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

The EPA Comptox Chemistry Dashboard: A Web-Based Data Integration Hub

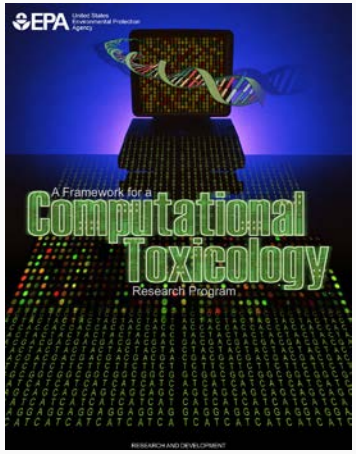
Antony Williams

U.S. Environmental Protection Agency, RTP, NC

Luxembourg Centre for Systems Biomedicine

27th November 2017

National Center for Computational Toxicology



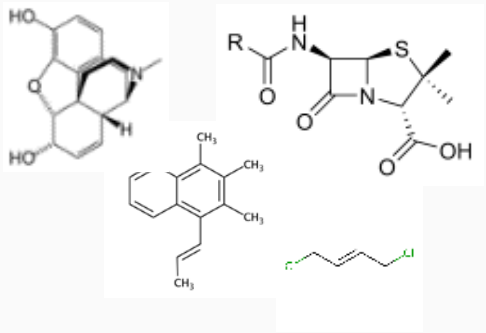
- National Center for Computational Toxicology established in 2005 to integrate:
  - High-throughput and high-content technologies
  - Modern molecular biology
  - Data mining and statistical modeling
  - Computational biology and chemistry
- Staffed by ~60 employees and contractors as part of EPA's Office of R&D
- 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

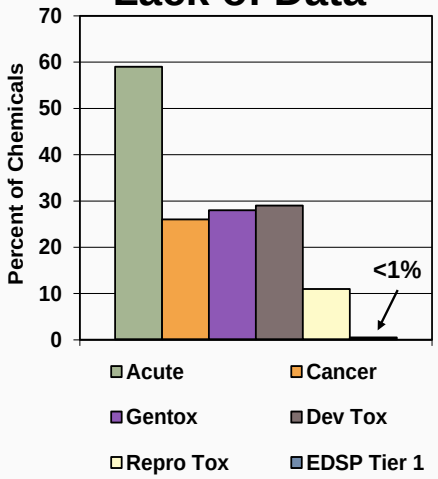
- Our team is broad in skills
  - Biologists, chemists and toxicologists
  - Bioinformaticians and cheminformaticians - modelers (QSAR, Deep Learning), text miners
  - Information technology team, software developers
  - A large IT support team for production applications (National Computing Center)

Why we must innovate…

Number of Chemicals /Combinations

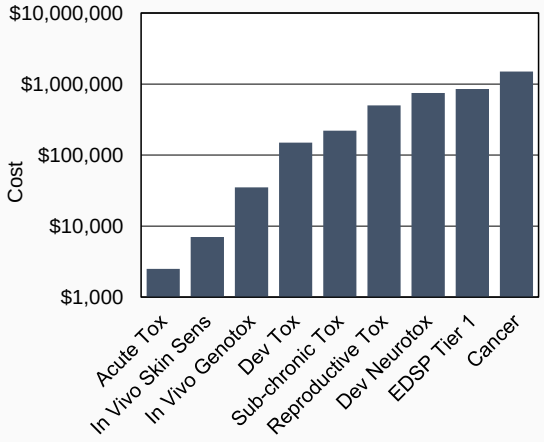


Lack of Data



Ethical Concerns

Economics



Toxic Substances Control List

- Inventory was initially published in 1979
- Second version, containing about 62k chemical substances, was published in 1982
- Continues to grow and now lists ~85k chemicals, about 15k are confidential business information

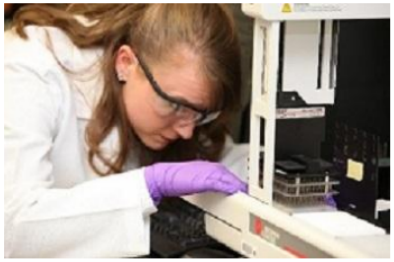
https://www.epa.gov/tsca-inventory/how-access-tsca-inventory

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.

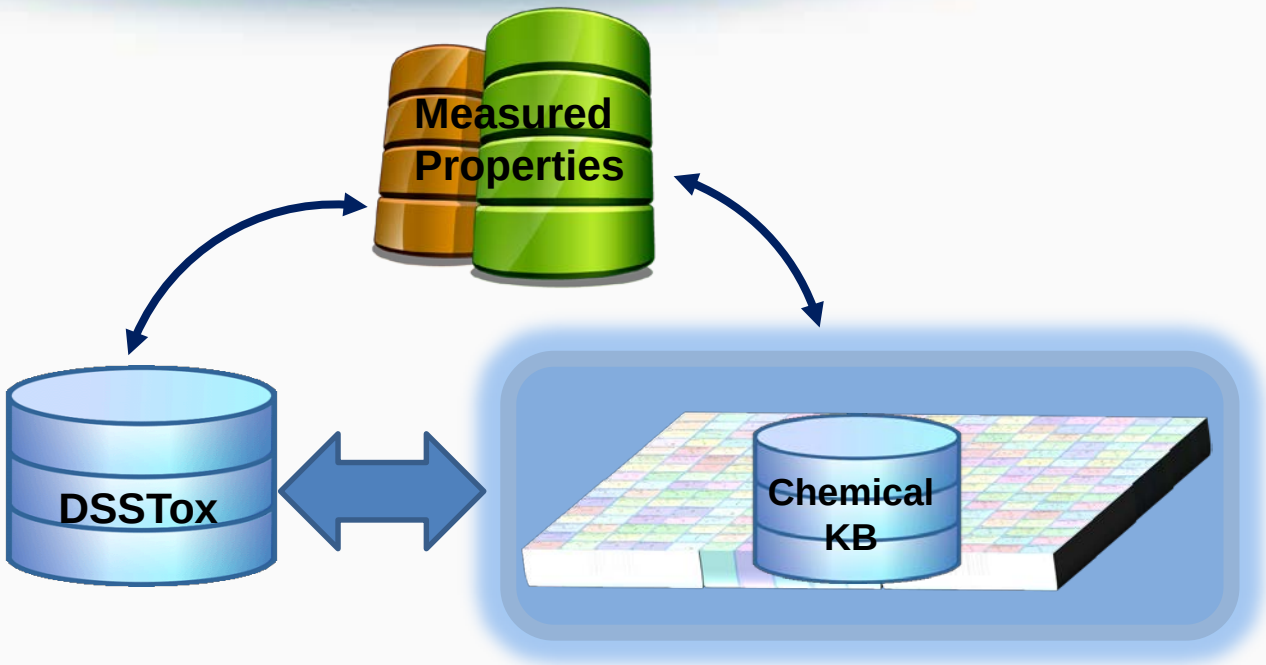
- ToxCast Data: High-throughput screening data on thousands of chemicals.

Rapid Exposure and Dose Data

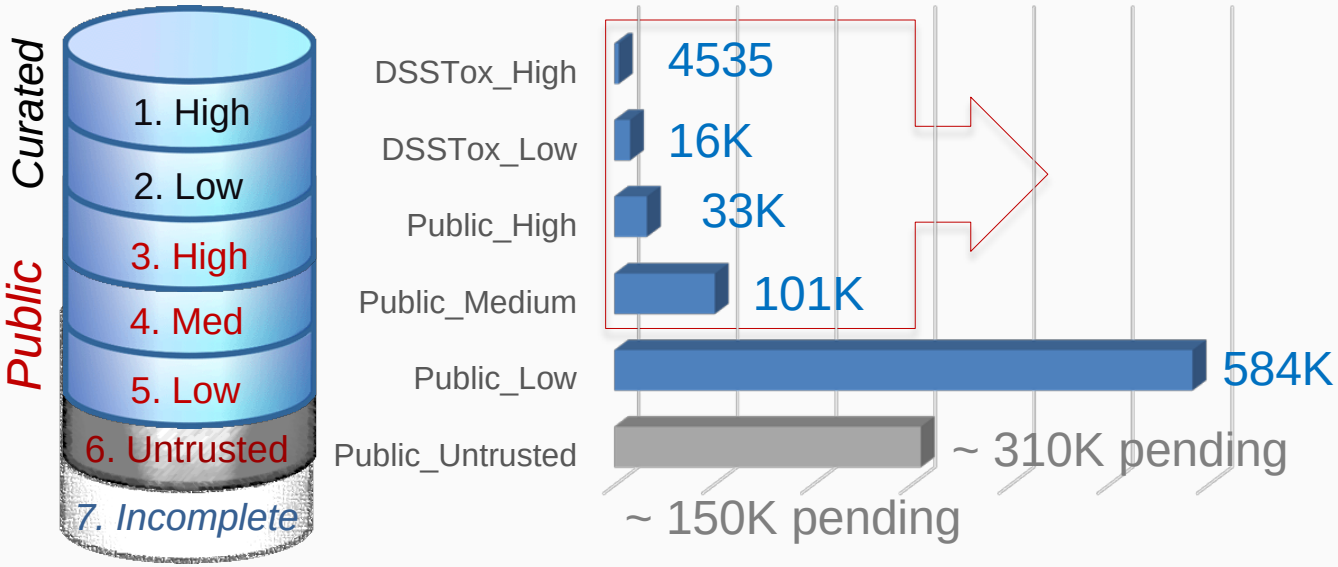
EPA researchers develop and use rapid exposure estimates to predict potential exposure for thousands of chemicals.

- 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



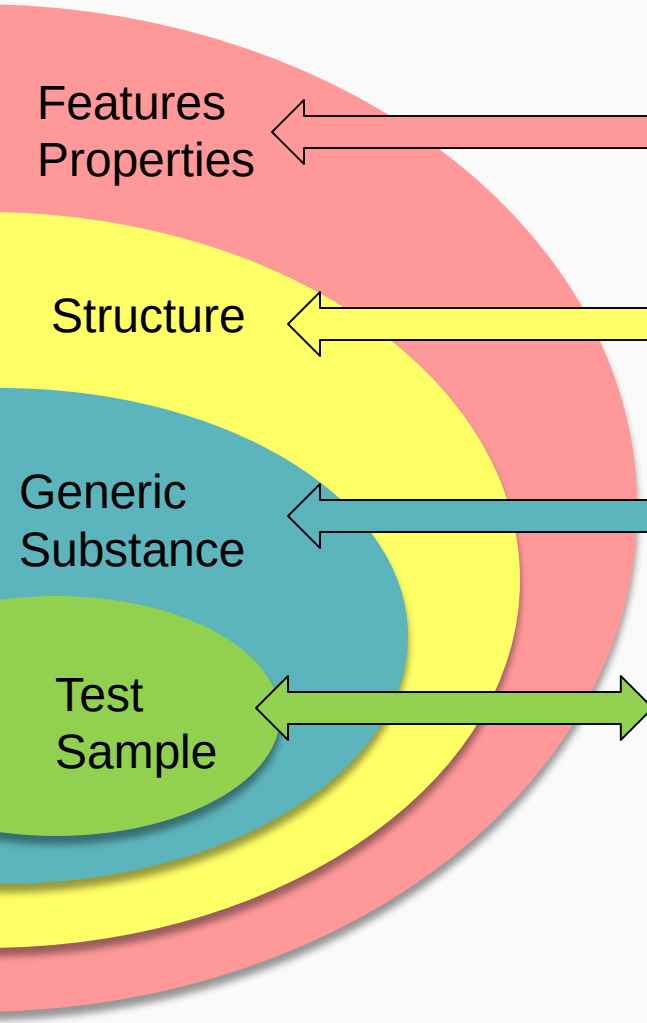
Approximately 17 Years of Data… Growing with daily curation



QC Levels

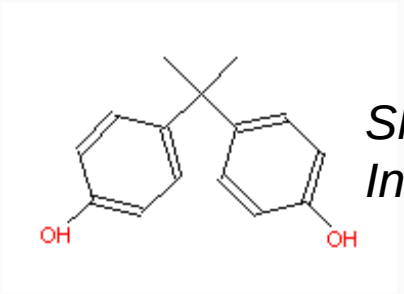
DSSTox_High:	Hand curated and validated
DSSTox_Low:	Hand curated and confirmed using multiple public sources
Public_High:	Extracted from EPA SRS and confirmed to have no conflicts in ChemID and PubChem
Public_Medium:	Extracted from ChemID and confirmed to have no conflicts in PubChem
Public_Low:	Extracted 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, ...

