor Pressure	3.45e-01 (1)	5.21e-01 (4)	3.45e-01	5.21e-01	3.45e-01	1.71e-01 to 9.99e-01	mmHg
Viscosity	-	3.61 (1)	-	3.61	-	-	cP
LogKoa: Octanol-Air	-	3.84 (1)	-	3.84	-	-	-
Henry's Law	3.20e-09 (1)	9.15e-09 (1)	3.20e-09	9.15e-09	-	-	atm-m3/mole
Index of Refraction	-	1.47 (1)	-	1.47	-	-	-

Available Properties

- Solubility
- Melting Point
- Boiling Point
- LogP (Octanol-water partition coefficient)
- Atmospheric Hydroxylation Rate
- LogBCF (Bioconcentration Factor)
- Biodegradation Half-life
- Henry's Law Constant
- Fish Biotransformation Half-life
- LogKOA (Octanol/Air Partition Coefficient)
- LogKOC (Soil Adsorption Coefficient)
- Vapor Pressure

- Data and models are updated ~90 days

Developing “NCCT Models”

- Our approach to modeling:
  - Obtain **high quality** training sets
  - Apply appropriate modeling approaches
  - **Validate** performance of models
  - Define the applicability domain and model limitations
  - Use models to predict properties across our full datasets
  - Release as **Open Data and Open Models**

Multiple Prediction Algorithms **Transparency** is Important

EPA United States Environmental Protection Agency

- Summary
- **LogP: Octanol-Water**
- Water Solubility
- Density
- Flash Point
- Melting Point
- Boiling Point
- Surface Tension
- Thermal Conductivity
- Vapor Pressure
- Viscosity
- LogKoa: Octanol-Air
- Henry's Law

LogP: Octanol-Water			
	Average	Median	Range
Experimental	-0.380 (1)	-0.380	-0.380
Predicted	-0.329 (5)	-0.329	-0.494 to -0.110

Download as: TSV | Excel | SDF

Experimental

Source	Result
PhysPropNCCT	-0.380

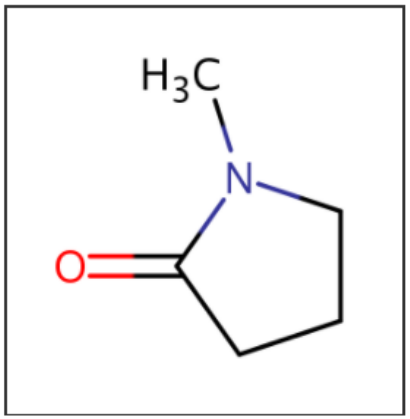
Predicted

Source	Result	Calculation Details	QMRF
EPISUITE	-0.110	Not Available	Not Available
NICEATM	-0.494	Not Available	Available
ACD/Labs Consensus	-0.345	Not Available	Not Available
ACD/Labs	-0.398	Not Available	Not Available
OPERA	-0.300	OPERA Model Report	Available

OPERA Models: LogP: Octanol-Water

N-Methyl-2-pyrrolidone

872-50-4 | DTXSID6020856



Model Results

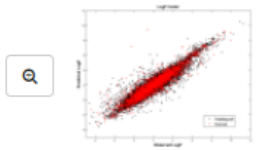
Predicted value: -0.300

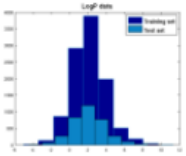
Global applicability domain: Inside

Local applicability domain index: 0.88

Confidence level: 0.81

Model Performance



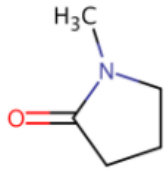


Weighted KNN model

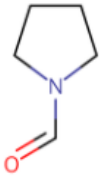
QMRF

5-fold CV (75%)		Training (75%)		Test (25%)	
Q2	RMSE	R2	RMSE	R2	RMSE
0.85	0.69	0.86	0.67	0.86	0.78

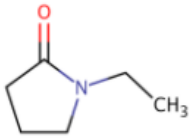
Nearest Neighbors from the Training Set



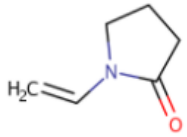
N-Methyl-2-pyrrolidone
Measured: -0.380
Predicted: -0.300



Prolinal
Measured: -0.320
Predicted: -0.226

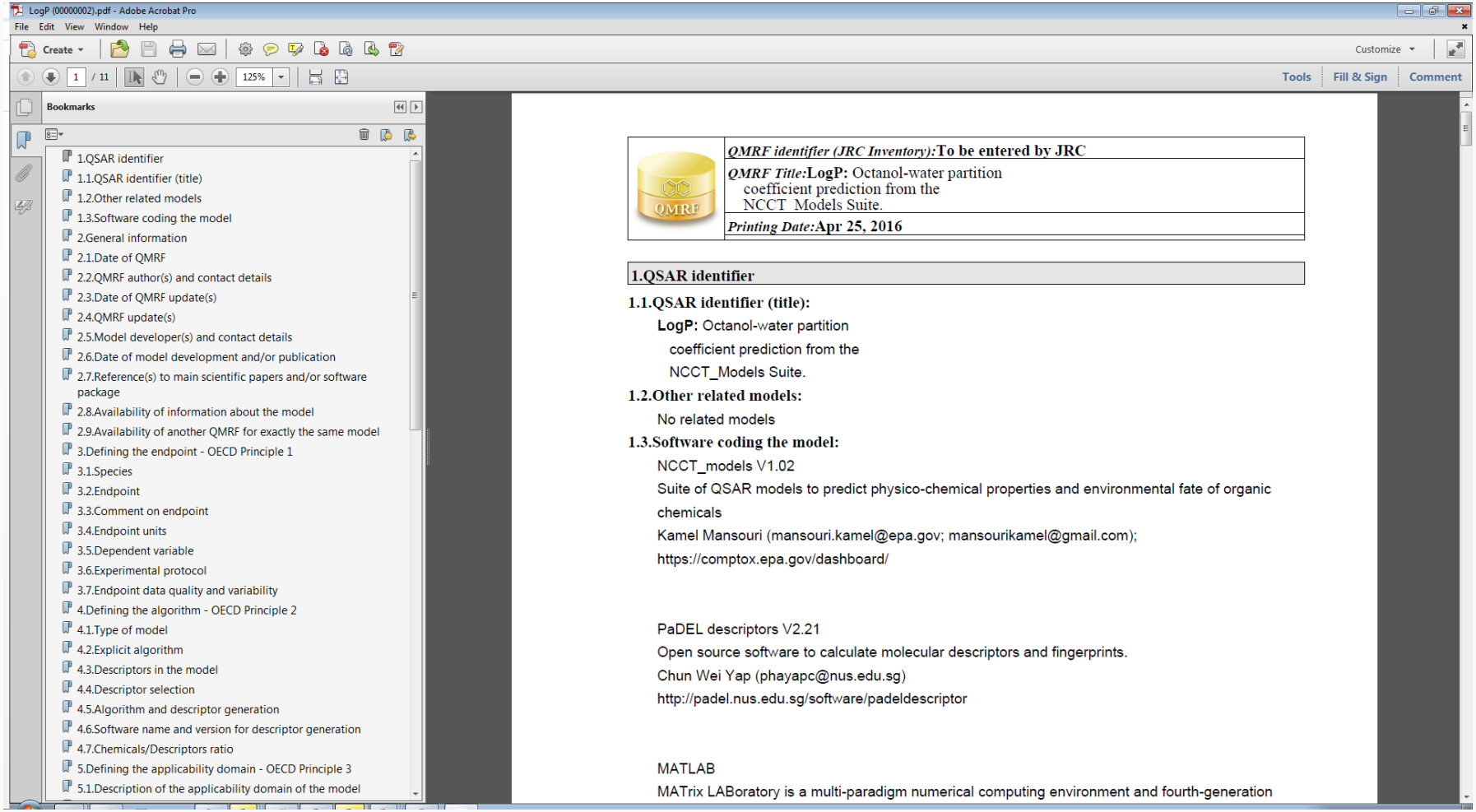


1-Ethyl-2-pyrrolidinone
Measured: -0.0400
Predicted: -0.211



N-Vinyl-2-pyrrolidone
Measured: 0.37
Predicted: 2.65e-01

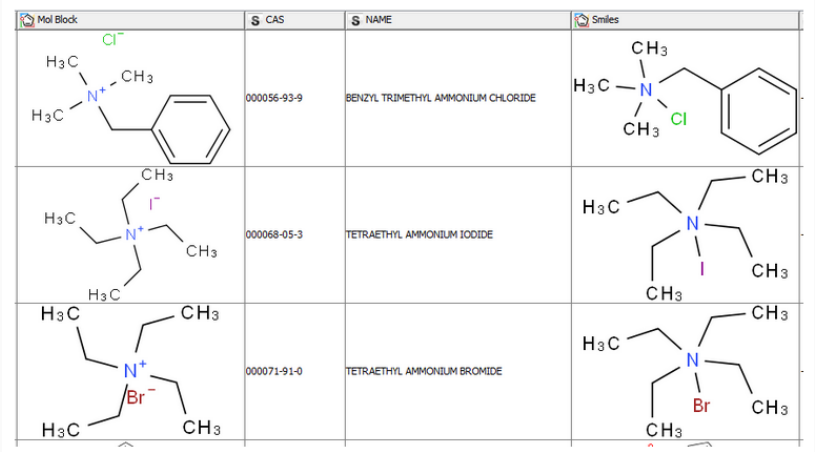
QSAR Modeling Reporting Format



Consuming and Curating Public Data

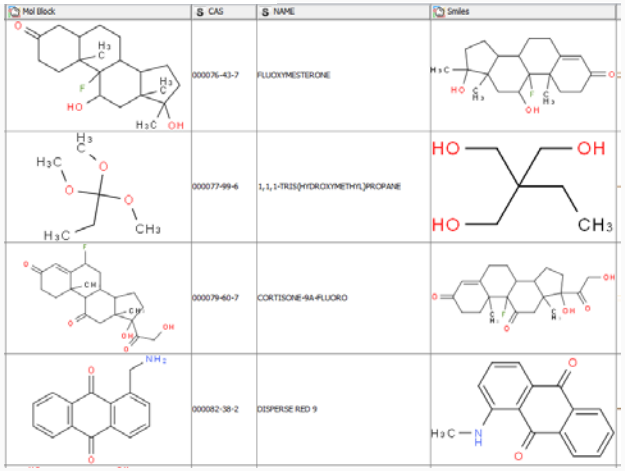
EPA United States Environmental Protection Agency

Public data should be curated prior to modeling



Covalent Halogens

Identical Chemicals



Mismatches

Workflow Details and Data

Journal
SAR and QSAR in Environmental Research ›
Volume 27, 2016 - Issue 11: 17th International Conference on QSAR in Environmental and Health Sciences (QSAR 2016) - Part II. Guest Editors: C.G. Barber and G.J. Myatt

Enter keywords, authors, DOI etc.

258
Views
4
CrossRef citations
16
Altmetric

Articles

An automated curation procedure for addressing chemical errors and inconsistencies in public datasets used in QSAR modelling $

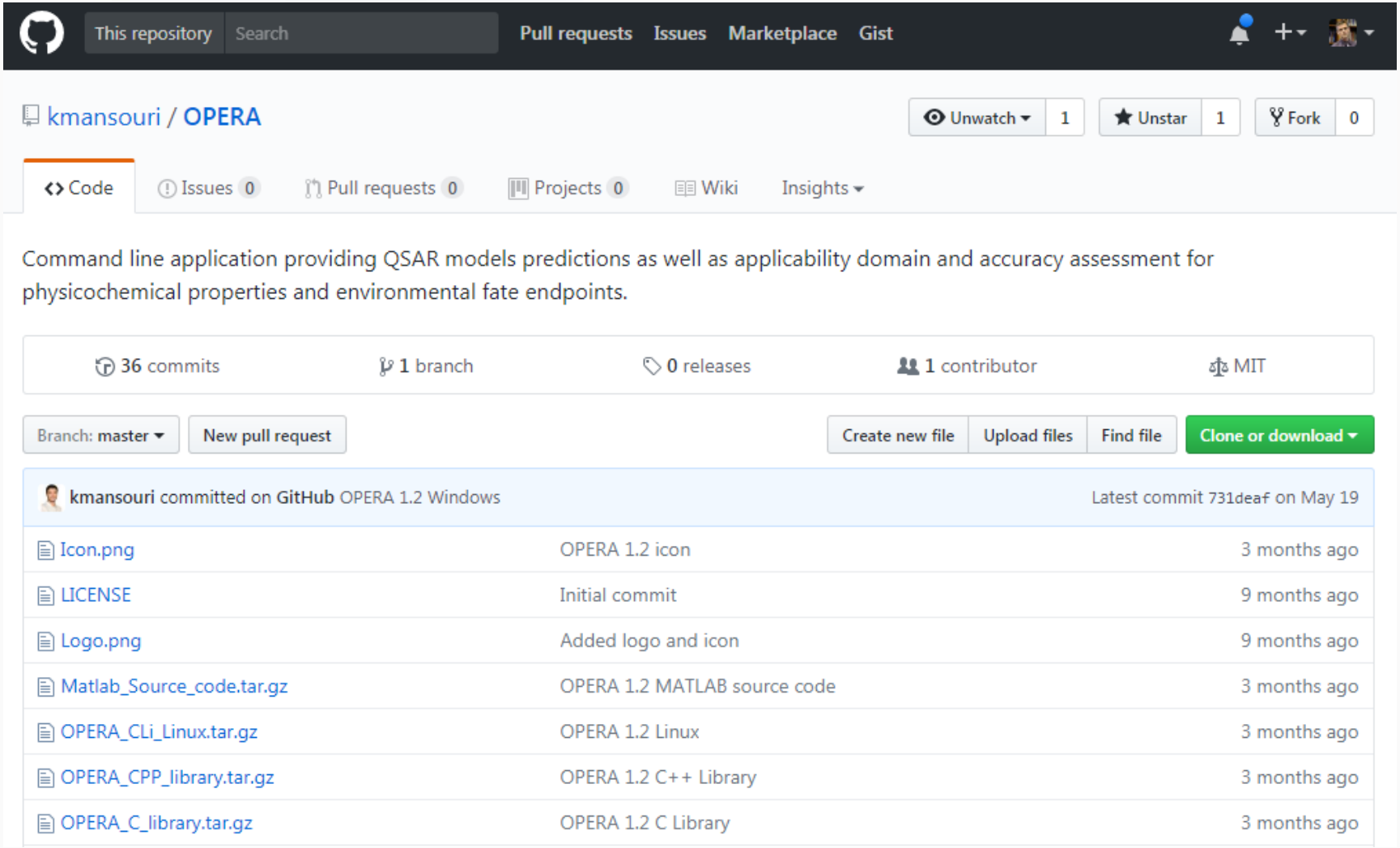
K. Mansouri, C. M. Grulke, A. M. Richard, R. S. Judson & A. J. Williams