Chemical analogs, Read-across, SAR modeling



SMILES
InChI

Structure searching & modeling

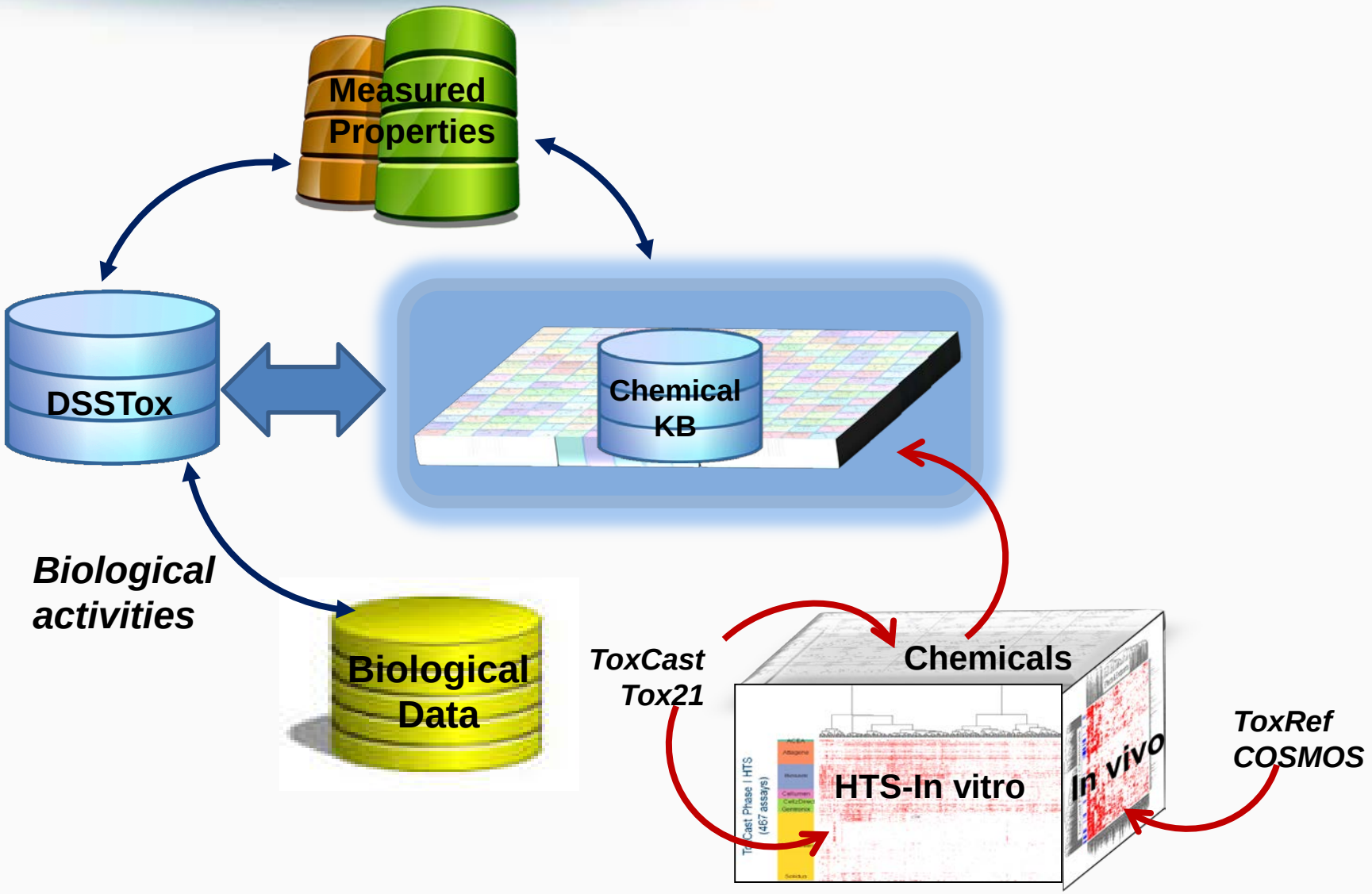
Chemical Name
CASRN

Public toxicity datasets

Supplier, Lot/Batch, physical description

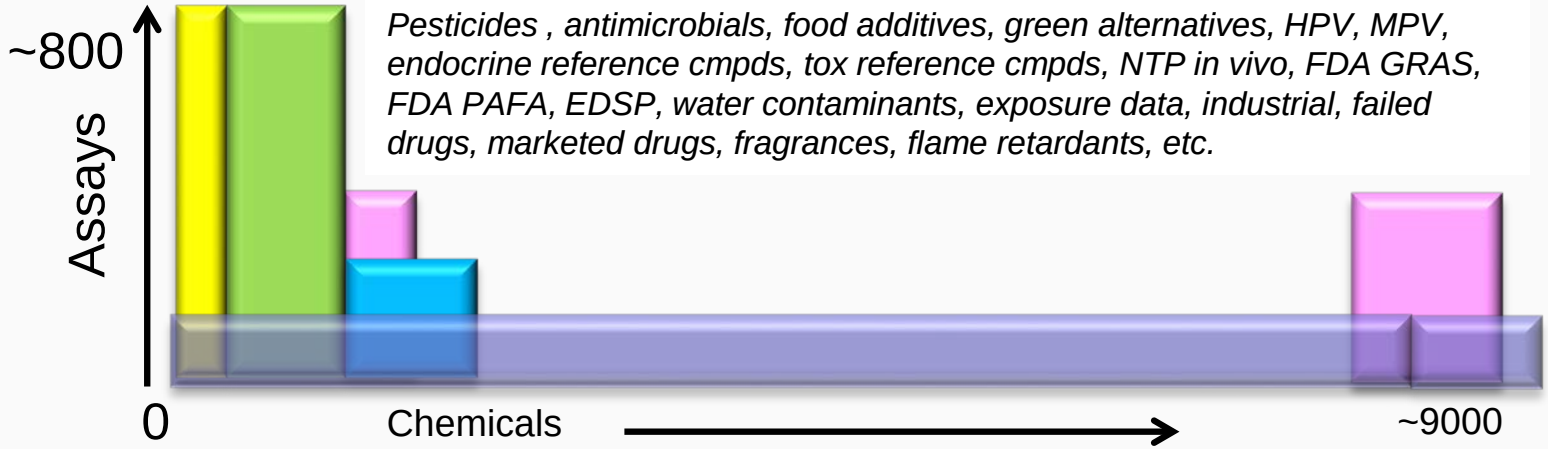
Experimental Endpoint Data

Integrating *in vitro* and *in vivo* data

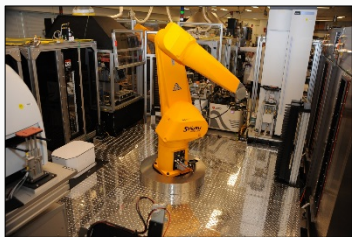


TEN YEARS of Assay Measurements: ToxCast & Tox21

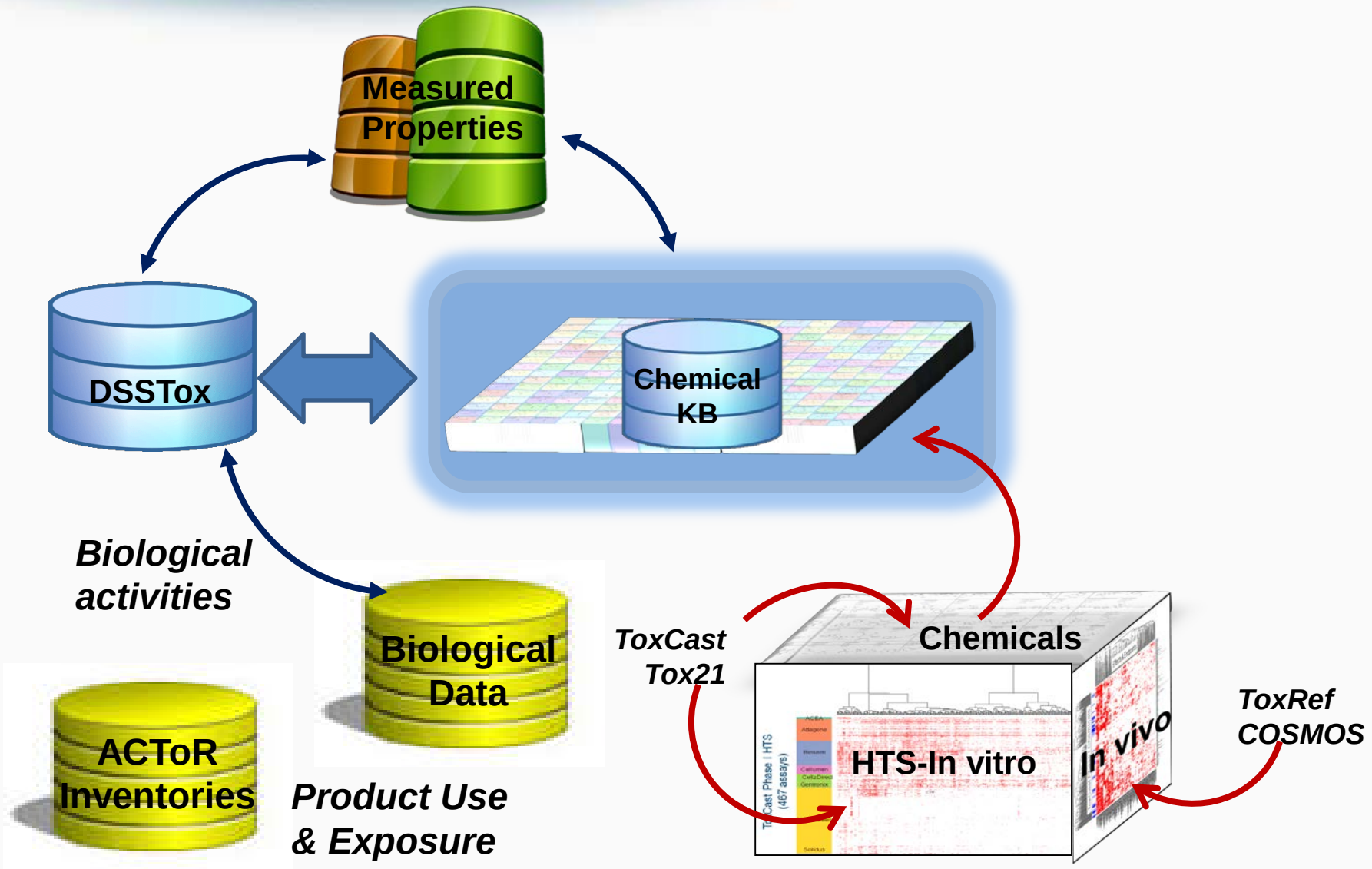
Set	Chemicals	Assays	Endpoints	Completion
ToxCast Phase I	293	~600	~700	2011
ToxCast Phase II	767	~600	~700	03/2013
ToxCast E1K	800	~50	~120	03/2013
ToxCast Phase III	~900	~300	~300	In progress
Tox21	~9000	~80	~150	In progress



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

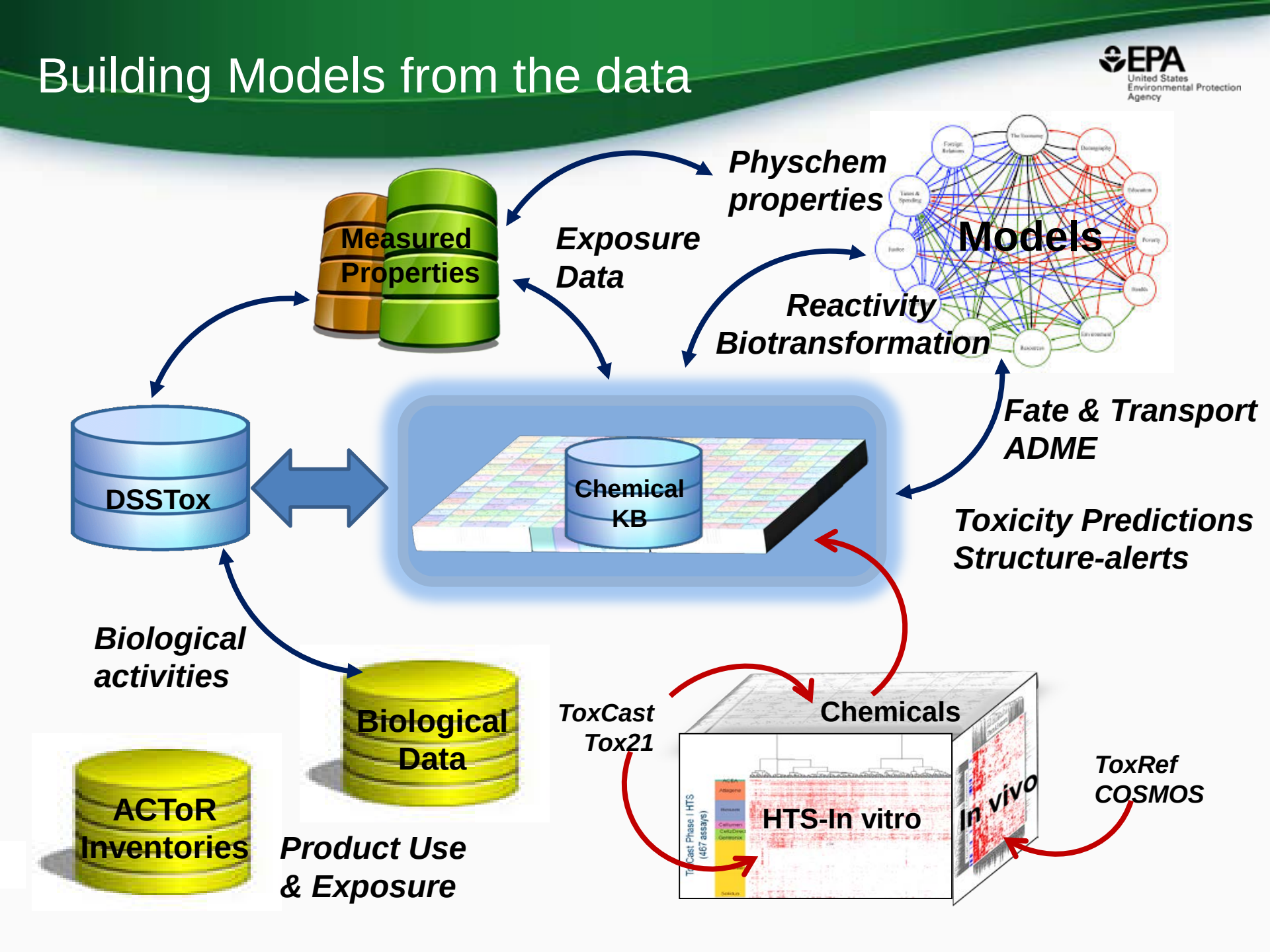
› Farfield Exposure Models

› Nearfield Exposure Models

Evaluating High-throughput Exposure Predictions

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

Building Models from the data



What we have learned…

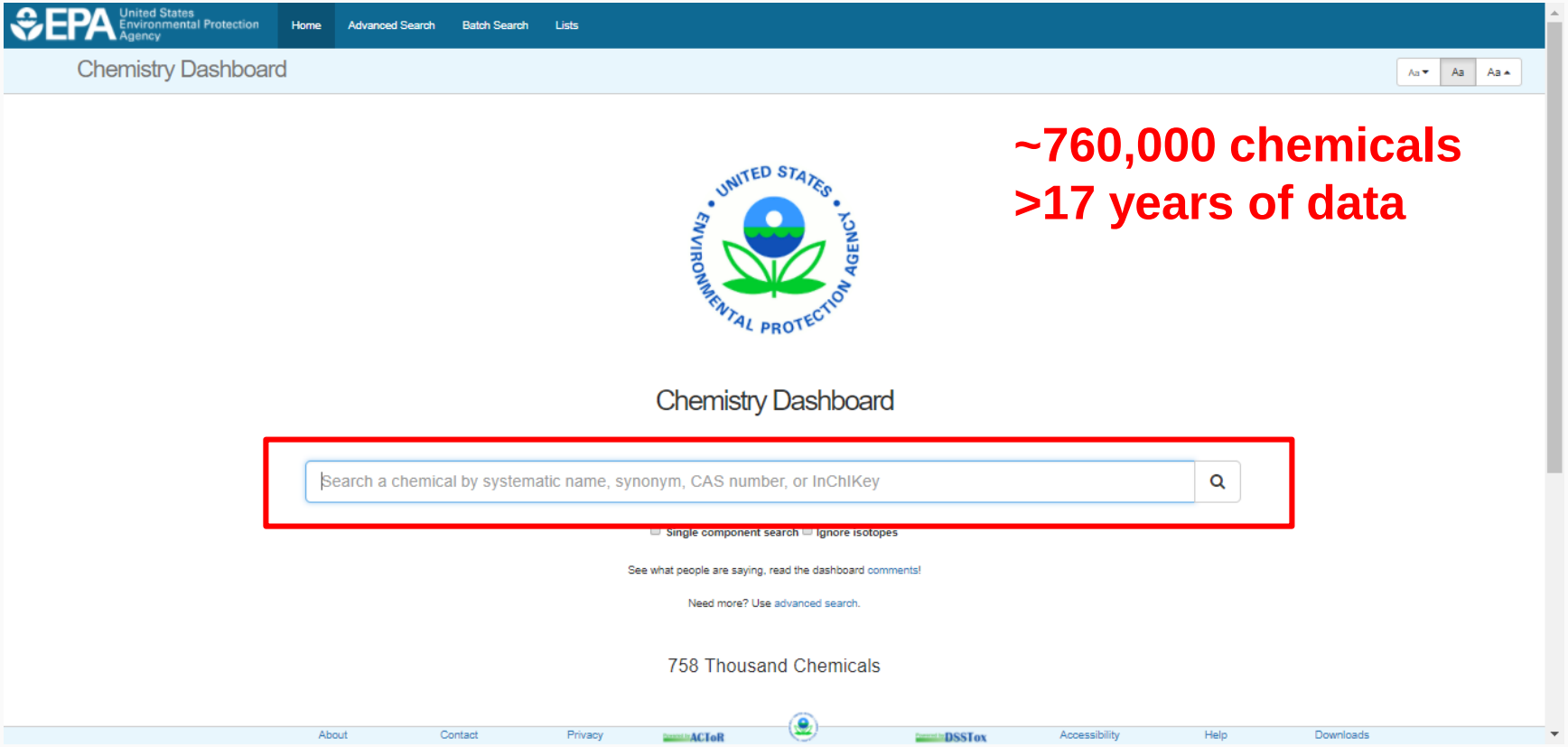
- Data curation, standardization and versioning is **essential**
- **Prototype** application development suffices **for research** projects
- **Production apps** need managed processes
- ODOSOS (Open Data, Open Source and Open Standards) endows many benefits