Pages 911-937 | Received 03 Sep 2016, Accepted 24 Oct 2016, Published online: 25 Nov 2016

Download citation http://dx.doi.org/10.1080/1062936X.2016.1253611 Check for updates

OPERA Models: https://github.com/kmansouri/OPERA

OPERA on GitHub



https://github.com/kmansouri/OPERA.git

Toxicity Values

EPA United States Environmental Protection Agency

- **Bioavailability Metric**
- Exposure Limit
- Point Of Departure
- Regulatory Toxicity Value
- Effect Level
- Misc Hazard Information
- Screening Level
- Uncertainty Factor

Download as: TSV | Excel

Type	Subtype	Value	Units	Study Type	Exposure Route	Study Duration	Species	Media	Details	Source
TD50	-	20.7	mM/kg-day	-	-	-	-	-	DSSTox C...	ACToR
TD50	-	2050	mg/kg-day	-	-	-	-	-	DSSTox C...	ACToR
LEL	systemic	619	mg/kg-day	subchronic	oral	subchronic	mouse	-	Study ID: ...	ToxRefDB
NEL	systemic	277	mg/kg-day	subchronic	oral	subchronic	mouse	-	Study ID: ...	ToxRefDB
LEL	systemic	25.0	mg/kg-day	subchronic	oral	subchronic	dog	-	Study ID: ...	ToxRefDB
NEL	systemic	25.0	mg/kg-day	subchronic	oral	subchronic	dog	-	Study ID: ...	ToxRefDB
LEL	systemic	173	mg/kg-day	chronic	oral	chronic	mouse	-	Study ID: ...	ToxRefDB
NEL	systemic	115	mg/kg-day	chronic	oral	chronic	mouse	-	Study ID: ...	ToxRefDB
LEL	systemic	678	mg/kg-day	chronic	oral	chronic	rat	-	Study ID: ...	ToxRefDB
NEL	systemic	283	mg/kg-day	chronic	oral	chronic	rat	-	Study ID: ...	ToxRefDB
LEL	systemic	1230	mg/kg-day	subacute	oral	subacute	rat	-	Study ID: ...	ToxRefDB
NEL	systemic	493	mg/kg-day	subacute	oral	subacute	rat	-	Study ID: ...	ToxRefDB
LEL	systemic	2130	mg/kg-day	subacute	oral	subacute	mouse	-	Study ID: ...	ToxRefDB
NEL	systemic	920	mg/kg-day	subacute	oral	subacute	mouse	-	Study ID: ...	ToxRefDB

Product Composition Details

EPA United States Environmental Protection Agency

- Product & Use Categor...
- Chemical Weight Fra...
- Chemical Functional Use
- Monitoring Data
- Exposure Predictions

Chemical Weight Fractions

Download as: TSV | Excel

Product Name	Product Use Category	Minimum Weight Fraction	Maximum Weight Fraction	Data Type	Source
10-02199- calico tip & ...	personal care: nail poli...	0.01	0.05	MSDS	Retail Product Categori...
6095-6096 minwax wo...	home maintenance: fin...	0.01	0.01	MSDS	Retail Product Categori...
6095/6096 minwax wo...	home maintenance: fin...	0.01	0.01	MSDS	Retail Product Categori...
ab artificial nail remove...	personal care: nail poli...	0.01	0.03	MSDS	Retail Product Categori...
artificial nail remover 7...	personal care: nail poli...	0.01	0.05	MSDS	Retail Product Categori...
calico tip & glue remov...	personal care: nail poli...	0.01	0.05	MSDS	Retail Product Categori...
citristrip canadian strip...	home maintenance: str...	0.65	0.7	MSDS	Retail Product Categori...
citristrip stripping gel q...	home maintenance: str...	0.4	0.55	MSDS	Retail Product Categori...
citristrip stripping gel q...	home maintenance: str...	-	-	MSDS	Retail Product Categori...

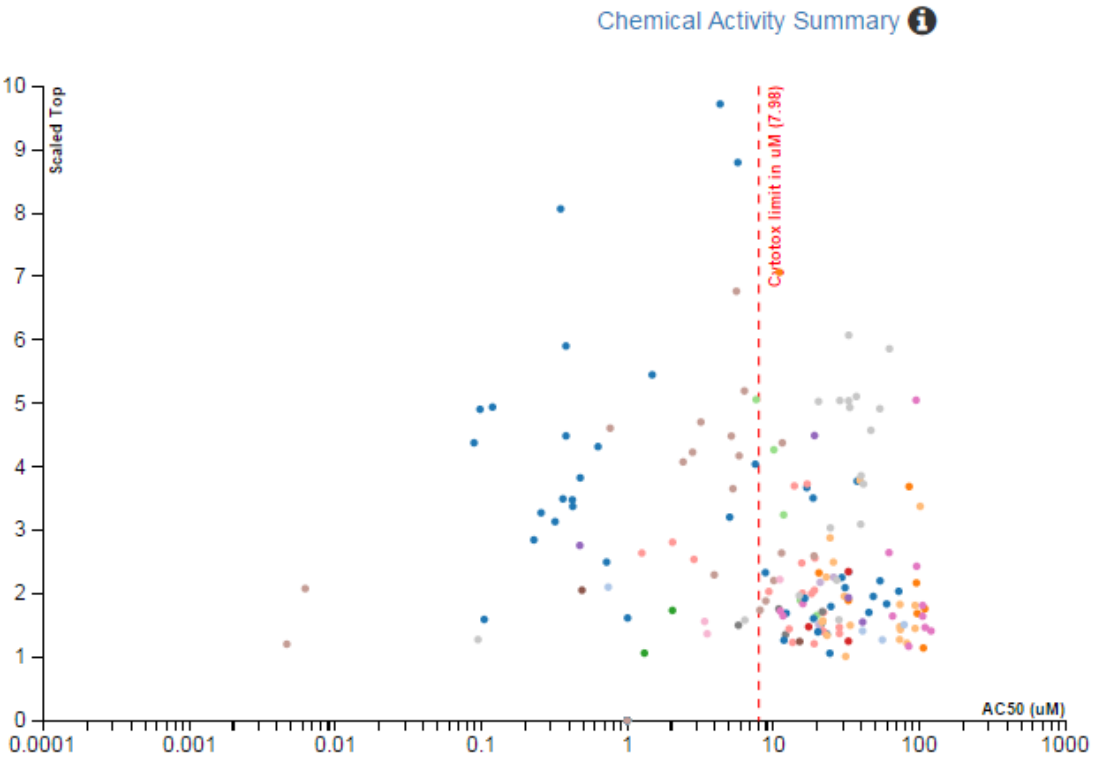
ToxCast and Tox21 Bioassays

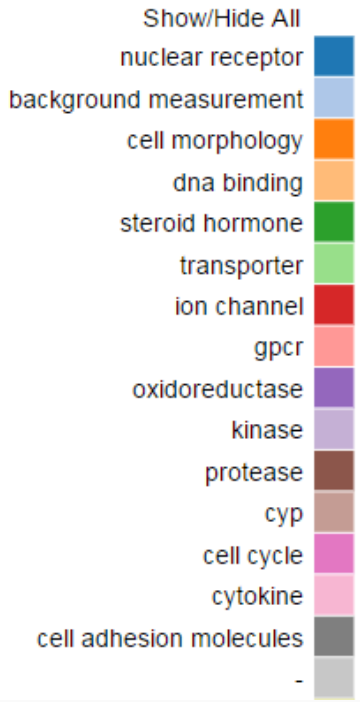
EPA United States Environmental Protection Agency

ToxCast

PubChem

ScrubChem (Beta)





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

ToxCast and Tox21 Bioassays

EPA United States Environmental Protection Agency

Download as: TSV | Excel Show: Inactive | Background

Assay Name	Hit Call	Top	Scaled Top	AC50	log AC50 ↓	Intended Target Family
APR_Hepat_CellLoss_48hr_dn	ACTIVE	1.41	1.41	120	2.08	cell cycle
APR_HepG2_MitoMass_24h_dn	ACTIVE	0.874	1.76	109	2.04	cell morphology
APR_HepG2_OxidativeStress_24h_up	ACTIVE	1.20	1.47	110	2.04	cell cycle
APR_Hepat_DNADamage_48hr_up	ACTIVE	1.84	1.14	107	2.03	cell morphology
APR_HepG2_CellLoss_24h_dn	ACTIVE	1.20	1.81	106	2.02	cell cycle
APR_HepG2_OxidativeStress_72h_up	ACTIVE	1.80	1.65	106	2.02	cell cycle
ATG_HSE_CIS_up	ACTIVE	1.59	3.38	102	2.01	dna binding

Names and Identifiers

EPA United States Environmental Protection Agency

Found 40 synonyms

Legend: **Valid Synonyms** *Good Synonyms* Other Synonyms Copy all Synonyms

N-Methyl-2-pyrrolidone
1-Methylpyrrolidin-2-one
2-Pyrrolidinone, 1-methyl-
872-50-4 Active CAS-RN
2-Pyrrolidinone, 1-methyl-
1-Methyl-2-pyrrolidinone
1-Methyl-2-pyrrolidon
1-Methyl-2-pyrrolidone
1-Methyl-5-pyrrolidinone
1-Methylazacyclopentan-2-one
1-Methylpyrrolidone
1-metil-2-pirrolidona
2-Pyrolidinone, 1-methyl-
2-Pyrrolidone, 1-methyl-
Microposit 2001
M-Pyrol

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

Integrated Literature Searching

EPA United States Environmental Protection Agency

- Google Scholar
- PubMed Abstract S...
- PubChem Articles
- PubChem Patents

Select Term: Hazard **AND** RfD OR reference dose

Google "RfD OR reference dose" AND "Hazard" AND "872-50-4" OR "N-Methyl-2-py

Scholar About 22 results (0.05 sec)

Articles
Case law
My library

Any time
Since 2017
Since 2016
Since 2013
Custom range...

Sort by relevance
Sort by date

include patents
include citations

Create alert

Solvent Substitution: An Analysis of Comprehensive Hazard Screening Indices
M Debia, D Bégin, M Gérin - Journal of occupational and ..., 2011 - oeh.tandfonline.com
... Reliability and weighting factors were not used. Armenti and Moure-Eraso(22) used the same FHS index, using six **hazard** categories, to compare replacement options. ... **872-50-4 N-Methyl-2-pyrrolidone** 45.4 13 1.46E-03 11 4.55 6 20 4 ...
Related articles All 4 versions Cite Save

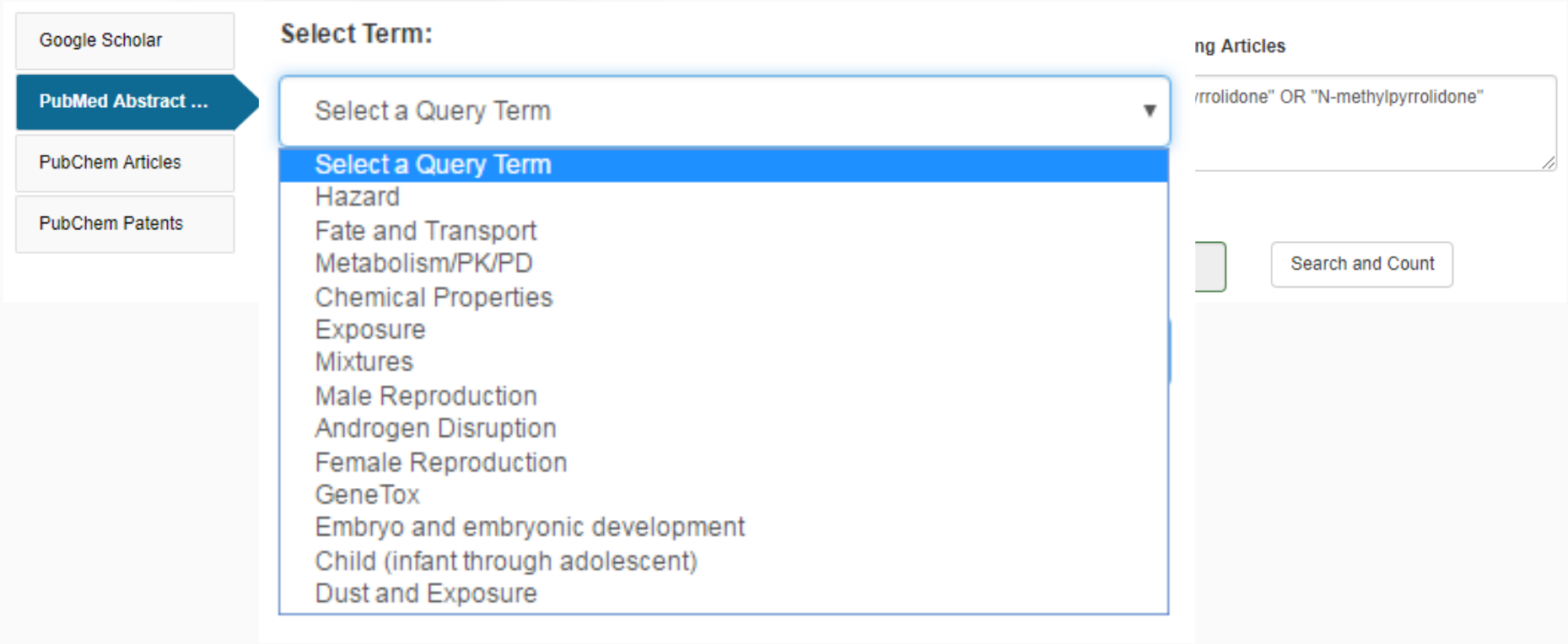
Effects of 'inactive'ingredients on bees
CA Mullin - Current Opinion in Insect Science, 2015 - Elsevier
... tallow amines, organosilicone ethoxylates and co-solvents such as **N-methyl-2-pyrrolidone** (NMP) [7 ... of tools to environmentally monitor residues and determine their relative **hazard** to bees. ... Defining a benchmark or **reference dose**, particularly for bee viruses [51], for what can ...
Cited by 9 Related articles All 2 versions Web of Science: 6 Cite Save More