The CompTox Chemistry Dashboard: An Overview

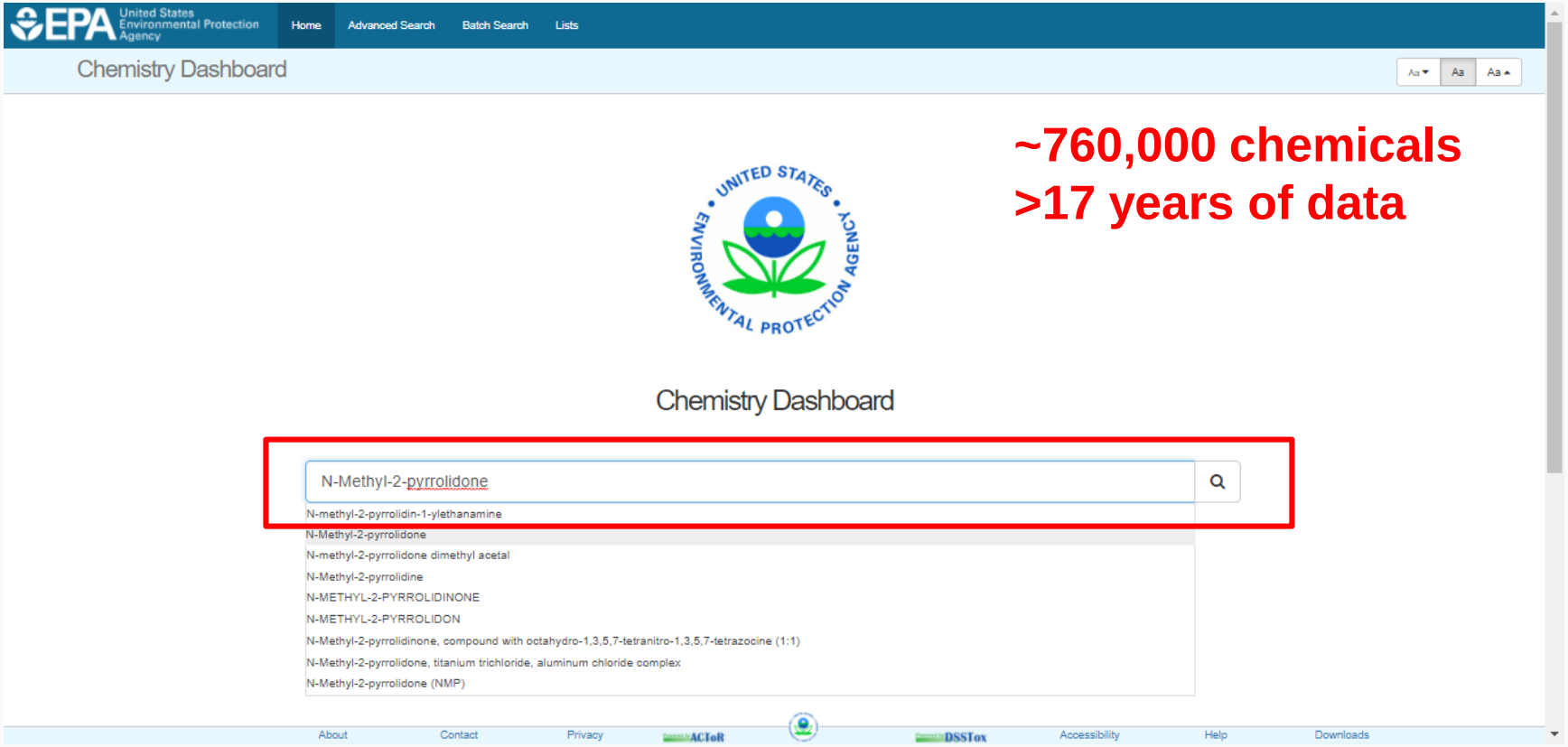
- A **publicly accessible website** delivering access:
  - ~760,000 chemicals and related property data
  - Links to other agency websites and public data resources
  - “Literature” searches for chemicals using public resources
  - Integration to “biological assay data” for 1000s of chemicals
  - Information regarding consumer products containing chemicals
  - “Batch searching” for thousands of chemicals

- Day-to-day curation efforts for data quality

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



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

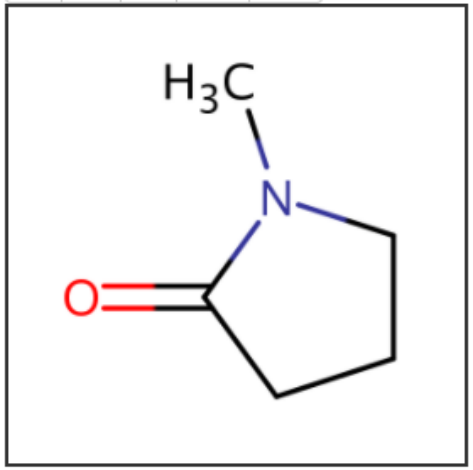


Chemical Page

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

Multiple Prediction Algorithms

- 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

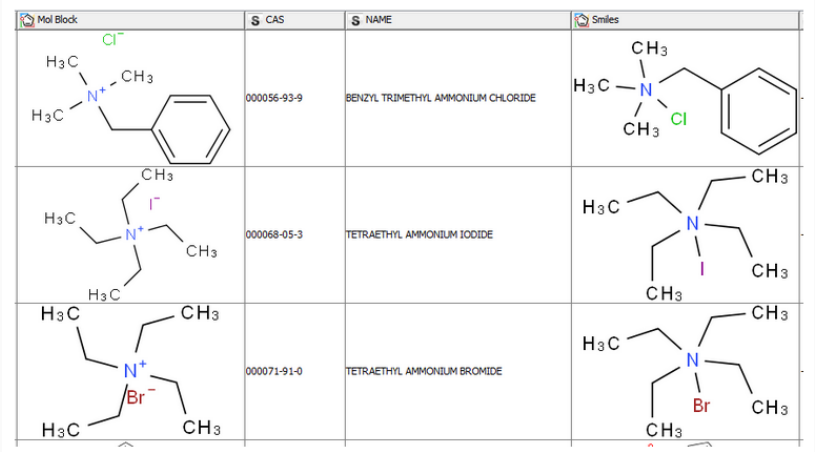
Predicted

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

Consuming and Curating Public Data

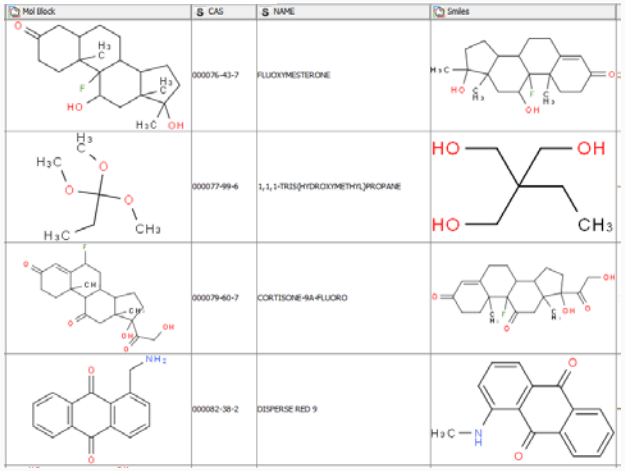
EPA United States Environmental Protection Agency

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

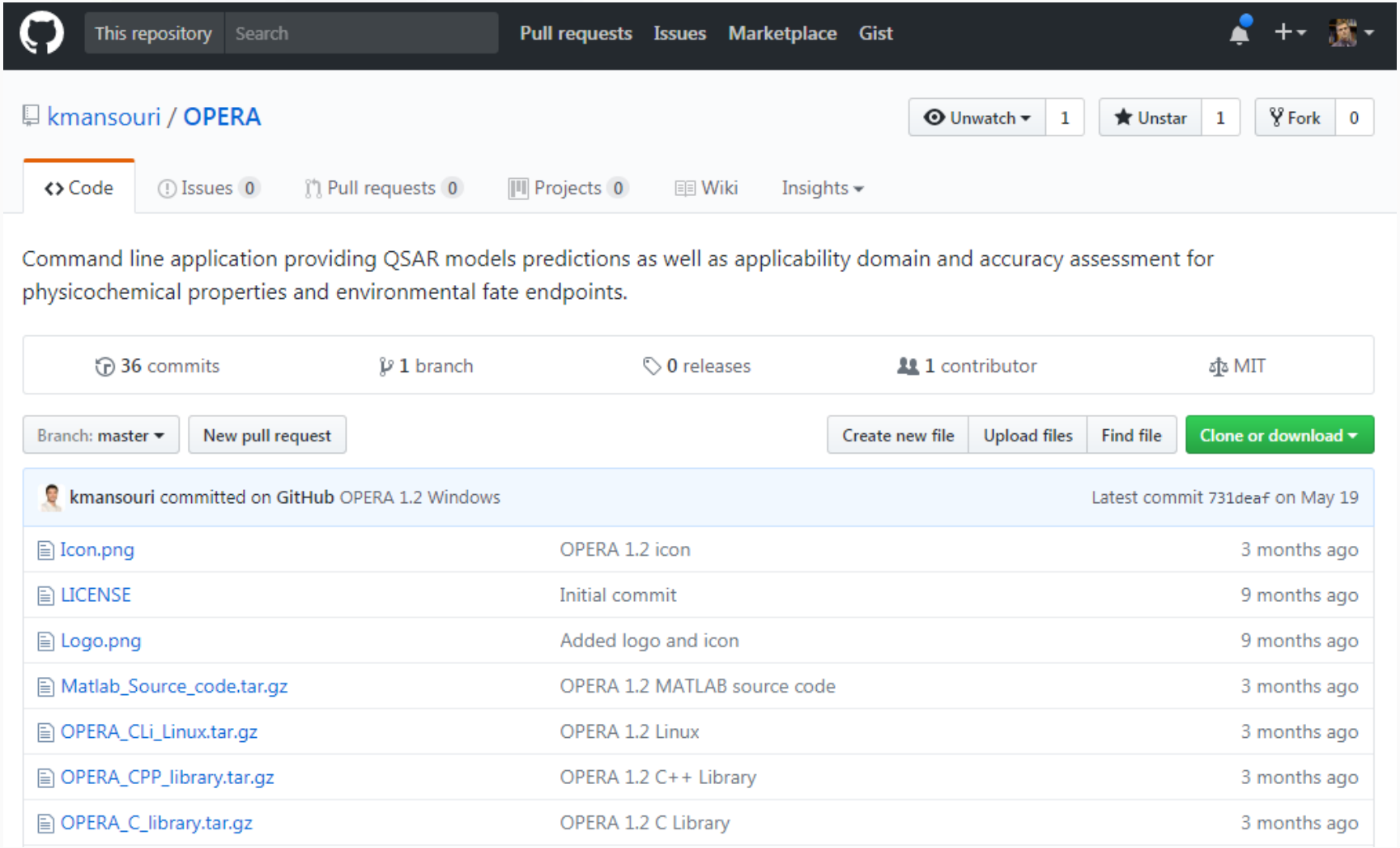
An automated curation procedure for addressing chemical errors and inconsistencies in public datasets used in QSAR modelling $

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 on GitHub

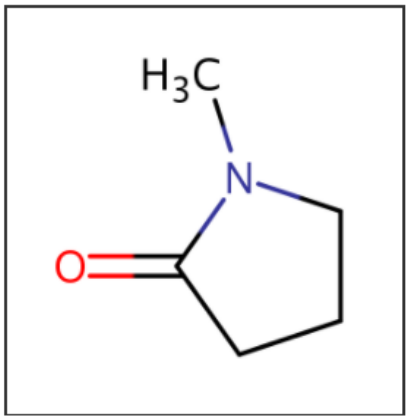


https://github.com/kmansouri/OPERA.git

OPERA Models: LogP: Octanol-Water

N-Methyl-2-pyrrolidone

872-50-4 | DTXSID6020856



Model Results

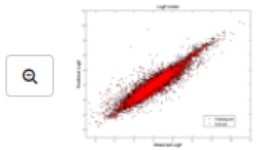
Predicted value: -0.300

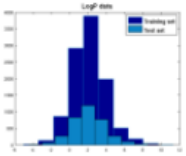
Global applicability domain: Inside

Local applicability domain index: 0.88

Confidence level: 0.81

Model Performance



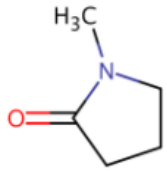


Weighted KNN model

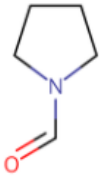
QMRF

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

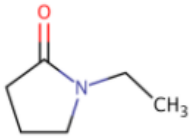
Nearest Neighbors from the Training Set



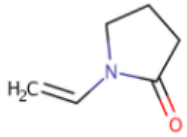
N-Methyl-2-pyrrolidone
Measured: -0.380
Predicted: -0.300



Prolinal
Measured: -0.320
Predicted: -0.226

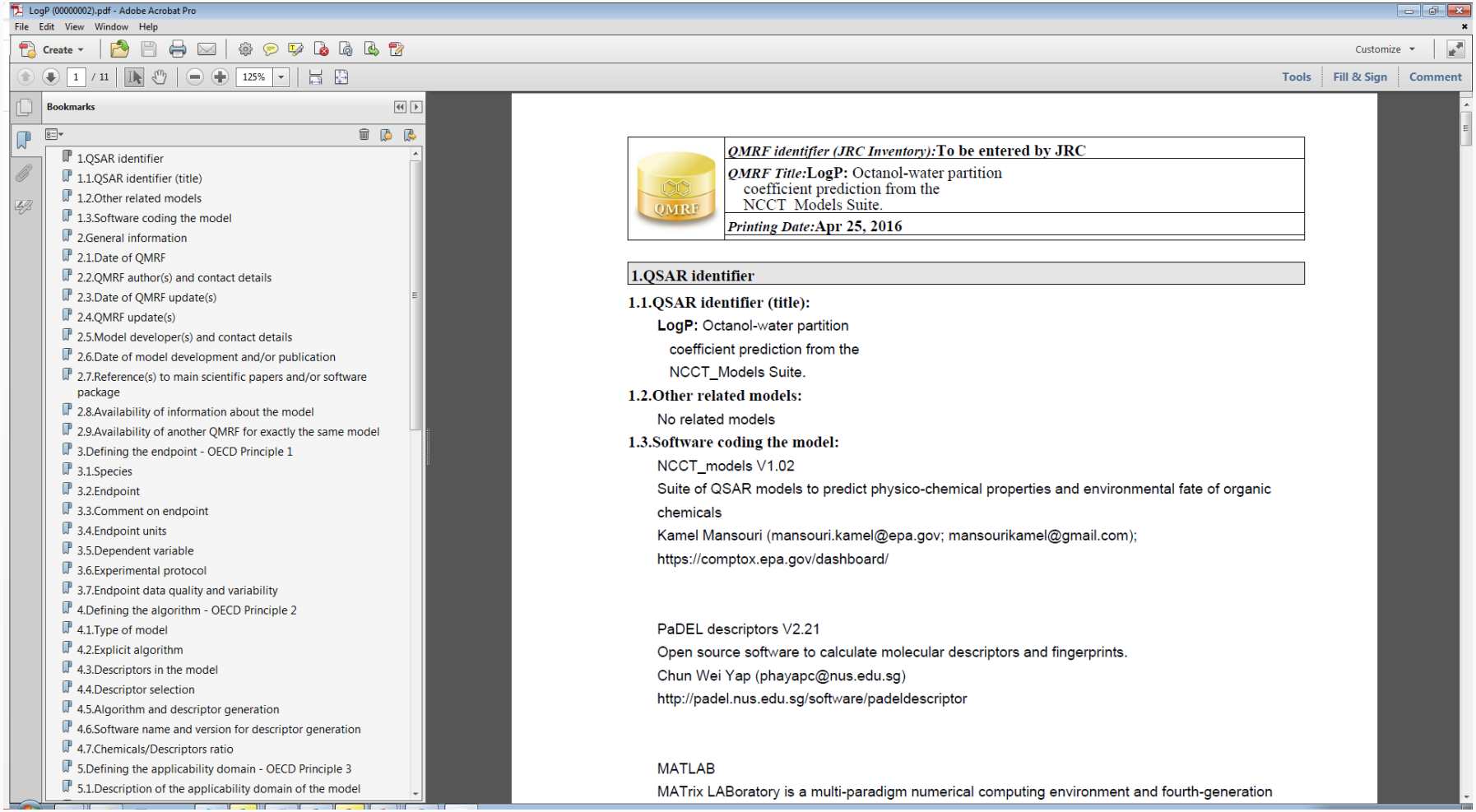


1-Ethyl-2-pyrrolidinone
Measured: -0.0400
Predicted: -0.211



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

QSAR Modeling Reporting Format



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

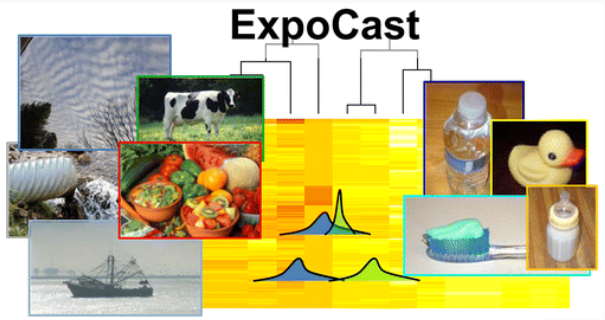
Exposure

EPA United States Environmental Protection Agency

National Health and Nutrition Examination Survey

High-Throughput Models for Exposure-Based Chemical Prioritization in the ExpoCast Project

Environ. Sci. Technol., **2013**, *47* (15), pp 8479–8488



Product & Use Categori...

Chemical Weight Fraction

Chemical Functional Use

Monitoring Data

Exposure Predictions

Download as: TSV | Excel

National Health and Nutrition Examination Survey (NHANES) Inferences (mg/kg-bw/day)

	Lower 95th Limit	Upper 95th Limit	Median
Ages 6-11	3.80e-05	4.92e-05	4.33e-05
Ages 12-19	2.55e-05	3.38e-05	2.93e-05
Ages 20-65	2.79e-05	3.27e-05	3.02e-05
Ages 65+	1.91e-05	2.31e-05	2.10e-05
BMI > 30	2.38e-05	2.74e-05	2.55e-05
BMI < 30	3.02e-05	3.30e-05	3.16e-05
Repro. Age Females	2.83e-05	3.31e-05	3.06e-05
Females	2.58e-05	3.03e-05	2.80e-05
Males	2.94e-05	3.37e-05	3.15e-05
Total	2.86e-05	3.08e-05	2.97e-05

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

Product Composition Details

EPA United States Environmental Protection Agency

Chemical Weight Fractions

Download as: TSV | Excel

Product Name	Product Use Category	Minimum Weight Fraction	Maximum Weight Fraction	Data Type	Source
citristrip canadian stripping...	home maintenance: stripper	0.65	0.7	MSDS	Retail Product Categories/...
citristrip stripping gel qcg7...	home maintenance: stripper	0.4	0.55	MSDS	Retail Product Categories/...
gumout 2 part professional...	auto products: auto fluids a...	0.3	0.4	MSDS	Retail Product Categories/...
minwax water based wipe ...	home maintenance: finish	0.06	0.06	MSDS	Retail Product Categories/...
10-02199- calico tip & glue...	personal care: nail polish r...	0.01	0.05	MSDS	Retail Product Categories/...
artificial nail remover 728 1	personal care: nail polish r...	0.01	0.05	MSDS	Retail Product Categories/...
calico tip & glue remover 1	personal care: nail polish r...	0.01	0.05	MSDS	Retail Product Categories/...
kiss nail remover 1	personal care: nail polish r...	0.01	0.05	MSDS	Retail Product Categories/...
waterborne clear wood fini...	home maintenance: finish	0.01	0.05	MSDS	Retail Product Categories/...

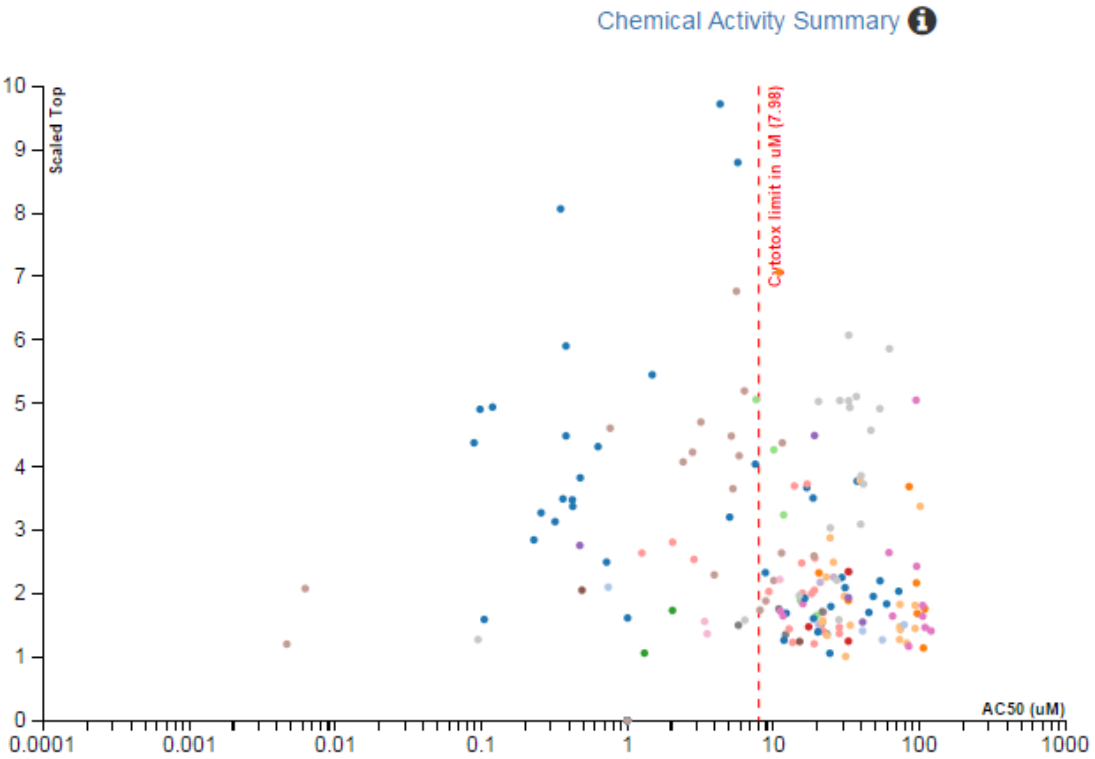
ToxCast and Tox21 Bioassays

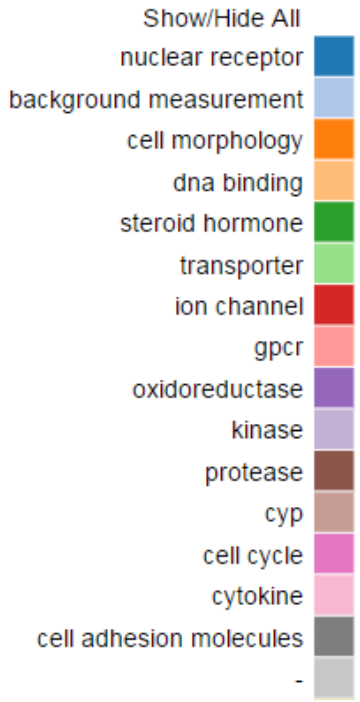
EPA United States Environmental Protection Agency

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

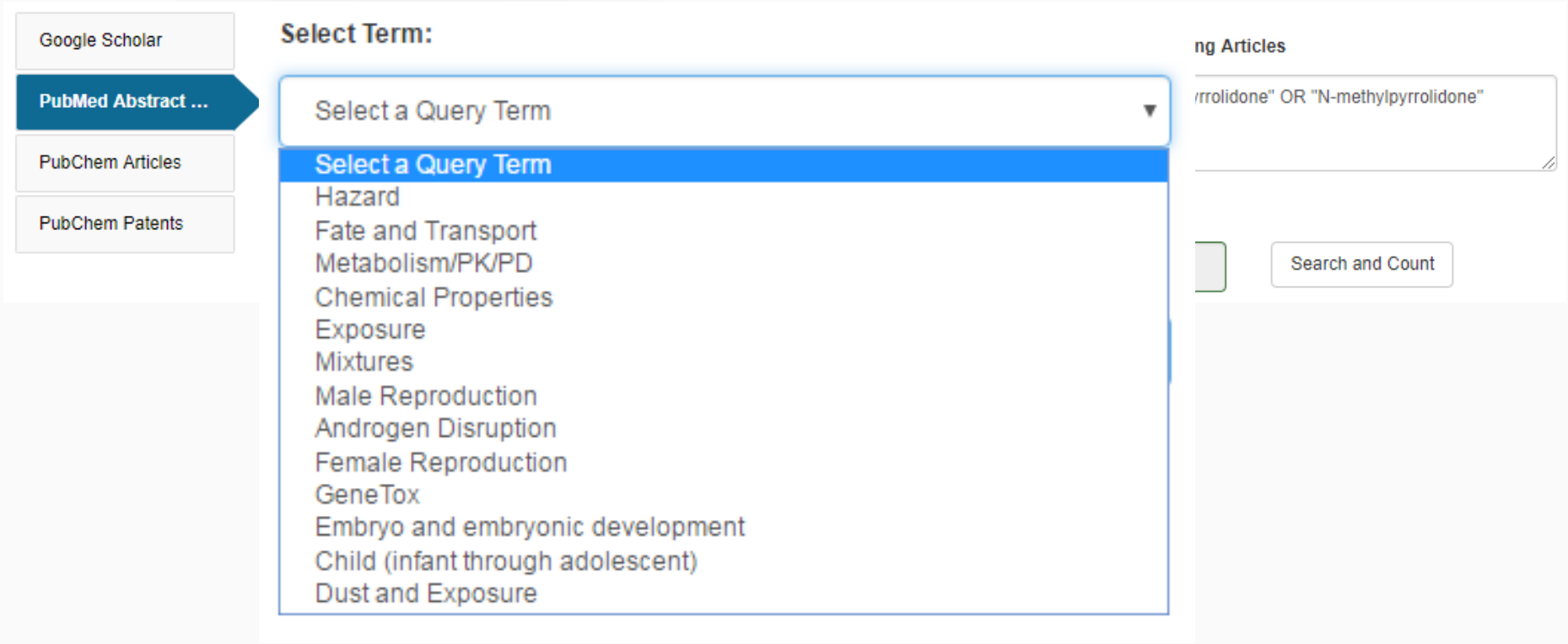
Effects of 'inactive'ingredients on bees
CA Mullin - Current Opinion in Insect Science, 2015 - Elsevier
... tallow amines, organosilicone ethoxylates and co-solvents such as **N-methyl-2-pyrrolidone** (NMP) [7 ... of tools to environmentally monitor residues and determine their relative **hazard** to 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

Pubmed Abstract Sifter

EPA United States Environmental Protection Agency

Select Term:

Exposure

Retrieve Articles | 0 Articles

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)

Add additional query terms to filter abstracts:

developmental | reproductive | occupational | Search and Count

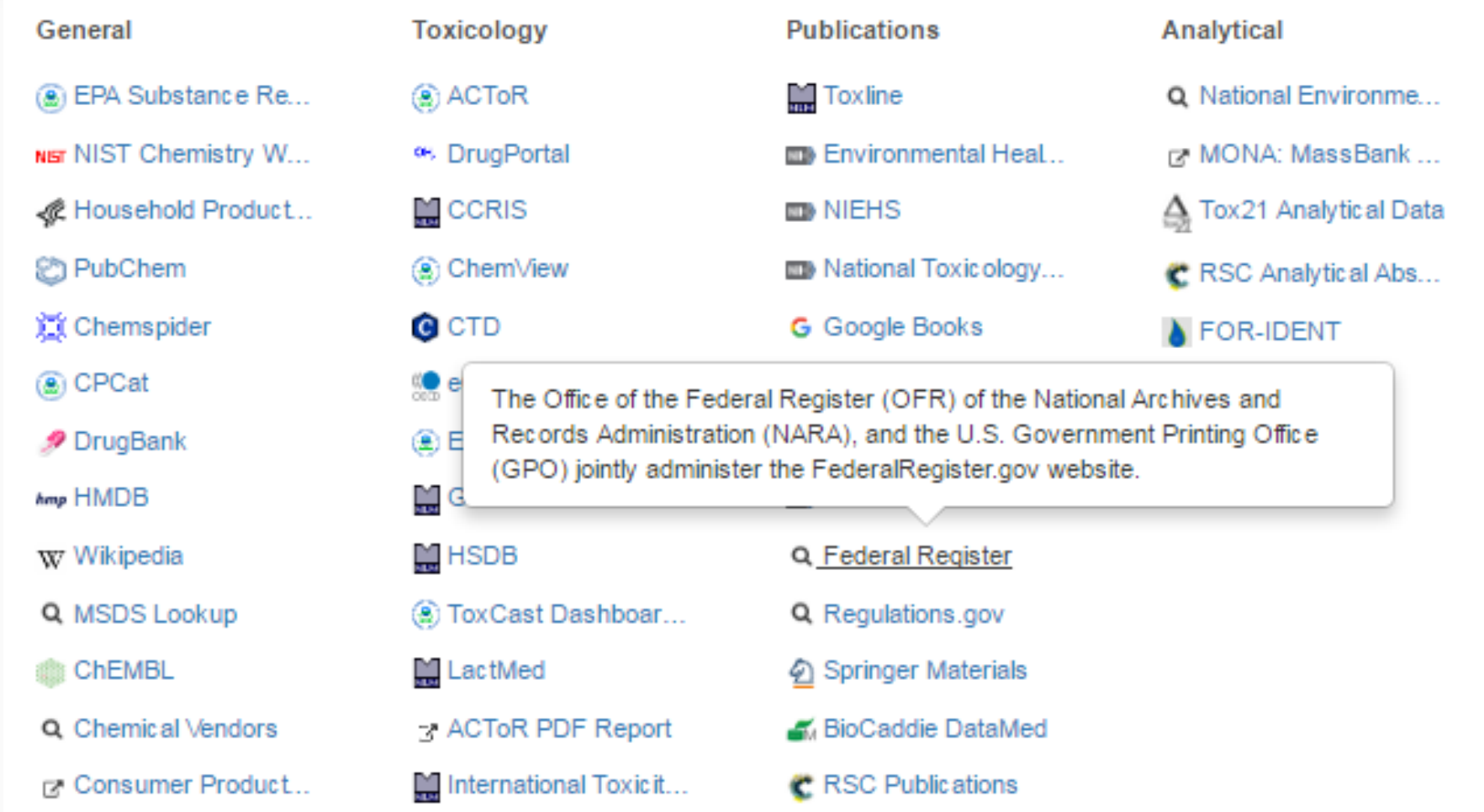
dev...	rep...	occ...	Total	PMID	Pu...	Title
0	0	1	1	24078144	2013	Biological monitoring and health effects of low-level exposure to N-methyl-2-pyrrolidone: a cross-sectional study.
1	0	0	1	23337464	2013	Biomonitoring of exposure to N-methyl-2-pyrrolidone in workers of the automobile industry.
0	0	2	2	19875680	2009	Quantitative risk analysis for N-methyl pyrrolidone using physiologically based pharmacokinetic and benchmark dos...
0	0	1	1	16897094	2006	Human volunteer study on the influence of exposure duration and dilution of dermally applied N-methyl-2-pyrrolidon...
0	0	0	0	16362322	2005	Ambient monitoring and biomonitoring of workers exposed to N-methyl-2-pyrrolidone in an industrial facility