Biological monitoring: exquisite research probes, risk assessment, and routine exposure measurement
WM Draper - Analytical chemistry, 2001 - ACS Publications
ADVERTISEMENT. ...
Cited by 19 Related articles All 5 versions Web of Science: 10 Cite Save

Evolution of chemical-specific adjustment factors (CSAF) based on recent international experience; increasing utility and facilitating regulatory acceptance
VS Bhat, ME Meek, M Valcke, C English... - Critical Reviews in ..., 2017 - Taylor & Francis
... As shown in the right side of Figure 2, uncertainty is taken into consideration at all tiers of **hazard** assessment. ... For decades, developing "safe doses" such as the oral **reference dose** (RfD), the inhaled reference concentration (RfC) or the acceptable or tolerable daily intake (ADI ...
All 3 versions Cite Save

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

Integrated Literature Searching



Integrated Literature Searching

EPA United States Environmental Protection Agency

Google Scholar

PubMed Abstract Sifter

PubChem Articles

PubChem Patents

Edit the Query Before Retrieving Articles

("872-50-4" OR "N-Methyl-2-pyrrolidone" OR "N-methylpyrrolidone") AND (exposure OR near-field OR far-field OR SHEDS[tiab] AND ENVIRONMENTAL MONITORING)

1

0

0	0	0	0	24078144	2013	Biological monitoring and health effects of low-level exposure to N-methyl-2-pyrrolidone: a cross-sectional study.
0	0	0	0	23337464	2013	Biomonitoring of exposure to N-methyl-2-pyrrolidone in workers of the automobile industry.
0	0	0	0	19875680	2009	Quantitative risk analysis for N-methyl pyrrolidone using physiologically based pharmacokinetic and benchmark dos…
0	0	0	0	16897094	2006	Human volunteer study on the influence of exposure duration and dilution of dermally applied N-methyl-2-pyrrolidon…

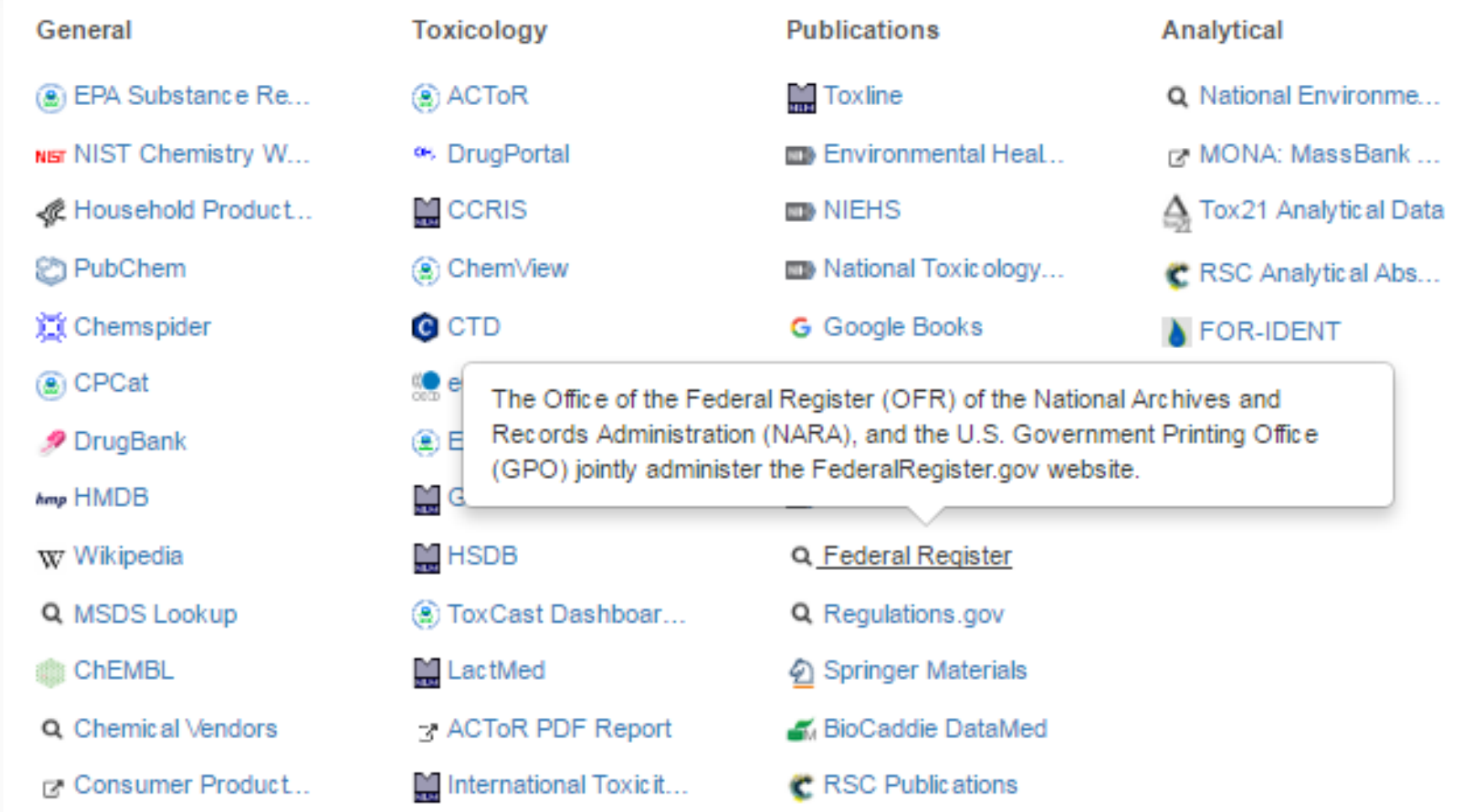
Record: 1 of 16

Title: [Not Available].

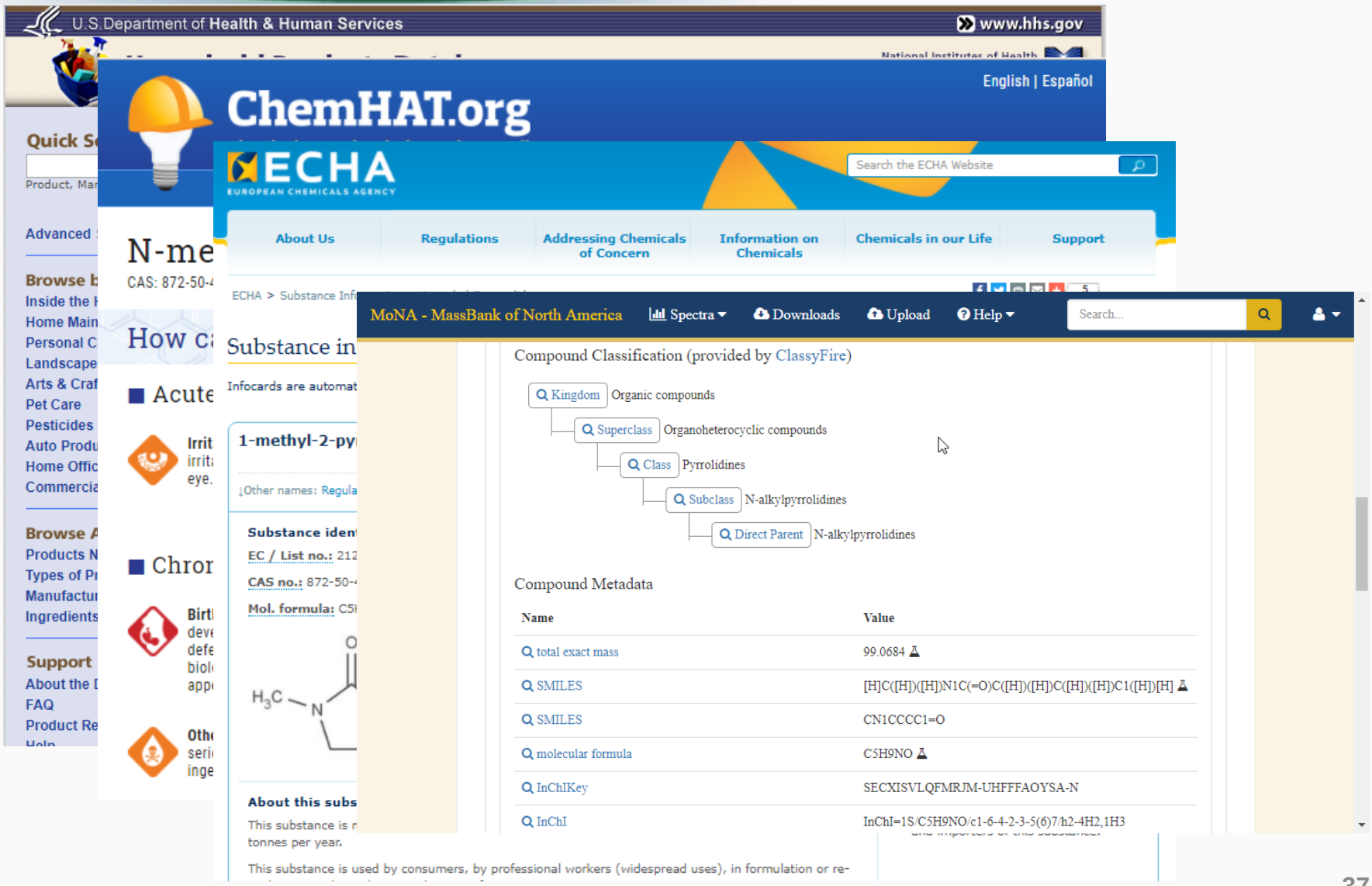
Abstract: N-Ethyl-2-pyrrolidone (NEP), a polar aprotic solvent, is used in many applications as substitute for the structural analogue N-methyl-2-pyrrolidone (NMP), e. g. for surface coatings, in cleaning agents and paint strippers. Monitoring studies indicate that individuals within the general public, without occupational exposure, may be exposed to NEP to an extent, which is comparable to NMP. As NMP, NEP presents a potential health hazard due to its developmental toxicity and teratogenicity. Exposure to NEP can be quantified by the determination of the excretion of its urinary metabolites 5-Hydroxy-N-ethyl-2-pyrrolidone (5-HNEP) and 2-Hydroxy-N-ethylsuccinimide (2-HESI). For the derivation of HBM values, the german Human Biomonitoring Commission (HBM commission) evaluated different toxicological endpoints and finally decided on the BMDL05 and the BMD10 for the endpoint "reduced grasp intensity" of a subchronic feeding study with rats as point of departure (POD) for further procedural steps. The resulting HBM-I and HBM-II values for the sum of the metabolites 5-HNEP and 2-HESI in the urine of children are 10 resp. 25 mg/l and in the urine of adults are 15 resp. 40 mg/l. If the HBM values are exceeded, a check-up will be necessary at first. Measurements above the HBM-II value give cause for concern, especially for pregnant women. Air meas ents to determine the source of exposure can be useful. The possibility of skin absorption from use of

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

Links to Other Resources



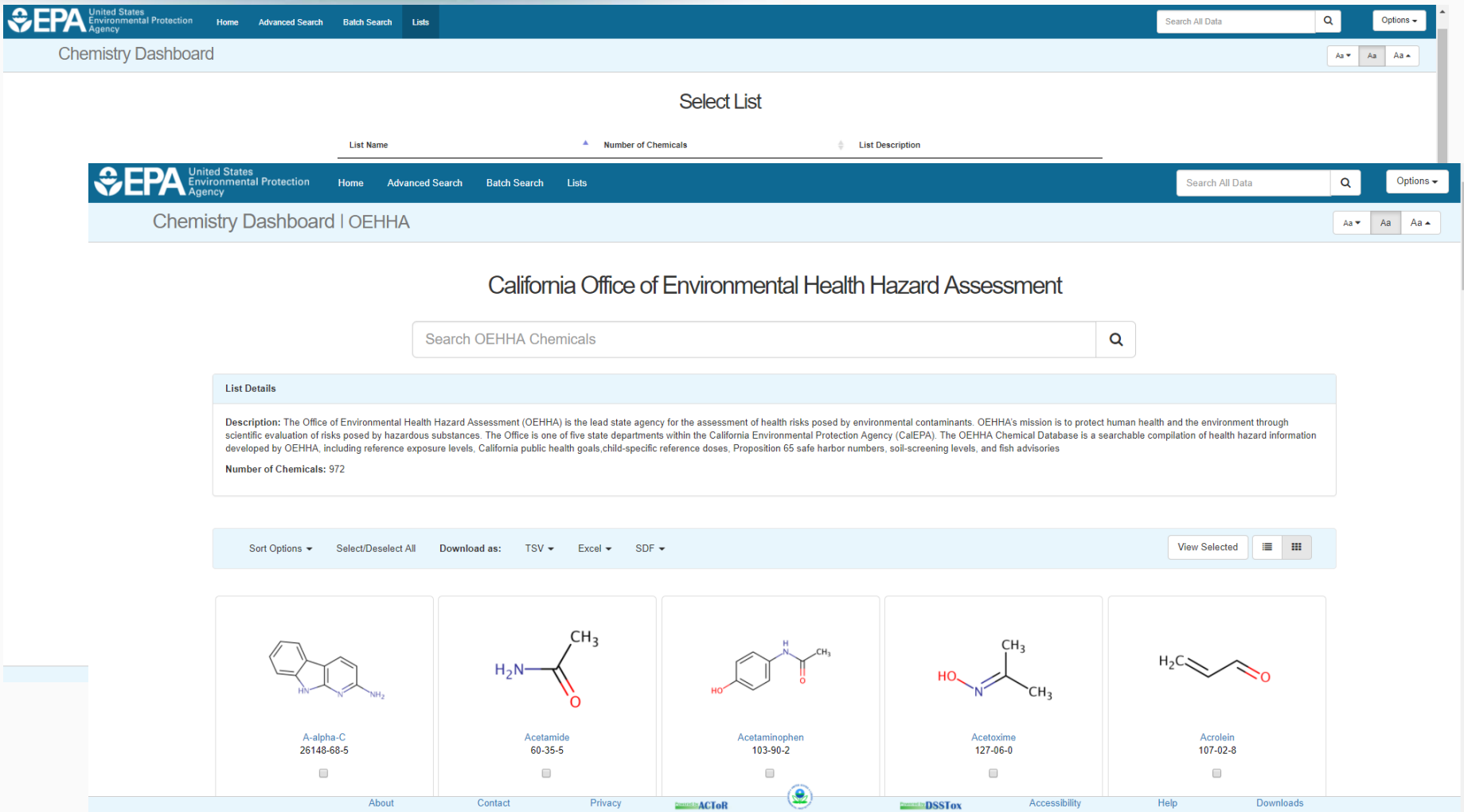
Example External Links…



Accessing Lists of Chemicals

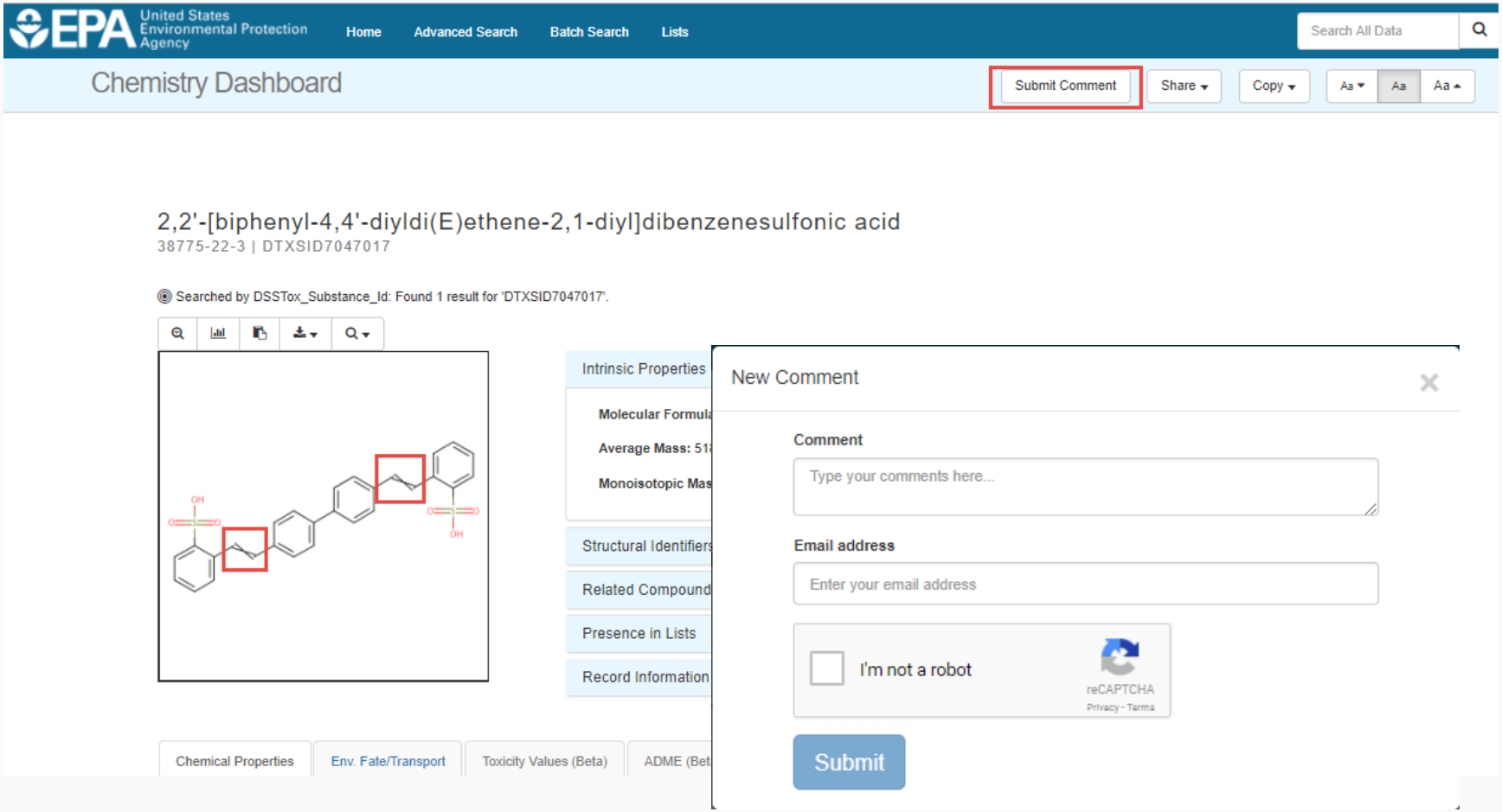
- Build out definitive “lists” of chemicals
  - Algal toxins
  - Poly/perfluorinated chemicals
  - Pesticides
  - Toxcast screening chemical collection

The Collection of Lists



Crowdsourced Curation – HELP!

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

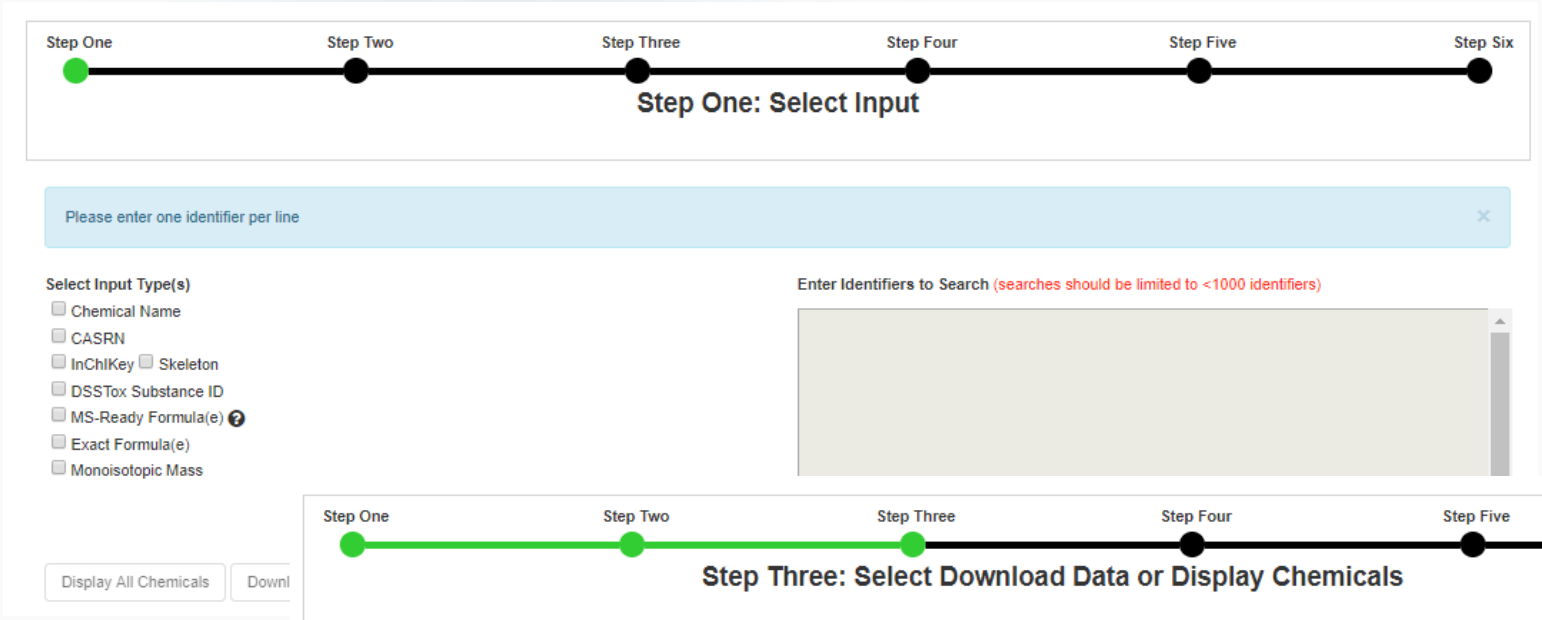


Answering Questions

- I have a 1000 CAS Numbers (or Names) – are there data available?
  - Has any Toxcast data been run?
  - Are there Toxicity Data values available?
  - Are there predicted exposure data (via Expocast)?
  - Can I get predicted physchem data for my model?

Batch Searching for Data for Thousands of Chemicals

EPA United States Environmental Protection Agency



Batch Searching for Data for Thousands of Chemicals

EPA United States Environmental Protection Agency

Step One — Step Two — Step Three — Step Four — Step Five — Step Six

Step Six: Click "Download"

Display All Chemicals | Download Chemical Data

Select Output Format

Excel

Customize Results

- Select All
- Select All In Lists

Chemical Identifiers

- DTXSID
- Chemical Name
- CAS-RN
- InChIKey

Presence In List

- 40CFR355
- Algal Toxins
- ATSDR Toxic Substances Portal Chemical List
- Bisphenol Compounds