Record: 1 of 16

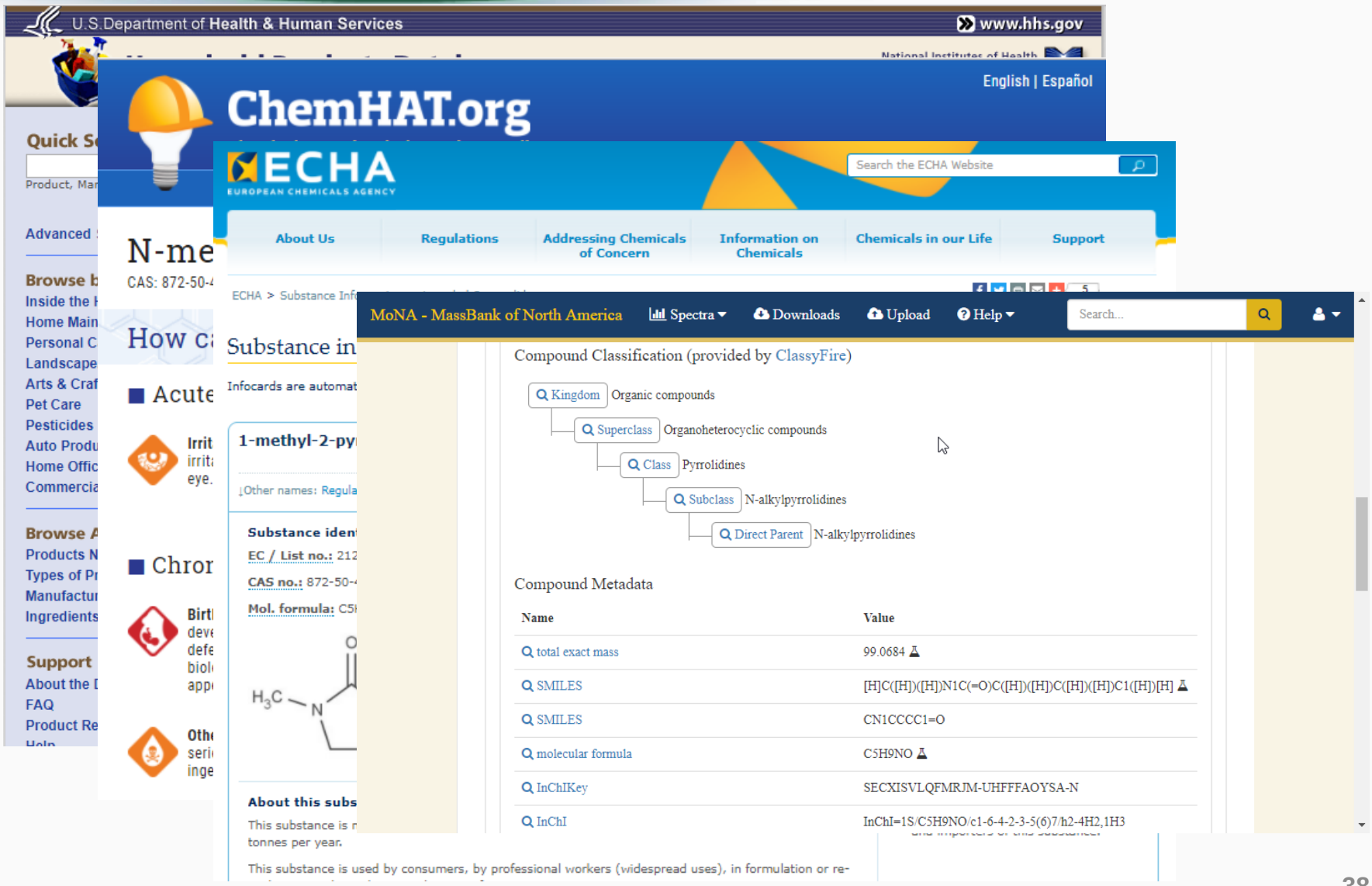
Title: Stillbirth after occupational exposure to N-methyl-2-pyrrolidone. A case report and review of the literature.

Abstract: N-methyl-2-pyrrolidone is a solvent that is increasingly used in a variety of industries, including petroleum refining, microelectronics, pesticide formulation, and veterinary medicine. Animal studies have demonstrated fetotoxic effects after maternal exposure to doses that have minimal to no adverse effect on the mothers. The fetotoxicity comprises resorption, stillbirth, and low birthweight and delayed ossification in surviving young. We report a human case of intrauterine growth retardation followed by fetal demise at 31 weeks gestation. The mother was a laboratory worker with no other apparent risk factors, who sustained occupational exposure to N-methyl-2-pyrrolidone throughout the first trimester of pregnancy. Laboratory work and solvent exposure have both previously been associated with adverse reproductive outcomes. Laboratories and other industries that use suspected reproductive toxins should have reproductive health policies in place that allow for decision-making based on toxicologic review, exposure assessment, and medical evaluation. These policies should allow for voluntary removal of prospective parents until environmental assessment and controls are instituted.

Links to Other Resources



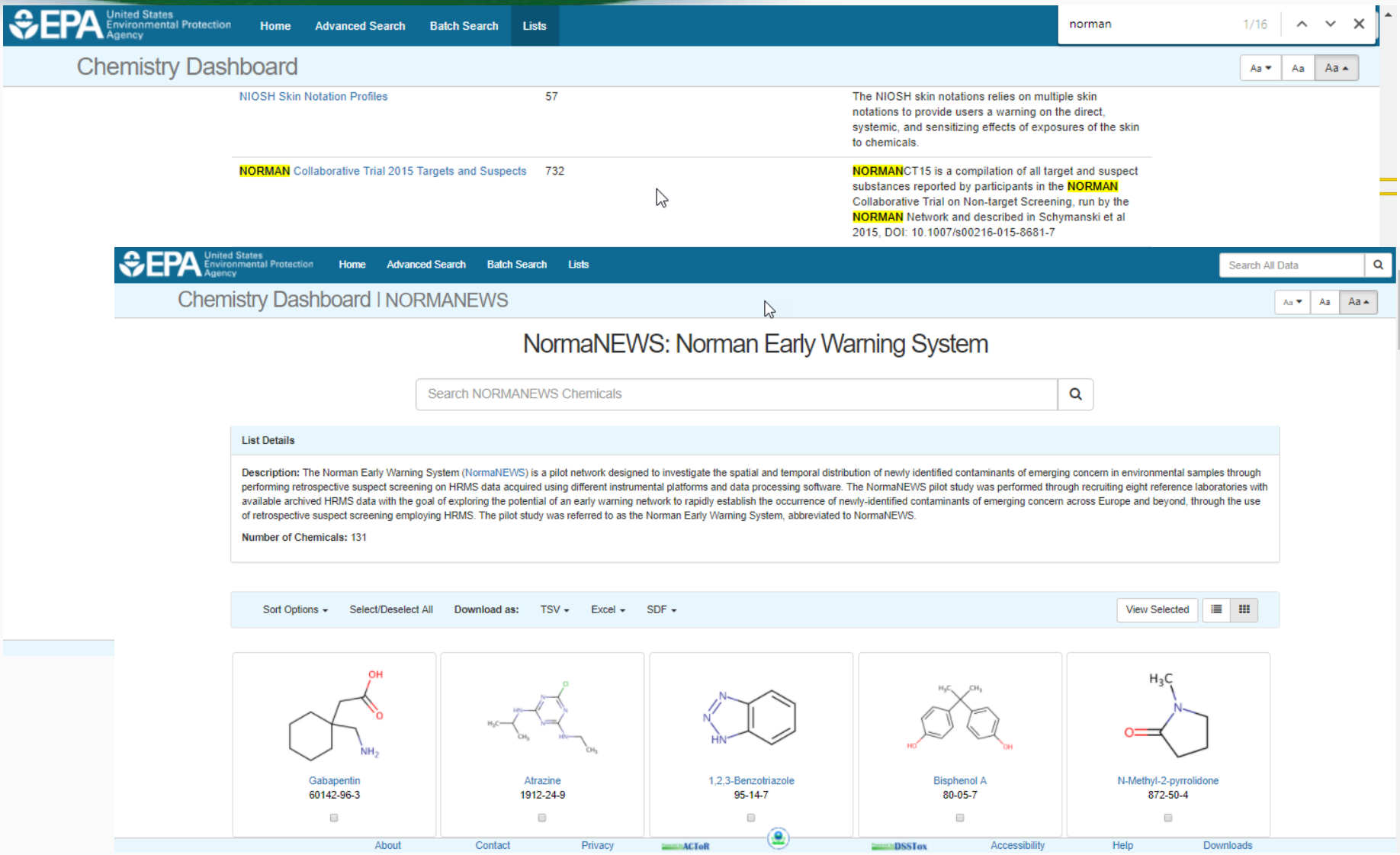
Example External Links…



Accessing Lists of Chemicals

- Build out definitive “lists” of chemicals
  - Algal toxins
  - Poly/perfluorinated chemicals
  - Pesticides
  - Toxcast screening chemical collection
  - Public data sets:
    - NORMAN Network data collections
    - MASSBANK data

The Collection of Lists



Batch Searching for Data for Thousands of Chemicals

- What are these chemicals?

Chemicals Used in the Hydraulic Fracturing Process in Pennsylvania
Prepared by the Department of Environmental Protection
Bureau of Oil and Gas Management

Updated June 10, 2010

Chemical	Product Name
2,2-Dibromo-3-Nitrilopropionamide	Bio Clear 1000/Bio Clear 2000/ Bio-Clear 200/BioRid20L/ EC6116A
2-methyl-4-isothiazolin-3-one	X-Cide 207
5-chloro-2-methyl-4-isothiazolin-3-one	X-Cide 207
Acetic Acid	Fe-1A Acidizing Composition/ Packer Inhibitor
Acetic Anhydride	Fe-1A Acidizing Composition
Acetylene	GT&S Inc./ Airco
Alcohol Ethoxylated	C12-16 NE-200
Alkyl benzene sulfonic acid	Tetrolite AW0007/ FR-46
Ammonia (aqueous)	FAW-5
Ammonium Bifluoride	ABF 37%
Ammonium Persulfate	AP Break
Ammonium Bisulfite	Techni-Hib 604/ Fe OXCLEAR/ Packer Inhibitor
Ammonium chloride	Salt Inhibitor
Ammonium Salt (alkylpolyether	

Batch Searching for Data for Thousands of Chemicals

EPA United States Environmental Protection Agency

Select Input Type(s)

- [x] Chemical Name
- [ ] CAS-RN
- [ ] InChIKey
- [ ] DSSTox Substance ID
- [ ] Exact Molecular Formula

Enter Identifiers to Search

Sodium l-glutamate
4-Hydroxy-3-nitrophenylarsonic acid
4-Nitro-1,2-phenylenediamine
Methoxypromazine
1-Phenyl-3-methyl-5-pyrazolone
1-Phenyl-2-thiourea
Phenylbutazone
1,4-Benzenediamine
Prednisolone
Probenecid