INPUT	FOUND_BY	DTXSID	PREFERRED_NAME	CASRN	IUPAC_NA	SMILES	MOLECULA	AVERAGE_	ATMOSPHI	BIOCONCE	BIODEGRA	BOILING_P	HENRYS_L	OPERA_KN	OCTANOL_	SOIL_ADS	OCTANOL_	MELTING_	VAPOR_PR	WATER_SOL	UBILITY
18699-02-0	CAS-RN, In	DTXSID0020020	4-Acetylaminophenylacetic acid	18699-02-0	(4-Acetamic	CC(=O)NC1	C10H11NO	193.202	7.544E-12	3.42204	3.53611	291.22	4.268E-10	0.432781	8.16821	40.7227	1.11077	169.466	5.945E-07	0.0027159	
50594-66-6	CAS-RN, In	DTXSID0020022	Acifluorfen	50594-66-6	5-[2-Chloro-	OC(=O)C1=	C14H7ClF3	361.66	1.38E-11	5.01914	3.54682	322.232	8.016E-09	0.894058	10.6329	375.753	3.75101	152.287	1.151E-09	0.0002794	
2113-61-3	CAS-RN, In	DTXSID0020072	4-Biphenylamine hydrochloride	2113-61-3	[1,1'-Biphen	Cl.NC1=CC	C12H12ClN	205.69	7.263E-11	58.7731	22.3831	302.005	2.234E-07	0.137444	9.26788	1450.75	2.86436	229.065	0.0001844	0.0004509	
61-82-5	CAS-RN, In	DTXSID0020076	Amitrole	61-82-5	1H-1,2,4-Tri	NC1=NNC=	C2H4N4	84.082	2.153E-11	3.20651	3.95444	258.064	3.079E-08	0.134589	7.0857	11.2387	-1.05255	147.209	4.798E-07	2.38468	
134-03-2	CAS-RN, In	DTXSID0020105	Sodium L-ascorbate	134-03-2	Sodium (2R	[Na+].OC[C	C6H7NaO6	198.106	8.252E-12	1.77879	3.54017	313.644	1.281E-11	0.109384	8.18129	303.789	-1.72664	168.898	5.036E-10	0.287515	
22839-47-0	CAS-RN, In	DTXSID0020107	Aspartame	22839-47-0	Methyl L-al	COC(=O)[C	C14H18N2O	294.307	1.69E-11	4.54579	3.53293	284.29	6.157E-11	0.110876	9.24976	166.002	-0.334933	234.896	1.373E-09	0.0126548	
140-11-4	CAS-RN, In	DTXSID0020151	Benzyl acetate	140-11-4	Benzyl acet	CC(=O)OC	C9H10O2	150.177	1.137E-11	6.68119	4.30298	216.69	3.895E-07	0.134105	4.83138	72.657	2.15746	-7.81986	0.181933	0.0139724	
58-08-2	CAS-RN, In	DTXSID0020232	Caffeine	58-08-2	1,3,7-Trimet	CN1C=NC2	C8H10N4O2	194.194	6.002E-12	1.86426	3.52242	286.771	1.595E-06	0.281842	8.51987	62.2886	-0.150871	202.556	2.08E-06	0.0991131	
814-80-2	CAS-RN, In	DTXSID0020236	Calcium lactate	814-80-2	Calcium bis	[Ca++].CC(	C6H10CaO6	218.218	-	-	-	-	-	-	-	-	-	-	-	-	
88-73-3	CAS-RN, In	DTXSID0020280	1-Chloro-2-nitrobenzene	88-73-3	1-Chloro-2-r	[O-][N+](=C	C6H4ClNO2	157.55	1.039E-12	22.1838	5.34707	245.57	9.396E-06	0.454339	4.31183	668.302	2.3191	48.0313	0.040638	0.0010906	
5131-60-2	CAS-RN, In	DTXSID0020282	4-Chloro-1,3-diaminobenzene	5131-60-2	4-Chlorober	NC1=CC(N)	C6H7ClN2	142.59	2.495E-11	8.06476	8.48537	238.553	2.438E-07	0.141138	7.84318	67.087	0.962899	89.0194	0.0006677	0.0057247	
61702-44-1	CAS-RN, In	DTXSID0020284	2-Chloro-1,4-diaminobenzene sulfate	61702-44-1	Sulfuric aci	OS(O)(=O)=	C6H9ClN2O	240.66	2.48E-11	8.07799	15.2095	238.674	2.439E-07	0.141213	7.84318	67.0933	0.979902	91.091	0.0006678	0.0021064	
95-74-9	CAS-RN, In	DTXSID0020286	3-Chloro-4-methylaniline	95-74-9	3-Chloro-4-r	CC1=C(Cl)C	C7H8ClN	141.6	2.486E-11	20.8027	3.66038	235.888	5.182E-07	0.166045	7.45692	164.083	1.92694	34.0556	0.0566718	0.004701	
3165-93-3	CAS-RN, In	DTXSID0020288	4-Chloro-2-methylaniline hydrochloride	3165-93-3	4-Chloro-2-r	Cl.CC1=C(N	C7H9Cl2N	178.06	2.473E-11	21.8011	3.56253	242.562	2.171E-07	0.162479	7.47775	164.241	2.13482	228.53	0.0536476	0.0046491	
150-68-5	CAS-RN, In	DTXSID0020311	Monuron	150-68-5	N'-(4-Chloro	CN(C)C(=O	C9H11ClN2	198.65	1.233E-11	21.0734	4.9644	300.033	1.574E-09	0.080547	9.30039	68.765	1.97223	156.714	2.733E-07	0.0009876	
76-06-2	CAS-RN, In	DTXSID0020315	Chloropicrin	76-06-2	Trichloro(nit	[O-][N+](=C	CCl3NO2	164.37	4.975E-15	19.1016	4.61961	126.822	0.005	2.61824	3.07713	61.6502	1.85845	-58.1899	90.3279	0.0176702	
1897-45-6	CAS-RN, In	DTXSID0020319	Chlorothalonil	1897-45-6	2,4,5,6-Tetr	ClC1=C(Cl)(	C8Cl4N2	265.9	2.835E-12	86.539	5.52071	339.516	3.532E-06	1.27499	8.49259	1866.18	3.1346	230.81	1.29E-06	3.211E-06	
4998-76-9	CAS-RN, In	DTXSID0020361	Cyclohexylamine hydrochloride	4998-76-9	Cyclohexan	Cl.NC1CCC	C6H14ClN	135.64	1.307E-11	6.49625	4.50266	132.539	1.042E-05	0.146863	4.74766	46.924	0.968819	216.087	7.42003	5.20689	
59865-13-3	CAS-RN, In	DTXSID0020365	Cyclosporin A	59865-13-3	(3S,6S,9S,1	CC[C@@H	C62H111N1	1202.635	1.855E-10	10.5684	13.4555	380.042	1.859E-11	0.336609	9.63201	30808.1	2.8611	149.875	2.258E-10	0.0002704	
52-89-1	CAS-RN, In	DTXSID0020367	l-Cysteine hydrochloride	52-89-1	L-Cysteine-	Cl.N[C@@H	C3H8ClNO2	157.61	5.306E-12	4.29025	3.54748	247.043	5.206E-09	0.13252	7.1697	12.2561	-2.16183	176.504	7.044E-06	1.72407	
120-36-5	CAS-RN, In	DTXSID0020440	Dichlorprop	120-36-5	2-(2,4-Dichl	CC(OC1=C(	C9H8Cl2O3	235.06	1.165E-11	3.54436	3.53189	298.478	3.028E-09	0.144781	9.19609	130.634	2.82696	128.196	4.587E-07	0.0015689	
94-75-7	CAS-RN, In	DTXSID0020442	2,4-Dichlorophenoxyacetic acid	94-75-7	(2,4-Dichlor	OC(=O)CO(	C8H6Cl2O3	221.03	8.509E-12	3.33933	3.54506	300.487	2.974E-09	0.214222	8.64934	81.4078	2.75594	131.291	5.25E-07	0.0026012	
330-54-1	CAS-RN, In	DTXSID0020446	Diuron	330-54-1	N'-(3,4-Dich	CN(C)C(=O	C9H10Cl2N	233.09	1.205E-11	17.1704	3.55395	323.013	1.062E-09	0.057673	9.95685	190.73	2.48235	151.313	8.966E-08	0.0002698	
756-79-6	CAS-RN, In	DTXSID0020494	Dimethyl methylphosphonate	756-79-6	Dimethyl m	COP(C)(=O	C3H9O3P	124.076	1.794E-11	3.57788	4.58659	172.811	9.154E-08	0.101991	4.52878	5.1856	-0.557951	-65.5885	0.474397	6.36062	
120-61-6	CAS-RN, In	DTXSID0020498	Dimethyl terephthalate	120-61-6	Dimethyl be	COC(=O)C1	C10H10O4	194.186	5.999E-12	3.75263	3.67612	279.314	7.991E-08	0.224423	5.6765	43.5705	1.98051	74.4094	0.0061344	0.0006996	
51-28-5	CAS-RN, In	DTXSID0020523	2,4-Dinitrophenol	51-28-5	2,4-Dinitrop	OC1=C(C=(	C6H4N2O5	184.107	8.576E-13	5.31642	4.59079	349.689	9.847E-08	0.134389	7.33727	221.884	1.87568	117.909	0.0002339	0.0017297	
121-14-2	CAS-RN, In	DTXSID0020529	2,4-Dinitrotoluene	121-14-2	1-Methyl-2,4	CC1=C(C=(	C7H6N2O4	182.135	1.626E-12	9.50638	3.55991	297.386	9.418E-08	0.191762	5.07604	402.949	1.91053	70.7766	0.0002412	0.0012579	
50-28-2	CAS-RN, In	DTXSID0020573	17beta-Estradiol	50-28-2	(17beta)-Es	[H][C@@]1	C18H24O2	272.388	1.66E-11	33.8438	58.4157	372.709	8.528E-06	1.49013	9.15213	212402	3.85324	195.408	3.413E-10	1.729E-05	
140-67-0	CAS-RN, In	DTXSID0020575	Estragole	140-67-0	1-Methoxy-4	COC1=CC=	C10H12O	148.205	7.077E-11	34.0071	3.53538	219.742	9.766E-06	1.3162	4.9796	1632.41	3.02077	-7.93361	0.173572	0.001369	
536-33-4	CAS-RN, In	DTXSID0020577	Ethionamide	536-33-4	2-Ethylpyric	CCC1=NC=	C8H10N2S	166.24	4.471E-11	5.38695	2.7777	298.219	1.828E-09	0.252948	8.49908	63.9763	1.20325	153.065	5.153E-06	0.139454	
106-87-6	CAS-RN, In	DTXSID0020604	4-Vinyl-1-cyclohexene dioxide	106-87-6	3-(Oxiran-2-	C1OC1C1C	C8H12O2	140.182	1.945E-11	3.56105	3.65865	218.846	2.014E-05	0.142757	4.30621	57.8208	0.803226	75.4067	0.0260644	0.566226	
103-23-1	CAS-RN, In	DTXSID0020606	Bis(2-ethylhexyl)hexanedioate	103-23-1	Bis(2-ethylh	CCCCC(CC	C22H42O4	370.574	1.203E-11	233.91	5.69976	403.187	3.923E-07	1.66751	10.86	14358.7	6.85077	-61.5844	1.158E-06	3.903E-06	
25812-30-0	CAS-RN, In	DTXSID0020652	Gemfibrozil	25812-30-0	5-(2,5-Dime	CC1=CC(O	C15H22O3	250.338	1.695E-11	3.16534	3.52148	326.707	8.615E-09	0.893056	9.60201	216.147	4.42964	74.6536	9.356E-06	0.0002483	
105-87-3	CAS-RN, In	DTXSID0020654	Geranyl acetate	105-87-3	(2E)-3,7-Dir	CC(C)=CCC	C12H20O2	196.29	1.645E-10	19.0144	4.49702	236.888	0.0002711	0.413938	5.79896	470.461	4.13931	-53.0564	0.0434534	0.0003752	

Supporting NTA Applications

- Exposure surveillance
  - What chemicals are in food, products, dust, blood, etc.?

- Chemical prioritization
  - What are relevant chemicals & mixtures?

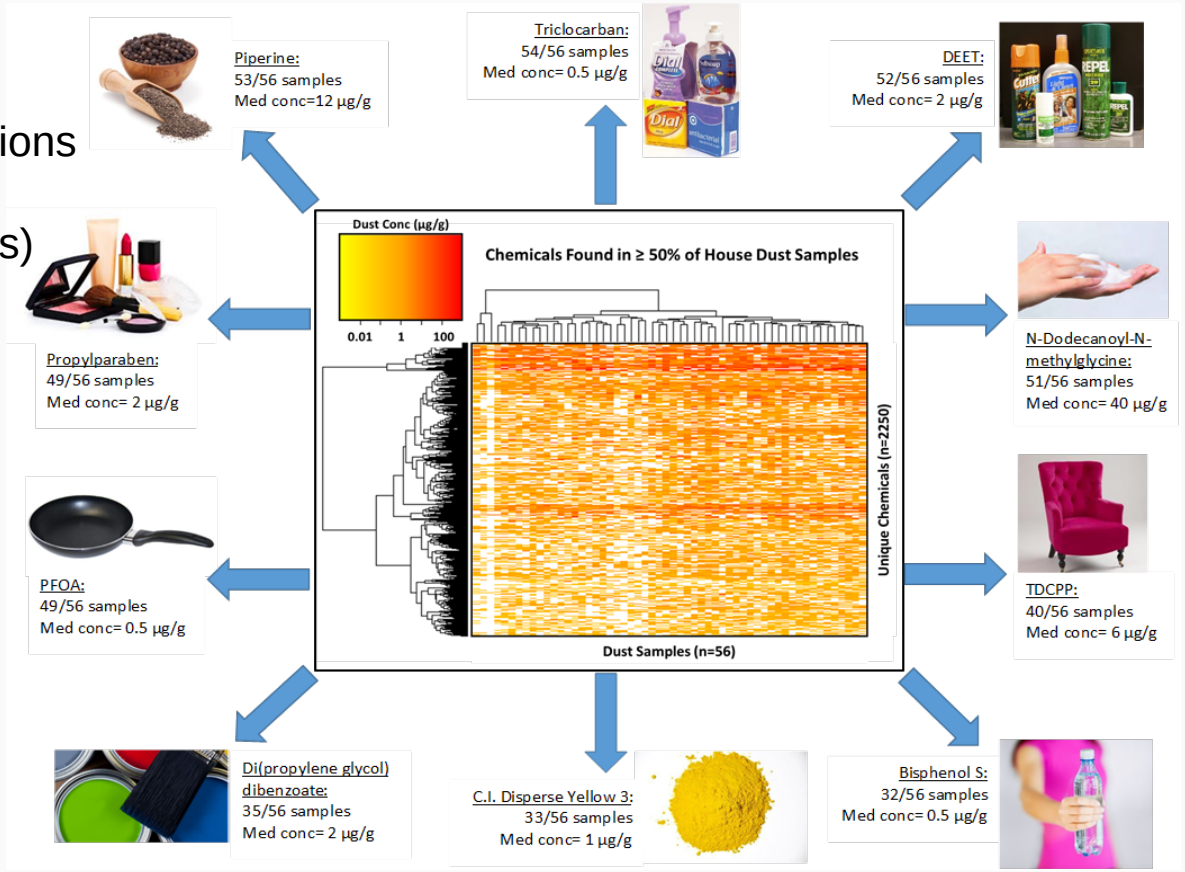
- Exposure forensics
  - What are chemical signatures of exposure sources?

- Effect-directed analysis
  - What are the biologically active chemicals in complex mixtures?

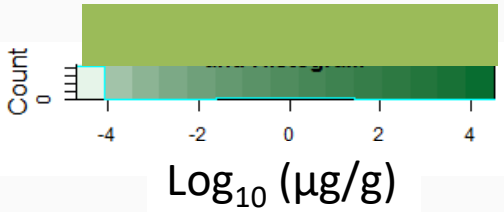
- Biomarker discovery
  - What chemicals are predictive of bioactivity/health impairment?

Suspect Screening Analysis
Chemicals in House Dust

- Chemical confirmation from ToxCast
- Detection frequency
- External calibration for concentrations
- Cluster analysis (homes/chemicals)
- Chemical functional use



Improving Exposure Estimates – Characterizing Commercial Products



- Air freshener
- Baby soap
- Carpet
- Carpet padding
- Cereals
- Cotton clothing
- Deodorant
- Fabric upholstery
- Glass cleaners
- Hand soap
- Indoor house paint
- Lipstick
- Plastic children's toys
- Shampoo
- Shaving cream
- Shower curtain
- Skin lotion
- Sunscreen
- Toothpaste
- Vinyl upholstery

- **GCXGC-MS with DCM Extraction**
- **1606 tentative and confirmed chemical identifications**

- Common Chemical (n>19)
- ToxCast
- Flame Retardant
- Potent ER

Wambaugh et al. Unpublished

Non-targeted Analysis Support

Environmental Pollution
Volume 234, March 2018, Pages 297-306

Suspect screening and non-targeted analysis of drinking water using point-of-use filters ☆

Seth R. Newton [a], Rebecca L. McMahen [a, b], Jon R. Sobus [a], Kamel Mansouri [b, c, 1], Antony J. Williams [c], Andrew D. McEachran [b, c], Mark J. Strynar [a]

Talanta
Available online 11 January 2018
In Press, Accepted Manuscript

A Comparison of Three Liquid Chromatography (LC) Retention Time Prediction Models

Andrew D. McEachran [a, b], Kamel Mansouri [a, b, 1], Seth R. Newton [c], Brandiese E.J. Beverly [a, c, 2], Jon R. Sobus [c], Antony J. Williams [b]

Environment International
Volume 88, March 2016, Pages 269-280

Linking high resolution mass spectrometry data with exposure and toxicity forecasts to advance high-throughput environmental monitoring

Julia E. Rager [a], Mark J. Strynar [b], Shuang Liang [a], Rebecca L. McMahen [a], Ann M. Richard [c], Christopher M. Grulke [d], John F. Wambaugh [c], Kristin K. Isaacs [b], Richard Judson [c], Antony J. Williams [c], Jon R. Sobus [b]

Journal of Exposure Science & Environmental Epidemiology

Review Article

Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA

Jon R. Sobus, John F. Wambaugh, Kristin K. Isaacs, Antony J. Williams, Andrew D. McEachran, Ann M. Richard, Christopher M. Grulke, Elin M. Ulrich, Julia E. Rager, Mark J. Strynar & Seth R. Newton

Analytical and Bioanalytical Chemistry
March 2017, Volume 409, Issue 7, pp 1729–1735 | Cite as

Identifying known unknowns using the US EPA's CompTox Chemistry Dashboard

Authors Authors and affiliations
Andrew D. McEachran, Jon R. Sobus, Antony J. Williams

Environmental Science & Technology

Open Science for Identifying "Known Unknown" Chemicals

Emma L. Schymanski and Antony J. Williams

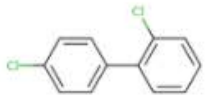
Chemical “Families”

- Sometimes the simplest of questions are difficult to answer!
  - What is the list of CAS Numbers for all PCBs?
  - Can I get an SDF file of all PCBs?
  - Do you have predicted properties for all PCBs?
  - What toxicity data is available for individual PCBS?
  - Have you measured ToxCast data for any PCBs?
  - Can I get all PCBs listed in an Excel Spreadsheet?