Display All Chemicals | Download Chemical Data

Select Output Format

Excel

Customize Results

- [ ] Select All

Chemical Identifiers

- [x] Chemical Name
- [x] DTXSID
- [ ] CAS-RN
- [ ] InChIKey
- [ ] IUPAC Name

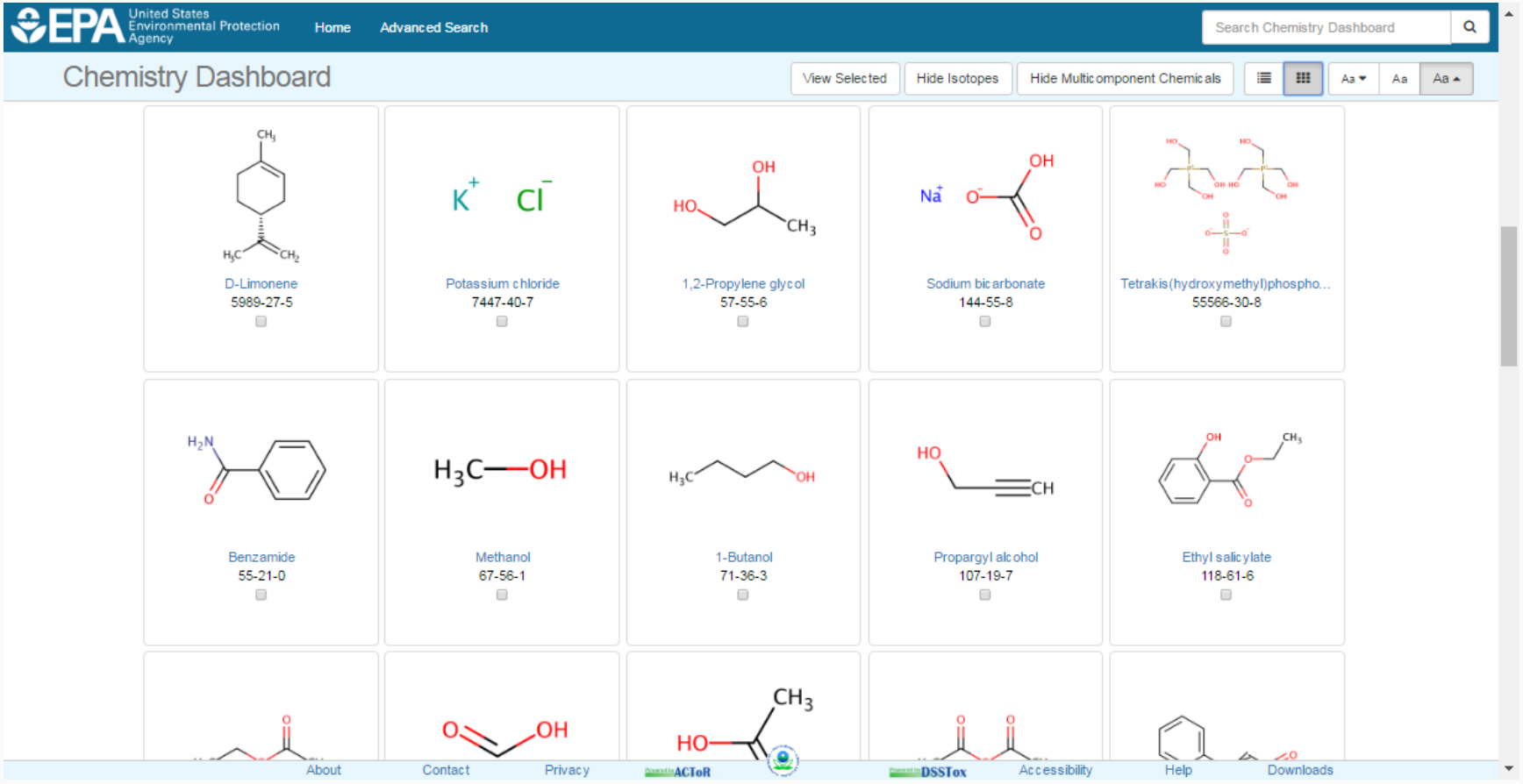
Structures

- [ ] Mol File
- [ ] SMILES
- [ ] InChI String

Intrinsic Properties

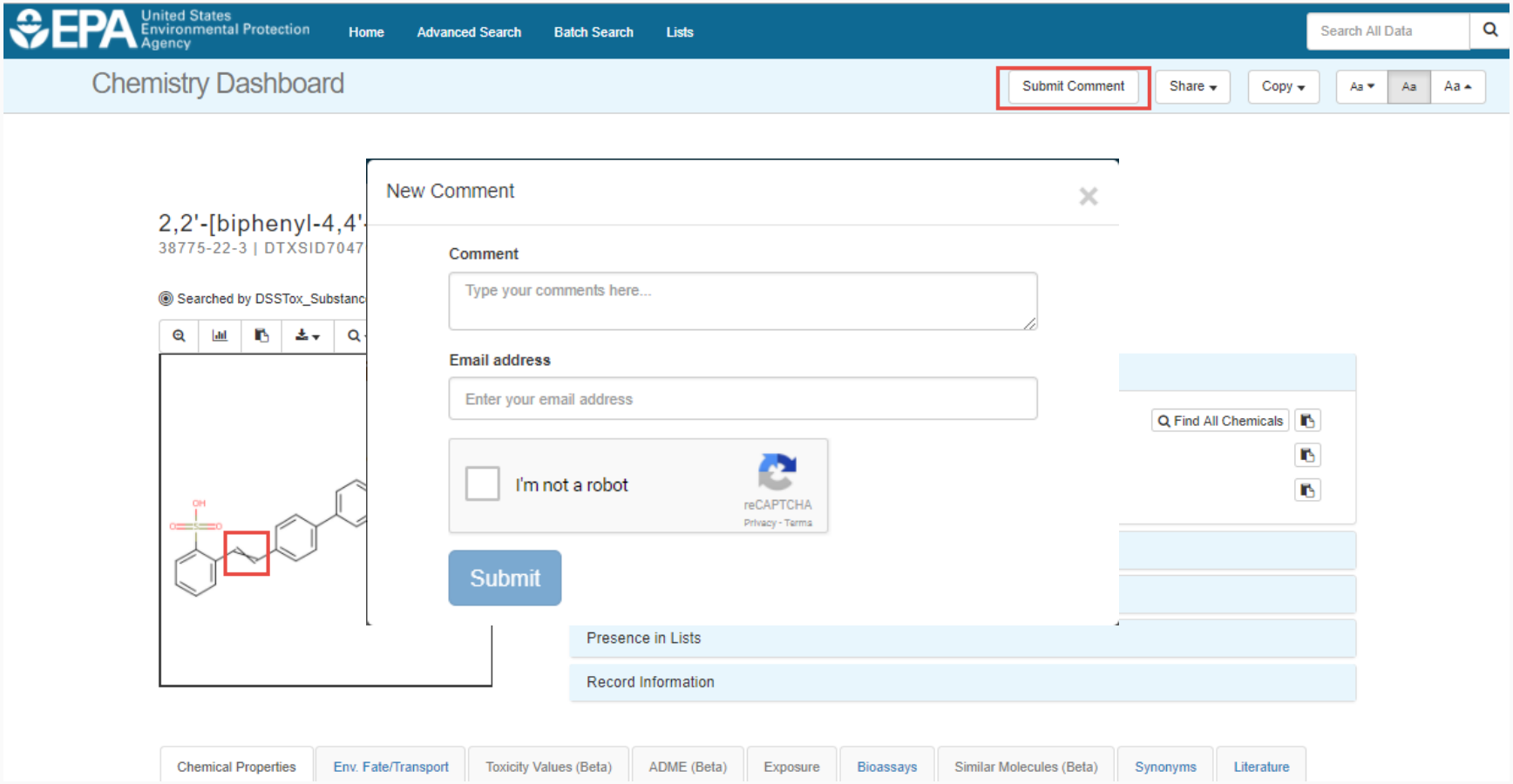
- [ ] Molecular Formula
- [ ] Average Mass
- [ ] Monoisotopic Mass
- [ ] OPERA and TEST Model Predictions

Access to associated data for review, modeling & download



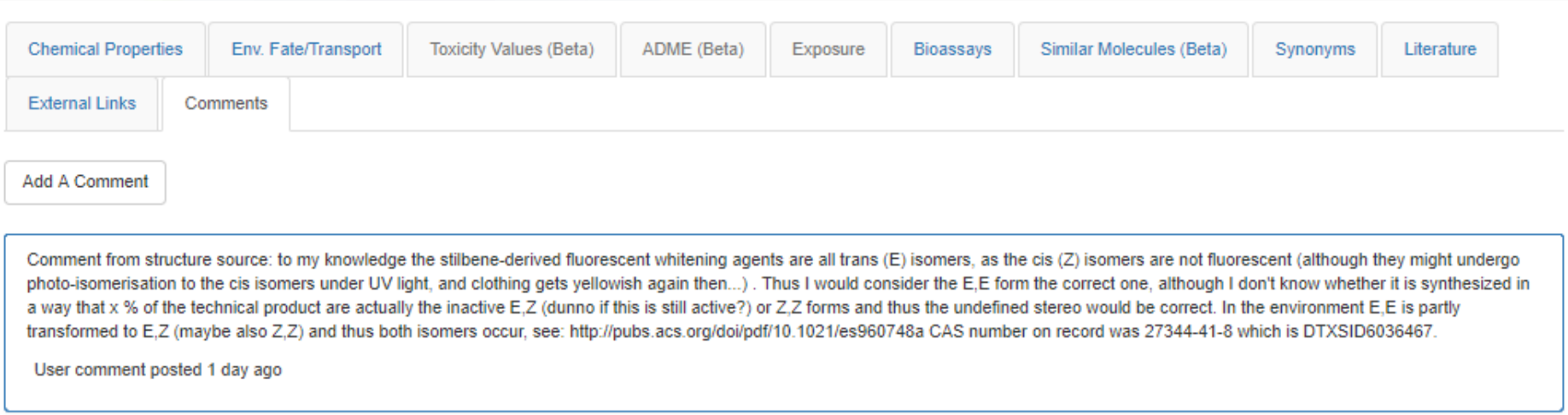
Crowdsourced Curation – HELP!

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



Crowdsourced Curation – HELP!

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



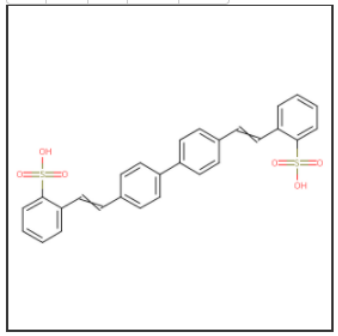
- The Question – is this a mixture of E/Z, E/E or Z/Z? Brighteners should be E/E…

- **Thanks for helping us Emma!**

Curation is laborious work

38775-22-3 | DTXSID7047017

Searched by DSSTox_Substance_Id: Found 1 result for 'DTXSID7047017'.



- So, it's clear what the correct E/Z orientation is yes???

CAS Registry Number 38775-22-3

~160 ~17

$C_{28}H_{22}O_6S_2$

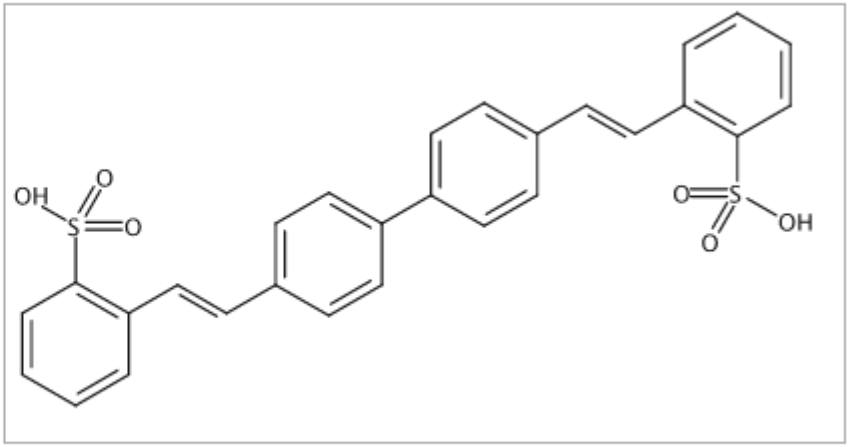
Benzenesulfonic acid, 2,2'-([1,1'-biphenyl]-4,4'-diyldi-2,1-ethenediyl) bis-

Molecular Weight
518.60

Density (Predicted)
Value: 1.414±0.06 g/cm3 | Condition: Temp: 20 °C Press: 760 Torr

pKa (Predicted)
Value: -0.92±0.27 | Condition: Most Acidic Temp: 25 °C

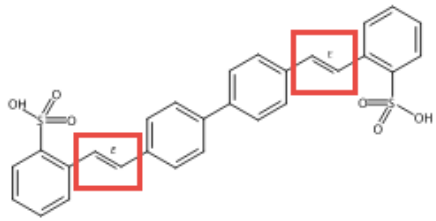
Other Names
2,2'-([1,1'-Biphenyl]-4,4'-diyldi-2,1-ethenediyl)bis[benzenesulfonic acid]
4,4'-Bis(2-sulfostyryl)biphenyl
4,4'-Bis[2-(2-sulfophenyl)ethenyl]biphenyl
C.I. 482200
C.I. Fluorescent Brightener 351



Curation is laborious work

1. **334756-45-5**

~5 ~1



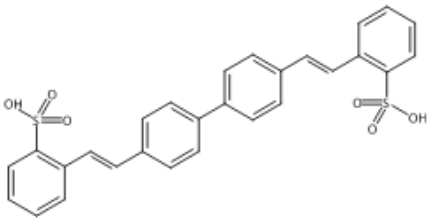
Double bond geometry as shown.

$C_{28}H_{22}O_6S_2$
Benzenesulfonic acid, 2,2'-[(1*E*)-[1,1'-biphenyl]-4,4'-diyldi-2,1-ethenediyl]bis- (9CI)

▸ **Key Physical Properties**

2. **38775-22-3**

~160 ~17



$C_{28}H_{22}O_6S_2$
Benzenesulfonic acid, 2,2'-([1,1'-biphenyl]-4,4'-diyldi-2,1-ethenediyl)bis-

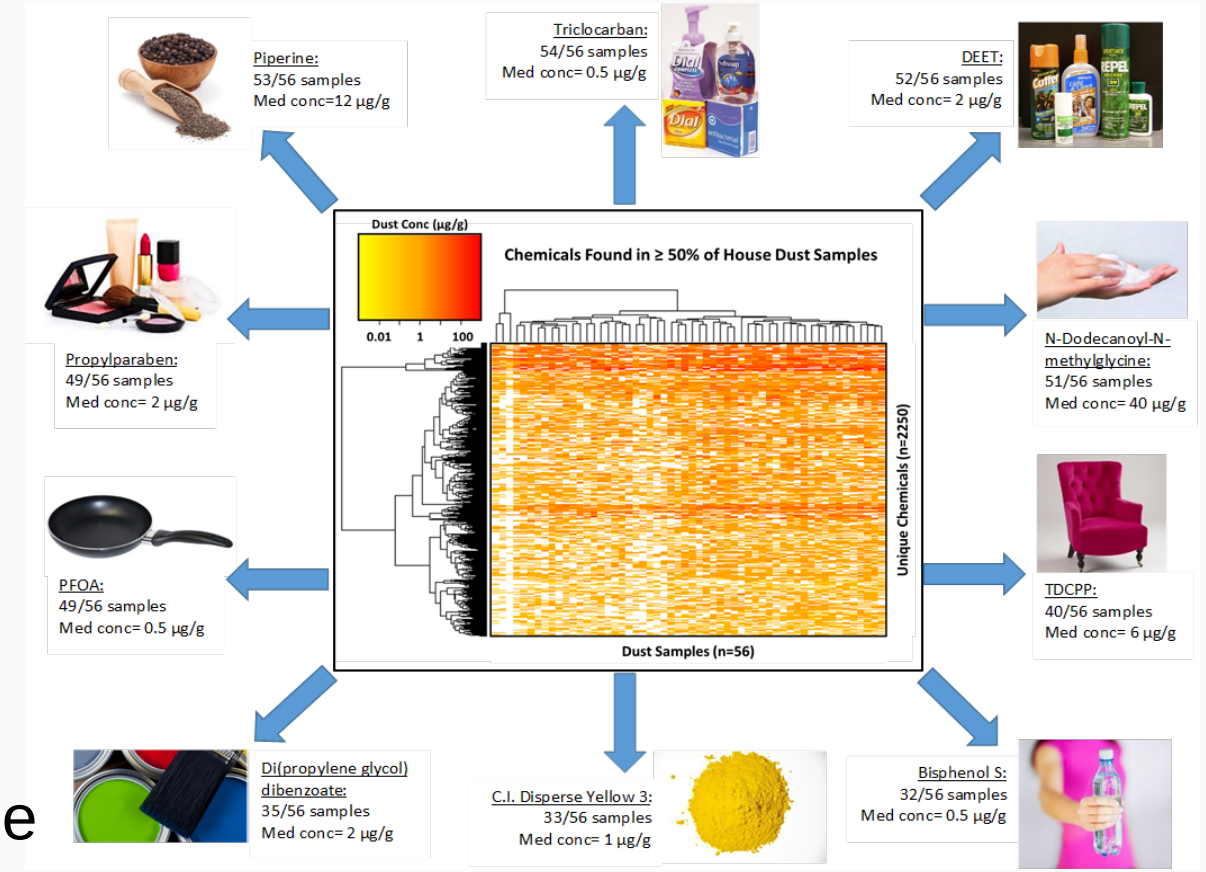
▸ **Key Physical Properties**
Regulatory Information

Answering Questions

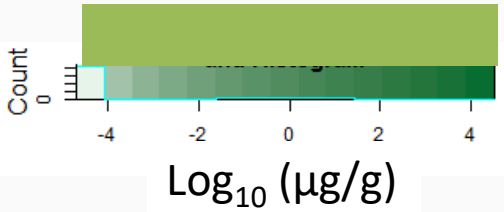
- I have a 1000 CAS Numbers (or Names) – is there data available?
  - Has any Toxcast data been run?
  - Are there Toxicity Data values available?
  - Are there predicted exposure data (via Expocast)?
  - Can I get predicted physchem data for my model?
- Identifying chemicals by analytical chemistry

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

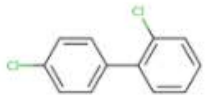
Chemical “Families”

- 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?