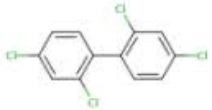
Chemical “Families”

EPA United States Environmental Protection Agency

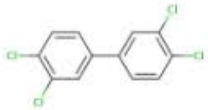
Download as: TSV | Excel | SDF



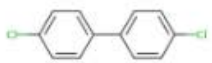
2,4'-Dichlorobiphenyl
34883-43-7



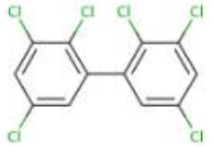
2,2',4,4'-Tetrachlorobiphenyl
2437-79-8



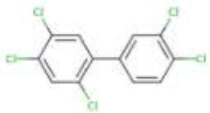
3,3',4,4'-Tetrachlorobinphenyl
32598-13-3



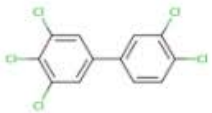
4,4'-Dichlorobiphenyl
2050-68-2



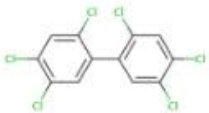
2,2',3,3',5,5'-Hexachlorobiphenyl
35694-04-3



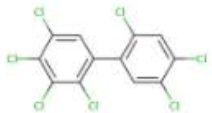
2,3',4,4',5-Pentachlorobiphenyl
31508-00-6



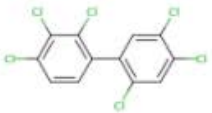
3,3',4,4',5-Pentachlorobiphenyl
57465-28-8



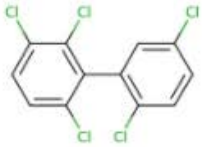
2,2',4,4',5,5'-Hexachlorobiphenyl
35065-27-1



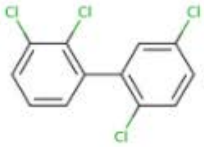
2,2',3,4,4',5,5'-Heptachlorobiphenyl
35065-29-3



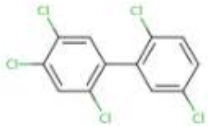
2,2',3,4,4',5'-Hexachlorobiphenyl
35065-28-2



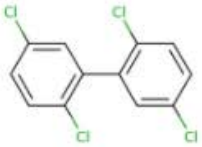
2,2',3,5',6-Pentachlorobiphenyl
38379-99-6



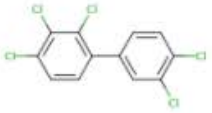
2,2',3,5'-tetrachlorobiphenyl
41464-39-5



2,2,4,5,5'-Pentochlorobiphenyl
37680-73-2



2,2',5,5'-Tetrachlorobiphenyl
35693-99-3

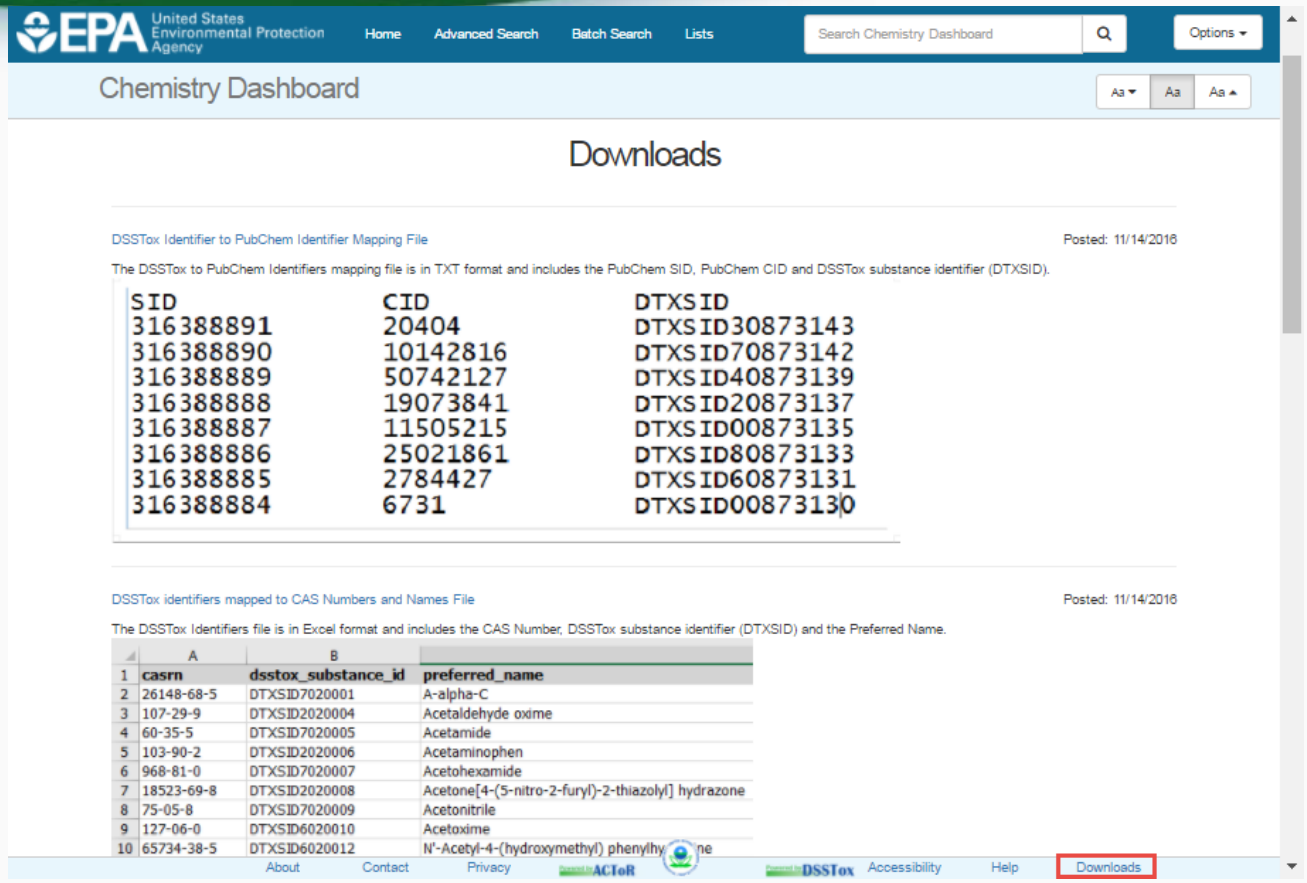


2,3,3',4,4'-Pentachlorobiphenyl
32598-14-4

One click download

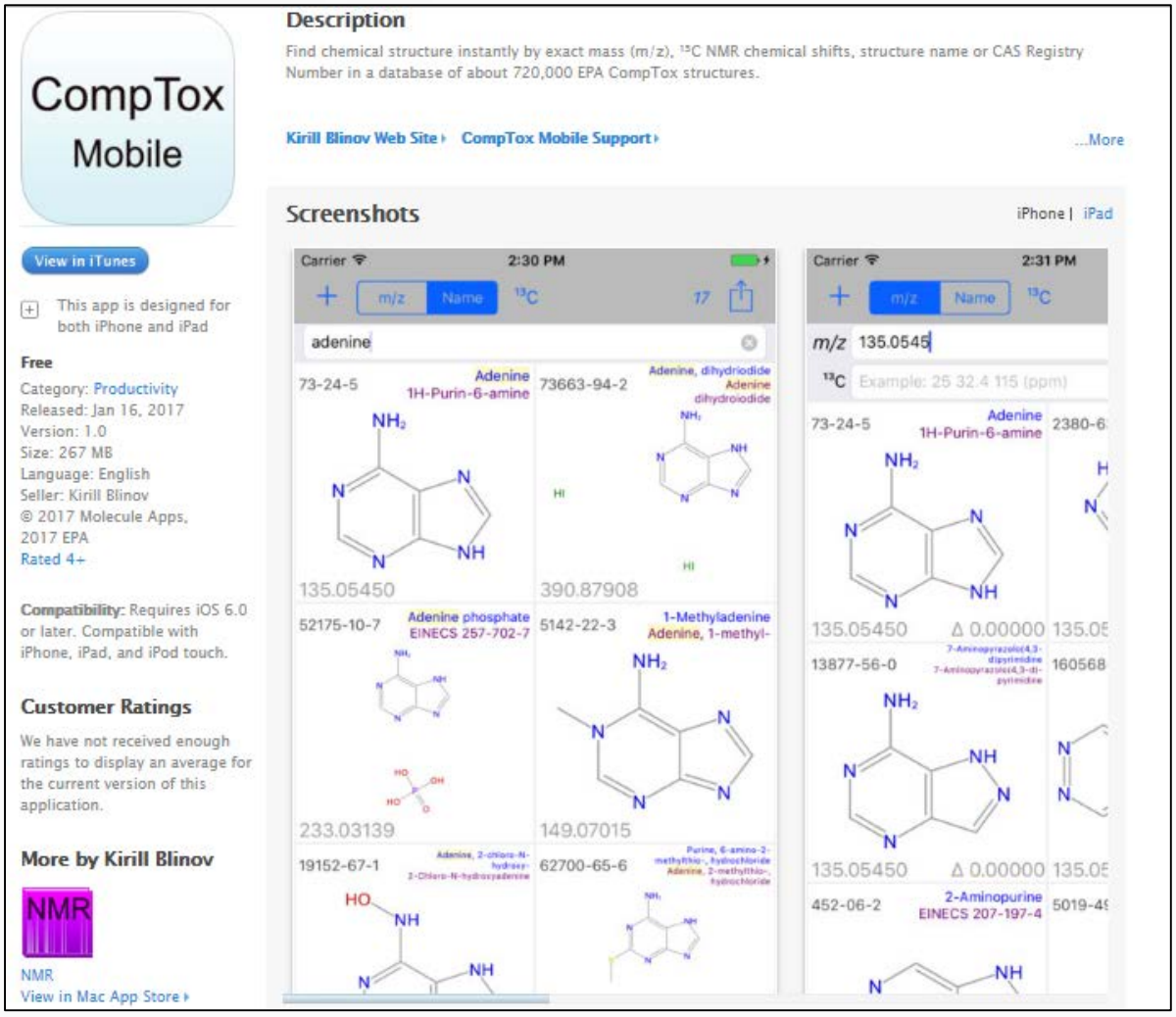
	A	B	C	D	E	F	G	H	I
1	DTXSID	CASRN	PREFERRED NAME	IUPAC NAME	MOLECULAR F	MONOISOT	AVERAGE	SMILES	INCHI KEY
2	DTXSID9074779	70362-45-7	PCB 045	2,2',3,6-Tetrachloro-1,1	C12H6Cl4	289.92236	291.98001	ClC1=CC=C(Cl)C(=C1Cl)C1=CC=CC=C1Cl	VHGHHZZT
3	DTXSID9074777	38444-73-4	PCB 019	2,2',6-Trichloro-1,1'-biph	C12H7Cl3	255.96133	257.54001	ClC1=CC=CC(Cl)=C1C1=C(Cl)C=CC=C1	MVXIJRBB(
4	DTXSID9074228	74472-39-2	2,3',4',5',6-Pentachlorol	2,3',4',5',6-Pentachloro-	C12H5Cl5	323.88339	326.42001	ClC1=CC(=CC(Cl)=C1Cl)C1=C(Cl)C=CC=C1Cl	WAZUWHG
5	DTXSID9074226	74472-37-0	2,3,4,4',5-Pentachlorob	2,3,4,4',5-Pentachloro-1	C12H5Cl5	323.88339	326.42001	ClC1=CC=C(C=C1)C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1	SXZSFWH(
6	DTXSID9074224	74472-35-8	2,3,3',4,6-Pentachlorob	2,3,3',4,6-Pentachloro-1	C12H5Cl5	323.88339	326.42001	ClC1=CC=CC(=C1)C1=C(Cl)C(Cl)=C(Cl)C=C1Cl	XGQBSVV'
7	DTXSID9074222	74472-33-6	2,3,3',6-Tetrachlorobiph	2,3,3',6-Tetrachloro-1,1	C12H6Cl4	289.92236	291.98001	ClC1=CC=CC(=C1)C1=C(Cl)C(Cl)=CC=C1Cl	WZNAMGY
8	DTXSID9074220	74338-23-1	2,3',5',6-Tetrachlorobipl	2,3',5',6-Tetrachloro-1,1	C12H6Cl4	289.92236	291.98001	ClC1=CC(=CC(Cl)=C1)C1=C(Cl)C=CC=C1Cl	HDULUCZR
9	DTXSID9074199	68194-11-6	2,3,4',5,6-Pentachlorob	2,3,4',5,6-Pentachloro-1	C12H5Cl5	323.88339	326.42001	ClC1=CC=C(C=C1)C1=C(Cl)C(Cl)=CC(Cl)=C1Cl	ZDDZPDTV
10	DTXSID9074197	68194-08-1	2,2',3,4',6,6'-Hexachlor	2,2',3,4',6,6'-Hexachlor	C12H4Cl6	357.84442	360.85999	ClC1=CC(Cl)=C(C(Cl)=C1)C1=C(Cl)C(Cl)=CC=C1Cl	RPPNJBZN
11	DTXSID9074195	68194-04-7	2,2',4,6'-Tetrachlorobipl	2,2',4,6'-Tetrachloro-1,1	C12H6Cl4	289.92236	291.98001	ClC1=CC=C(C(Cl)=C1)C1=C(Cl)C=CC=C1Cl	WVHNUGR
12	DTXSID9074193	60233-25-2	2,2',3,4',6'-Pentachlorol	2,2',3,4',6'-Pentachloro-	C12H5Cl5	323.88339	326.42001	ClC1=CC(Cl)=C(C(Cl)=C1)C1=C(Cl)C(Cl)=CC=C1	GOFFZTAP
13	DTXSID9074191	60145-23-5	2,2',3,4,4',5,6'-Heptach	2,2',3,4,4',5,6'-Heptachl	C12H3Cl7	391.80544	395.31	ClC1=CC(Cl)=C(C(Cl)=C1)C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1	RXRLRYZU
14	DTXSID9074149	41411-61-4	2,2',3,4,5,6-Hexachloro	2,2',3,4,5,6-Hexachloro	C12H4Cl6	357.84442	360.85999	ClC1=C(C=CC=C1)C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1Cl	RUEIBQJF(
15	DTXSID9074147	40186-70-7	2,2',3,3',4,5',6-Heptach	2,2',3,3',4,5',6-Heptachl	C12H3Cl7	391.80544	395.31	ClC1=CC(Cl)=C(Cl)C(=C1)C1=C(Cl)C(Cl)=C(Cl)C=C1Cl	KJBDZJFS'
16	DTXSID9074145	39635-33-1	3,3',4,5,5'-Pentachlorob	3,3',4,5,5'-Pentachloro-	C12H5Cl5	323.88339	326.42001	ClC1=CC(=CC(Cl)=C1)C1=CC(Cl)=C(Cl)C(Cl)=C1	MXVAYAXI
17	DTXSID9074143	38444-76-7	2,3',6-Trichlorobiphenyl	2,3',6-Trichloro-1,1'-biph	C12H7Cl3	255.96133	257.54001	ClC1=CC=CC(=C1)C1=C(Cl)C=CC=C1Cl	VQOFJPFY
18	DTXSID9074141	38380-05-1	2,2',3,3',4,6'-Hexachlor	2,2',3,3',4,6'-Hexachlor	C12H4Cl6	357.84442	360.85999	ClC1=CC=C(Cl)C(=C1Cl)C1=C(Cl)C(Cl)=C(Cl)C=C1	OKBJVIVE
19	DTXSID9073599	65510-45-4	2,2',3,4,4'-Pentachlorob	2,2',3,4,4'-Pentachloro-	C12H5Cl5	323.88339	326.42001	ClC1=CC(Cl)=C(C=C1)C1=C(Cl)C(Cl)=C(Cl)C=C1	LACXVZHA
20	DTXSID9073541	52744-13-5	2,2',3,3',5,6'-Hexachlor	2,2',3,3',5,6'-Hexachlor	C12H4Cl6	357.84442	360.85999	ClC1=CC(=C(Cl)C(Cl)=C1)C1=C(Cl)C(Cl)=CC=C1Cl	UUTNFLRS
21	DTXSID9073410	16606-02-3	2,4',5-Trichlorobiphenyl	2,4',5-Trichloro-1,1'-biph	C12H7Cl3	255.96133	257.54001	ClC1=CC=C(C=C1)C1=C(Cl)C=CC(Cl)=C1	VAHKBZSA
22	DTXSID80873557	36559-22-5	2,2',3,4'-Tetrachloro-1,1	2,2',3,4'-Tetrachloro-1,1	C12H6Cl4	289.92236	291.98	ClC1=CC(Cl)=C(C=C1)C1=CC=CC(Cl)=C1Cl	ALFHIHDQ
23	DTXSID8074780	61798-70-7	PCB 131	2,2',3,3',4,6-Hexachlorc	C12H4Cl6	357.84442	360.85999	ClC1=CC(Cl)=C(Cl)C(Cl)=C1C1=CC=CC(Cl)=C1Cl	WDLTVNW
24	DTXSID8074239	74472-51-8	2,3,3',4,5,5',6-Heptachl	2,3,3',4,5,5',6-Heptachl	C12H3Cl7	391.80544	395.31	ClC1=CC(=CC(Cl)=C1)C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1Cl	ZUTDUGMN
25	DTXSID8074237	74472-49-4	2,2',3,4,5,6,6'-Heptachl	2,2',3,4,5,6,6'-Heptachl	C12H3Cl7	391.80544	395.31	ClC1=CC=CC(Cl)=C1C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1Cl	FGDPOTMI
26	DTXSID8074235	74472-47-2	2,2',3,4,4',5,6-Heptachl	2,2',3,4,4',5,6-Heptachl	C12H3Cl7	391.80544	395.31	ClC1=CC(Cl)=C(C=C1)C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1Cl	DJEUXBQA
27	DTXSID8074233	74472-44-9	2,3,3',4',5,6-Hexachlorc	2,3,3',4',5,6-Hexachlorc	C12H4Cl6	357.84442	360.85999	ClC1=CC=C(C=C1Cl)C1=C(Cl)C(Cl)=CC(Cl)=C1Cl	ZAGRQXM

Delivering our Chemistry Data

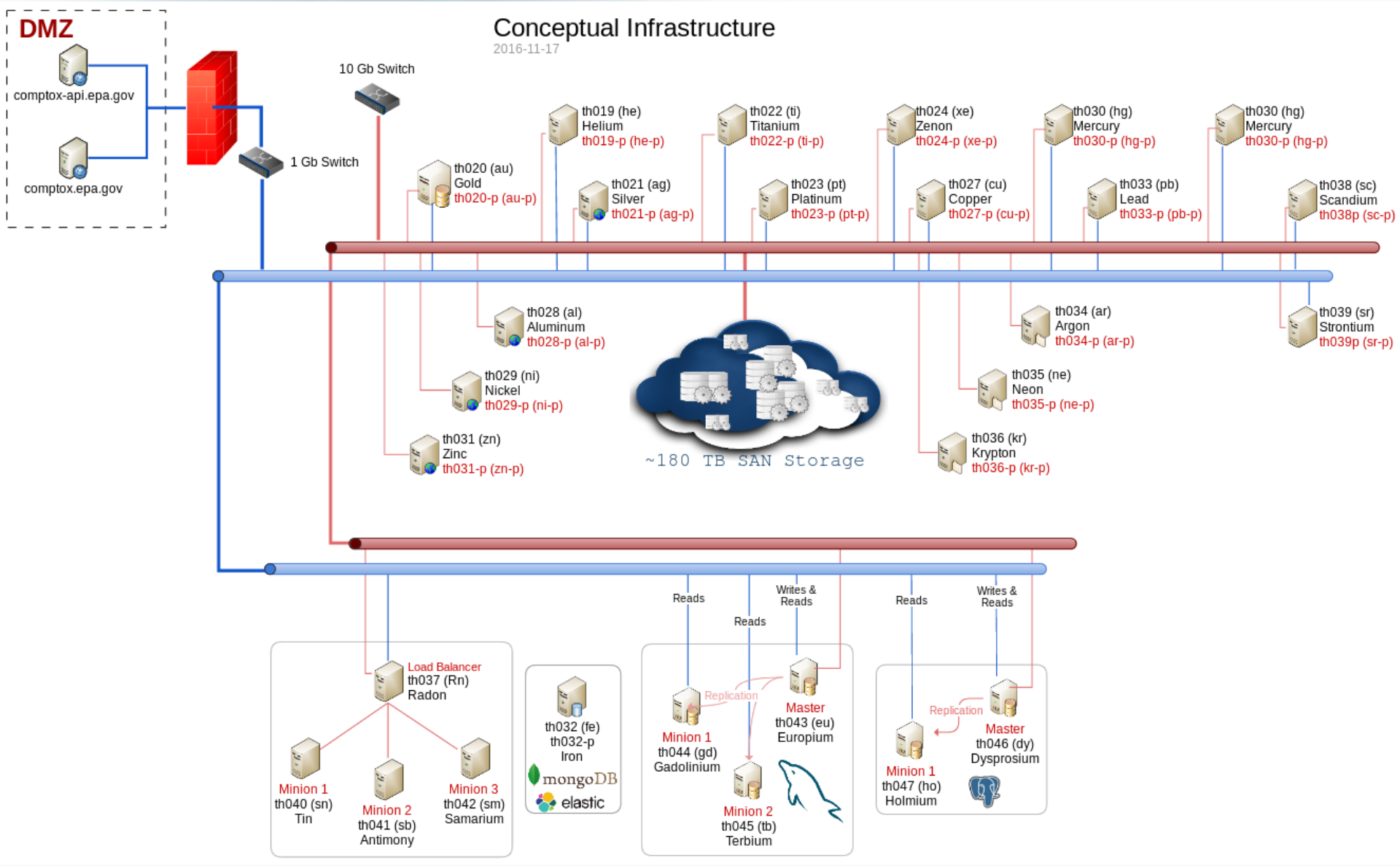


- Various types of data at FTP download site: ftp://newftp.epa.gov/COMPTOX/Sustainable_Chemistry_Data/Chemistry_Dashboard