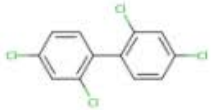
Chemical “Families”

EPA United States Environmental Protection Agency

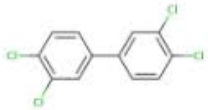
Download as: TSV | Excel | SDF



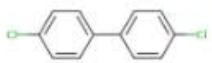
2,4'-Dichlorobiphenyl
34883-43-7



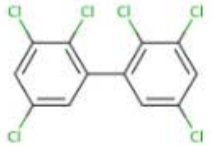
2,2',4,4'-Tetrachlorobiphenyl
2437-79-8



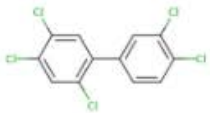
3,3',4,4'-Tetrachlorobinphenyl
32598-13-3



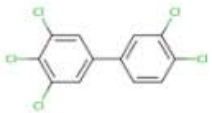
4,4'-Dichlorobiphenyl
2050-68-2



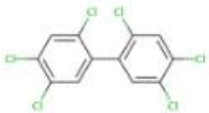
2,2',3,3',5,5'-Hexachlorobiphenyl
35694-04-3



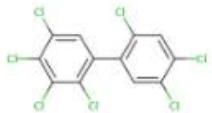
2,3',4,4',5-Pentachlorobiphenyl
31508-00-6



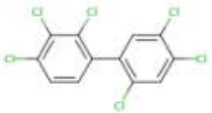
3,3',4,4',5-Pentachlorobiphenyl
57465-28-8



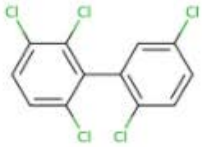
2,2',4,4',5,5'-Hexachlorobiphenyl
35065-27-1



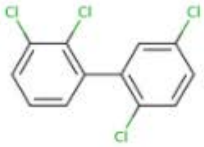
2,2',3,4,4',5,5'-Heptachlorobiphenyl
35065-29-3



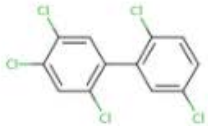
2,2',3,4,4',5'-Hexachlorobiphenyl
35065-28-2



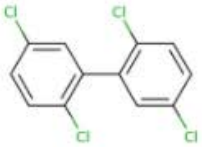
2,2',3,5',6-Pentachlorobiphenyl
38379-99-6



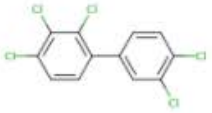
2,2',3,5'-tetrachlorobiphenyl
41464-39-5



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

How Did We Do This? DSSTox

EPA United States Environmental Protection Agency

ACToR-DSSTox Chemical Registration

View/Edit a Single Record | Structure Search | Browse/Curate Records | Export DSSTox | Chemotypes | Manage Chemical Lists | Manage Property Data | Add Deleted Casrns | Welcome, antony | Logout

Substance_ID: DTXSID5024267
CAS: 1336-36-3
Name: Polychlorinated biphenyls
Substance Type: Mixture/Formulation
QC Level: DSSTox_High
Data Source: STN(DSSTox)
QC Notes: biphenyl with multiple (unknown number) chlorines attached at unknown locations

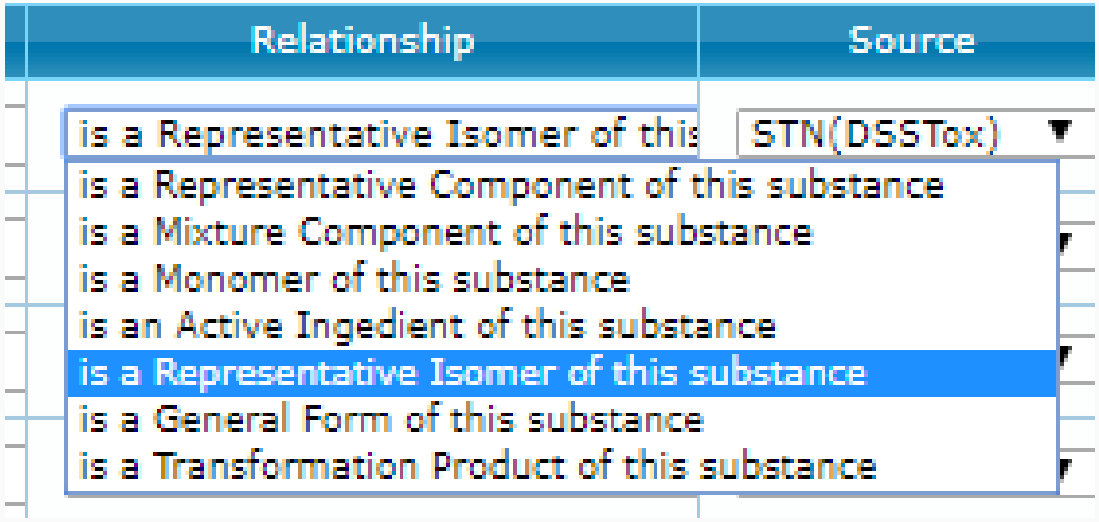
Compound_ID:
Chemical Shown: No Structure
Private Notes:
Source of CAS-Compound:
Double Stereo:
Chiral Stereo:
Chemical Form:

▸ Synonyms (31)
▸ Other Cas (0)
▾ Successor Substances (209)

	CAS-RN	Relationship	Source	Struc	Casrn	Comments
	32774-16-6	is a Representative Isomer of this	STN(DSSTox)	✓		structure shown 3,3',4,4',5,5
	2051-60-7	is a Representative Isomer of this	Public			
	2051-61-8	is a Representative Isomer of this	Public			
	2051-62-9	is a Representative Isomer of this	Public			
	13029-08-8	is a Representative Isomer of this	Public			
	16605-91-7	is a Representative Isomer of this	Public			
	25569-80-6	is a Representative Isomer of this	Public			
	33284-50-3	is a Representative Isomer of this	Public			

Relationship Mappings

- Various relationship mappings can be established. To this point all are manual.



- In progress – **metabolite** mappings (building metabolism competence into high-throughput assasy)

UVCB Chemicals

- UVCB chemical examples
  - Surfactants with undefined composition
  - Petroleum Distillates
  - Gelatins, hydrozylates
  - Formaldehyde, reaction products with diethanolamine
  - Fatty acids, linseed-oil, compds. with triethylamine

Managing UVCB Relationships

EPA United States Environmental Protection Agency

Alkylbenzenesulfonate, linear

42615-29-2 | DTXSID3020041

Searched by Synonym: Found 1 result for 'Linear alkylbenzene sulfonate'.

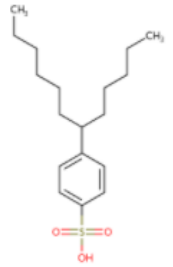
Presence in Lists

Surfactant List Screened in Swiss Wastewater (2014)

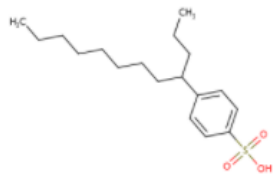
Surfactant List Screened in Swiss Wastewater (2014)

EAWAGSURF is a list of surfactants screened in Swiss wastewater effluents as part of a 2014 study. Structures/mixtures are being progressively curated and linked (Schymanski/Williams). Further details in Schymanski et al 2014, DOI: 10.1021/es4044374

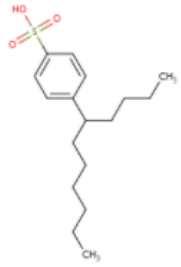
cals



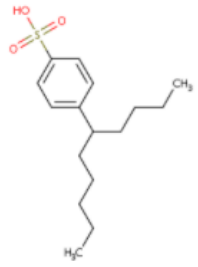
4-(Dodecan-6-yl)benzene-1-sulfon...
23003-92-1



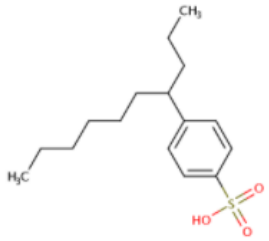
4-(dodecan-4-yl)benzene-1-sulfoni...
NOCAS_862870



C11-LAS
NOCAS_881097



4-(decan-5-yl)benzene-1-sulfonic ...
NOCAS_881146



4-(decan-4-yl)benzenesulfonic acid
NOCAS_891333

Deliver Data and Models for Reuse

EPA United States Environmental Protection Agency

Toxicity ForeCaster (ToxCast™) Data

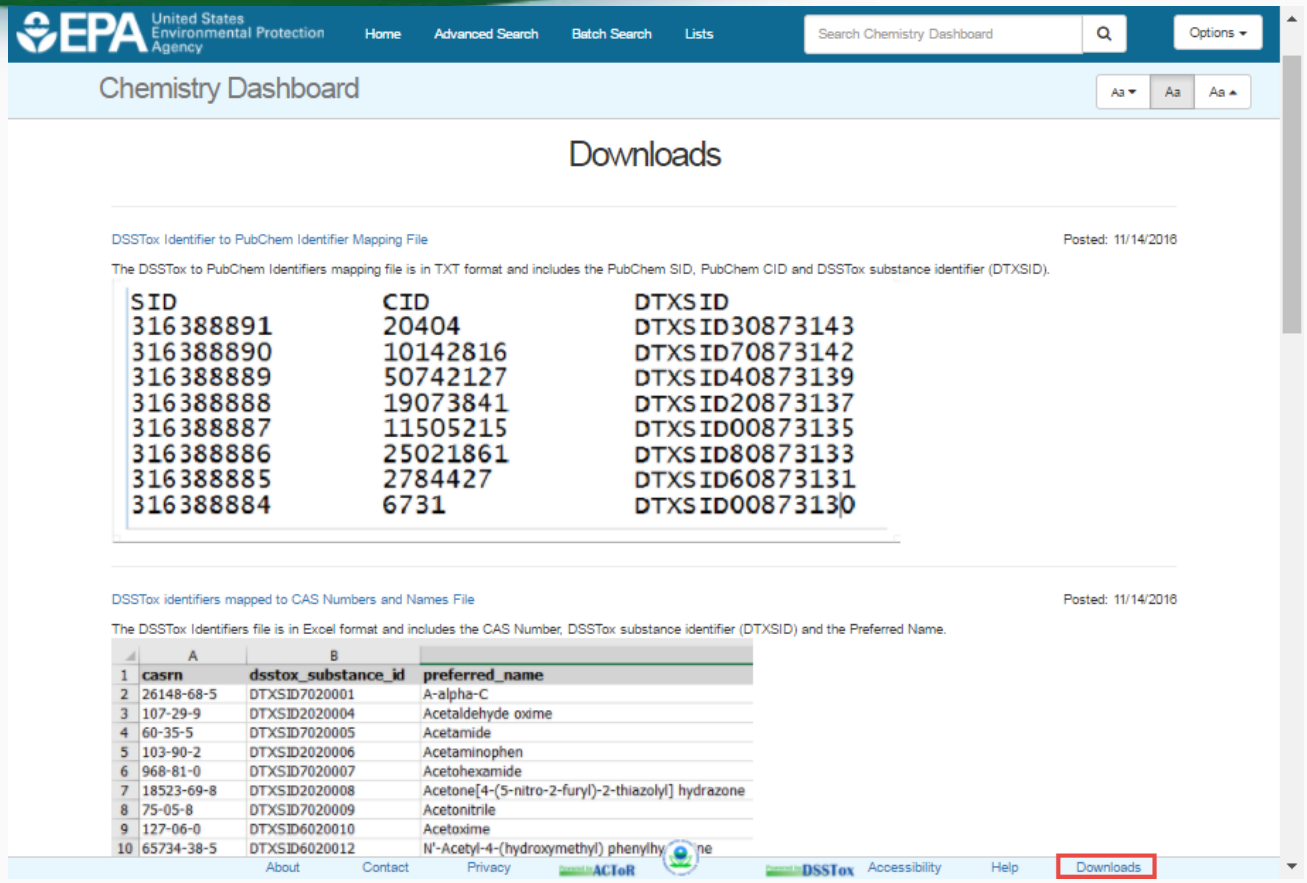
EPA's most updated, publicly available high-throughput toxicity data on thousands of chemicals. This data is generated through the EPA's ToxCast resesarch effort. ToxCast is part of the Toxicology in the 21st Century (Tox21) federal collaboration. All data is available for download and includes the following data sets. The release date and version names for the data sets are provided in the table below.

As part of EPA's commitment to share data, all of the computational toxicology data is publicly available for anyone to access and use. EPA's computational toxicology data is considered "open data", and thus all of the data below are free of all copyright restrictions, and fully and freely available for both non-commercial and commercial use.

- **ToxCast & Tox21 Chemicals**: A list of all chemicals screened, along with descriptions.
- **ToxCast & Tox21 High-throughput Assays**: Information about the hundreds of assays used to screen the chemicals.
- **ToxCast & Tox21 Summary Data**: Summary hit calls for all chemicals and concentration response curves for all assays.
- **MySQL Database**: Downloadable database that provides access to EPA's analysis of the ToxCast and Tox21 high-throughput screening data.
- **R Package**: The computer programming package used to process and model all EPA ToxCast and Tox21 chemical screening data.
- **Data Spreadsheet:** Spreadsheet that provides access to EPA's analysis of the ToxCast and Tox21 high-throughput screening data.
- **Concentration Response Curves:** Concentration response curves for all ToxCast & Tox21 assays.
- **Collaborative Estrogen Receptor Activity Prediction Project Data**: Data and supplemental files from CERAPP, a large-scale modeling project.