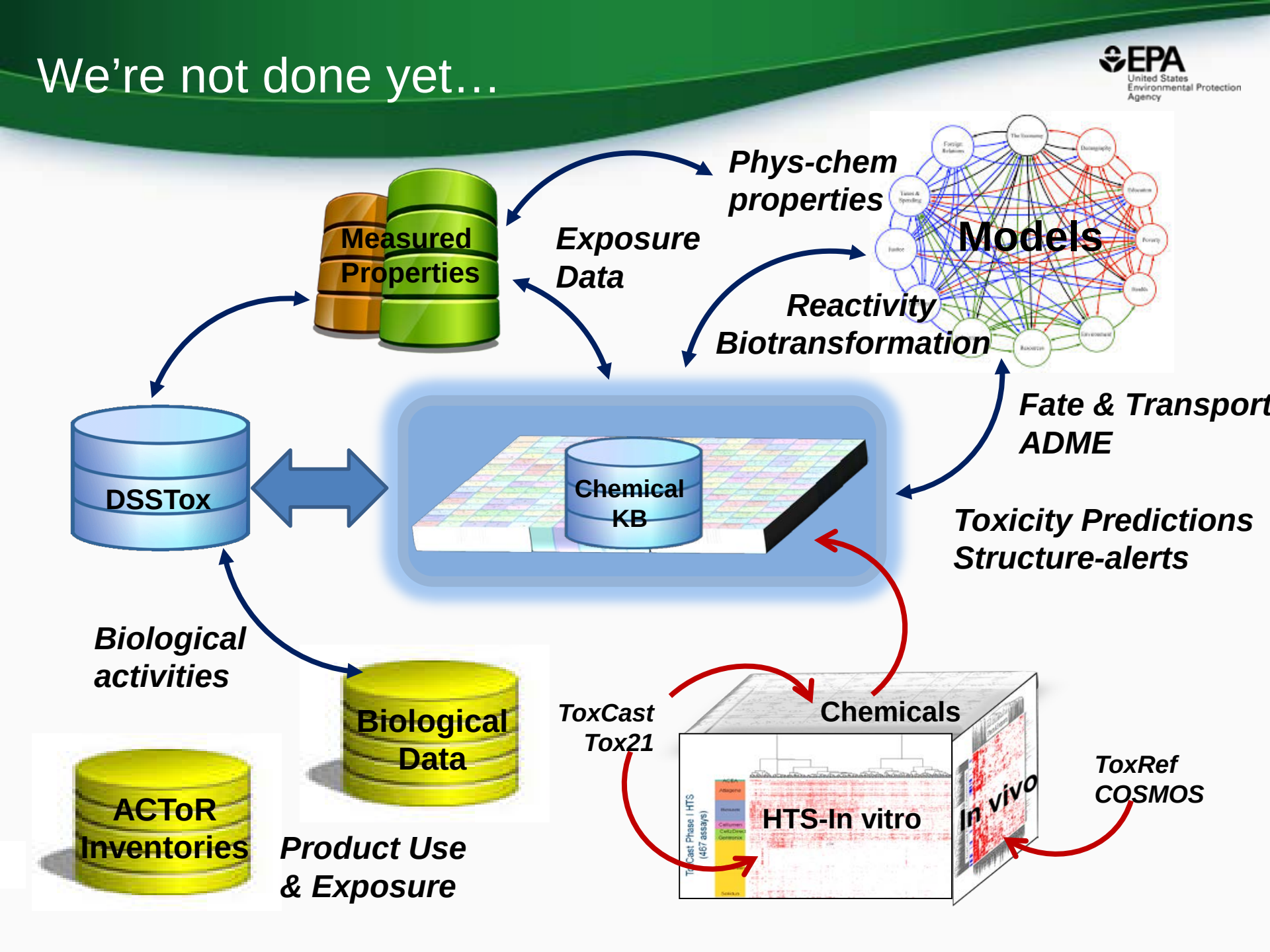
Open Data Reuse on iOS



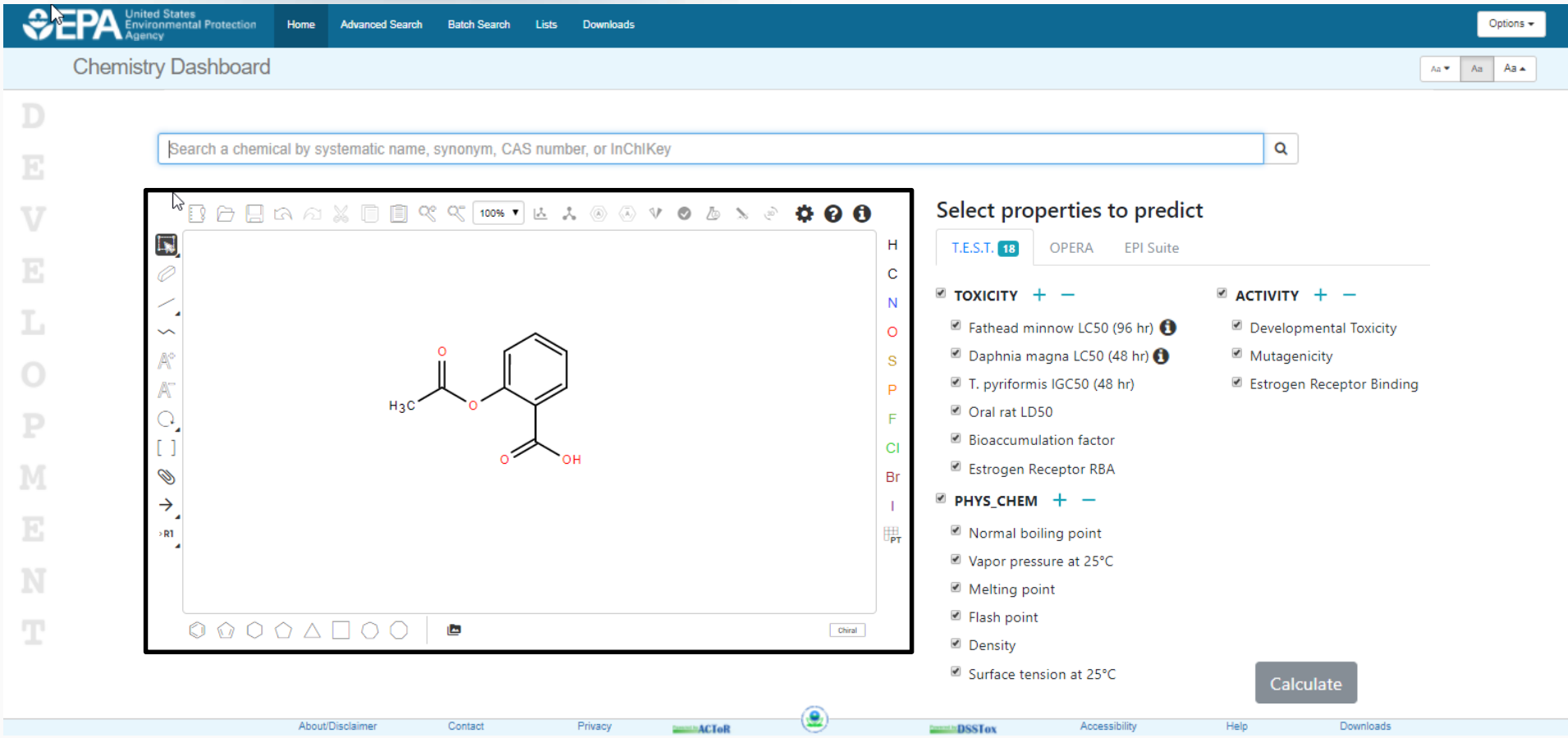
Computational Infrastructure



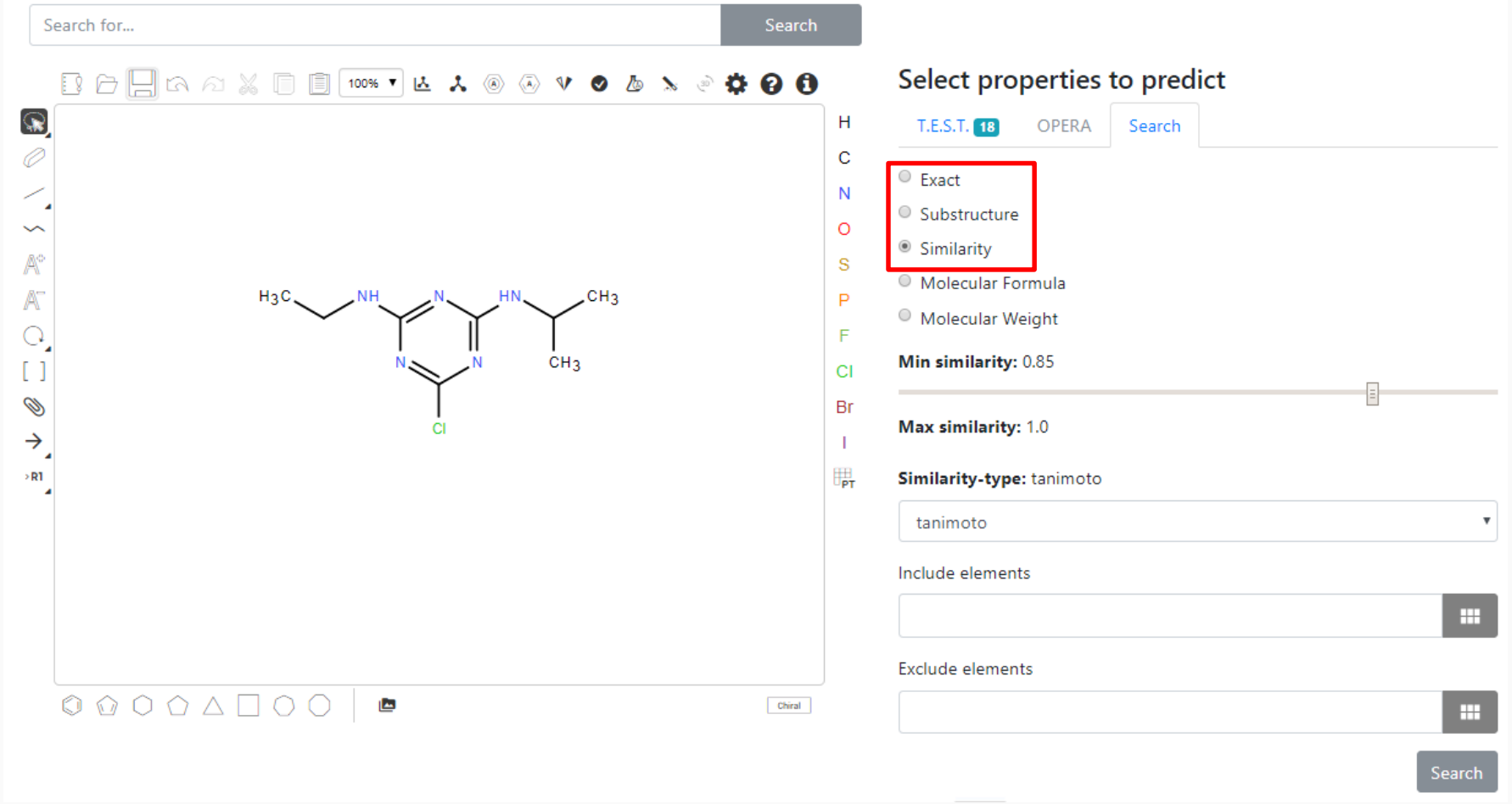
We're not done yet…



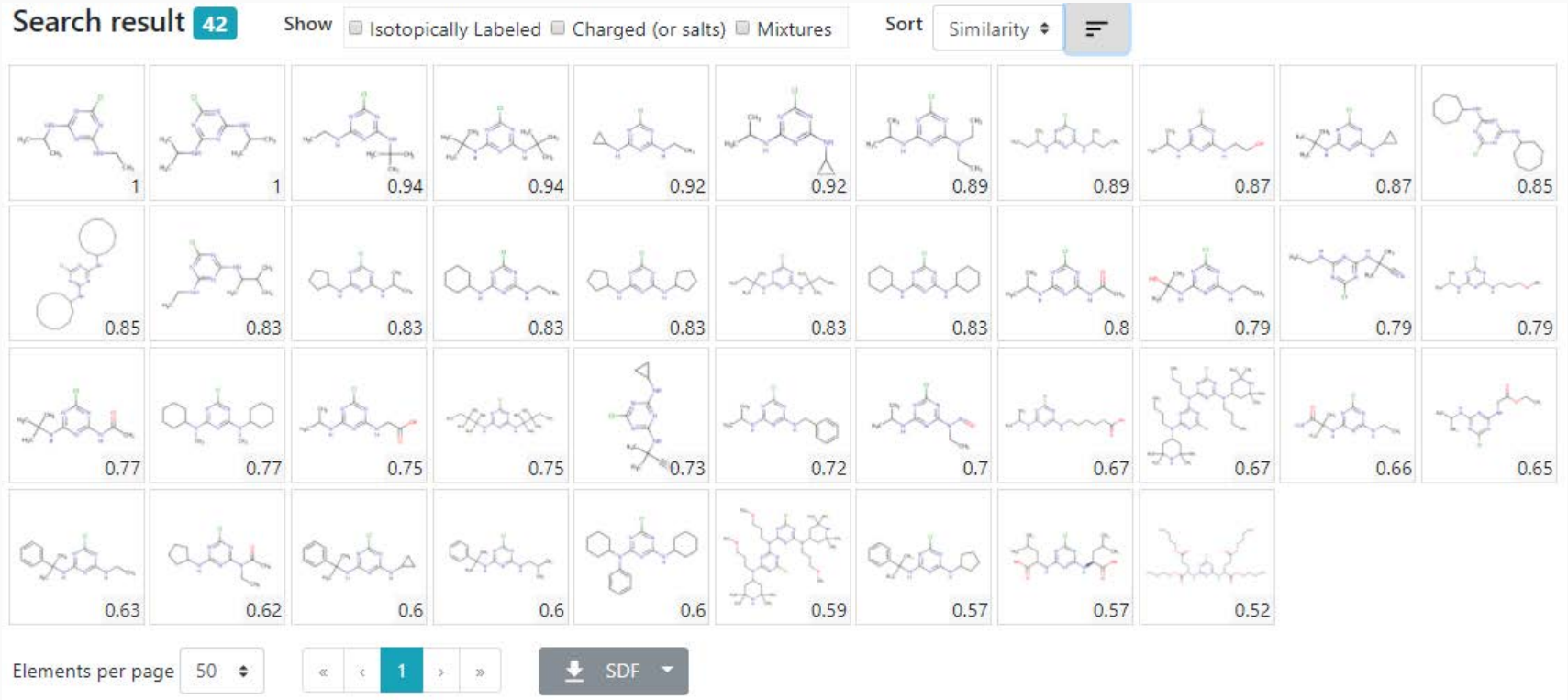
Real time predictions



When you have a structure editor…



Structure, **Substructure**, Similarity Searching



Future Work

- Existing dashboards functionality to be added to the “CompTox ~~Chemistry~~ Dashboard”
- Open Application Programming Interface and web services
- Integration to other agency databases

- NEXT Release 1st week of March – Society of Toxicology and ACS Spring Meetings

Acknowledgements

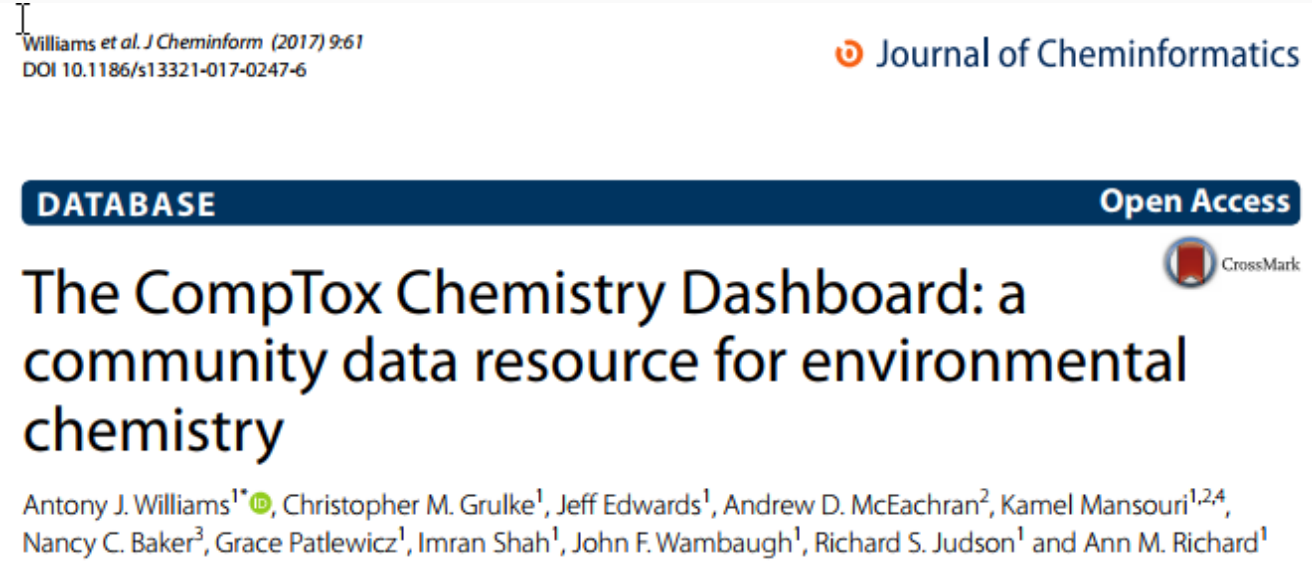
EPA-RTP

An enormous team of contributors from NCCT

and collaborators from
NERL
NHERL
NRMRL

Contact

Antony Williams
NCCT, US EPA Office of Research and Development,
Williams.Antony@epa.gov
ORCID: https://orcid.org/0000-0002-2668-4821



https://doi.org/10.1186/s13321-017-0247-6