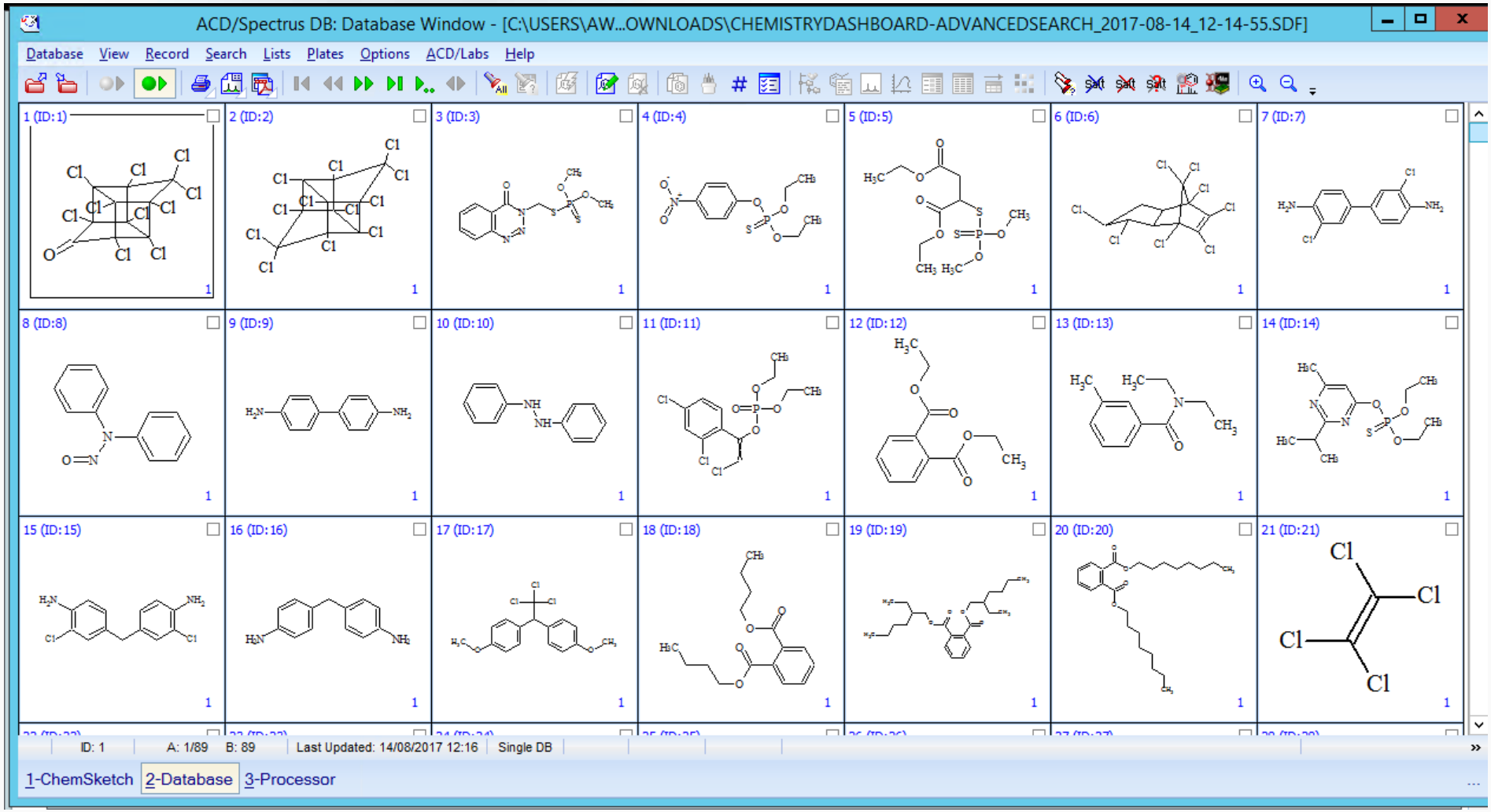
https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data

Delivering our Chemistry Data

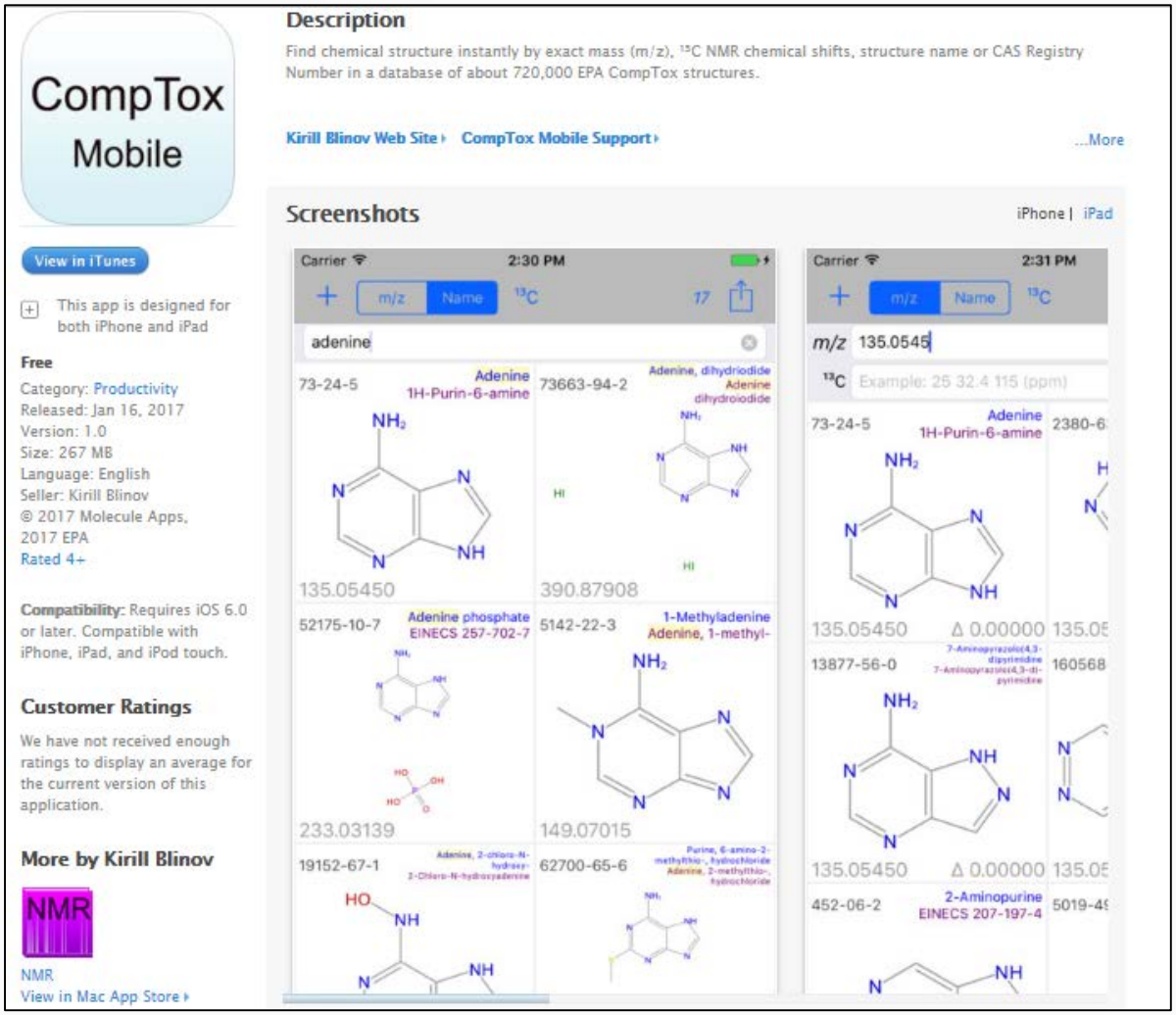


- Various types of data at FTP download site: ftp://newftp.epa.gov/COMPTOX/Sustainable_Chemistry_Data/Chemistry_Dashboard

SDF Download



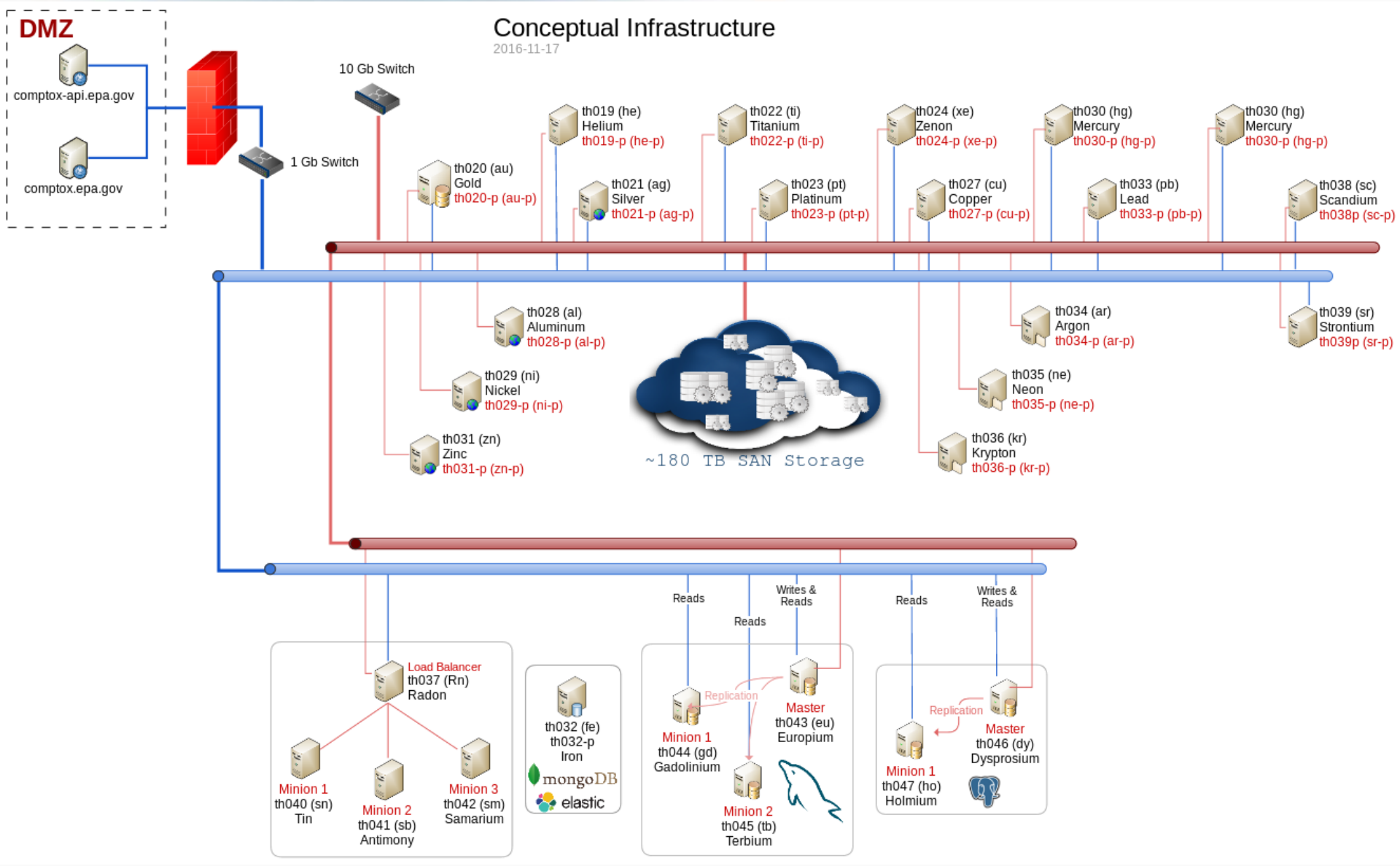
Open Data Reuse on iOS



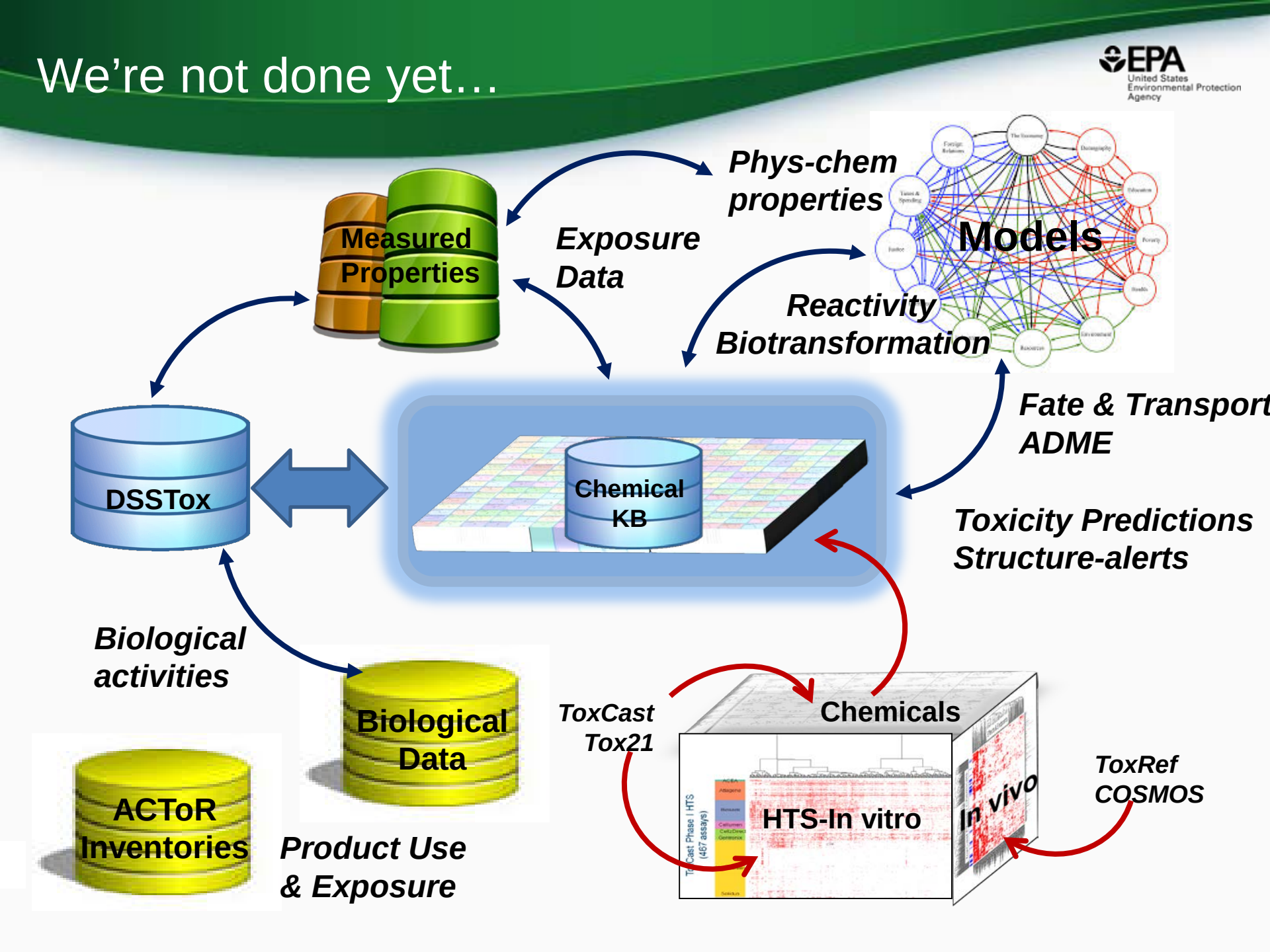
Computational Infrastructure and Processes

- 10 years of development – heterogeneous
  - MySQL, Java, R-code, Matlab, multiple Javascript environments, 100s of web services
  - Multiple projects, loose connectivity
- Migration to homogeneity for production Apps
  - MySQL
  - Java
  - JQuery/JavaScript
  - Versioned microservices
- Sprint methodology with Atlassian management tools
  - **Confluence** – knowledge management
  - **Jira** – ticketing/stories (work assignments)
  - **Fisheye** – collaborative peer code review
  - **Bitbucket** – code version control system

Computational Infrastructure



We're not done yet…



OPERA Services in development


```
{
    "meta": {
        "statusCode": 200,
        "status": "success",        // [success | fail | error]
        "success": true,            // [true | false]
        "message": "string",        // A specific informational message...
        "warnings": "string",       // A warning message like: "LEADING ZEROS STRIPPED"
        "selfUrl":  "http://api.zn.epa.gov/dsstox/chemicals",     // The URL for THIS resource.
    },
    "data": [{
        "dsstox_cid": "DTXCID101",
        "inchi": "InChI=1S/C11H9N3/c12-10-6-5-8-7-3-1-2-4-9(7)13-11(8)14-10/h1-6H,(H3,12,13,14)",
        "inchiKey": "FJTNLJLPLJDTRM-UHFFFAOYSA-N",
        "model_name": "OPERA_LogP",
        "model_version": "v1.1.4",
        "predicted_property_id": 45433098
        "predicted_value": 2.53385
        "global_ad": 1
        "local_ad": 0.865944
        "confidence_score": 0.618723
    }],
}
```

T.E.S.T services (ALPHA)

https://comptox.epa.gov/dashboard/web-test/WS?smiles=ClC(Cl)(Cl)Cl

JSON | Raw Data | Headers

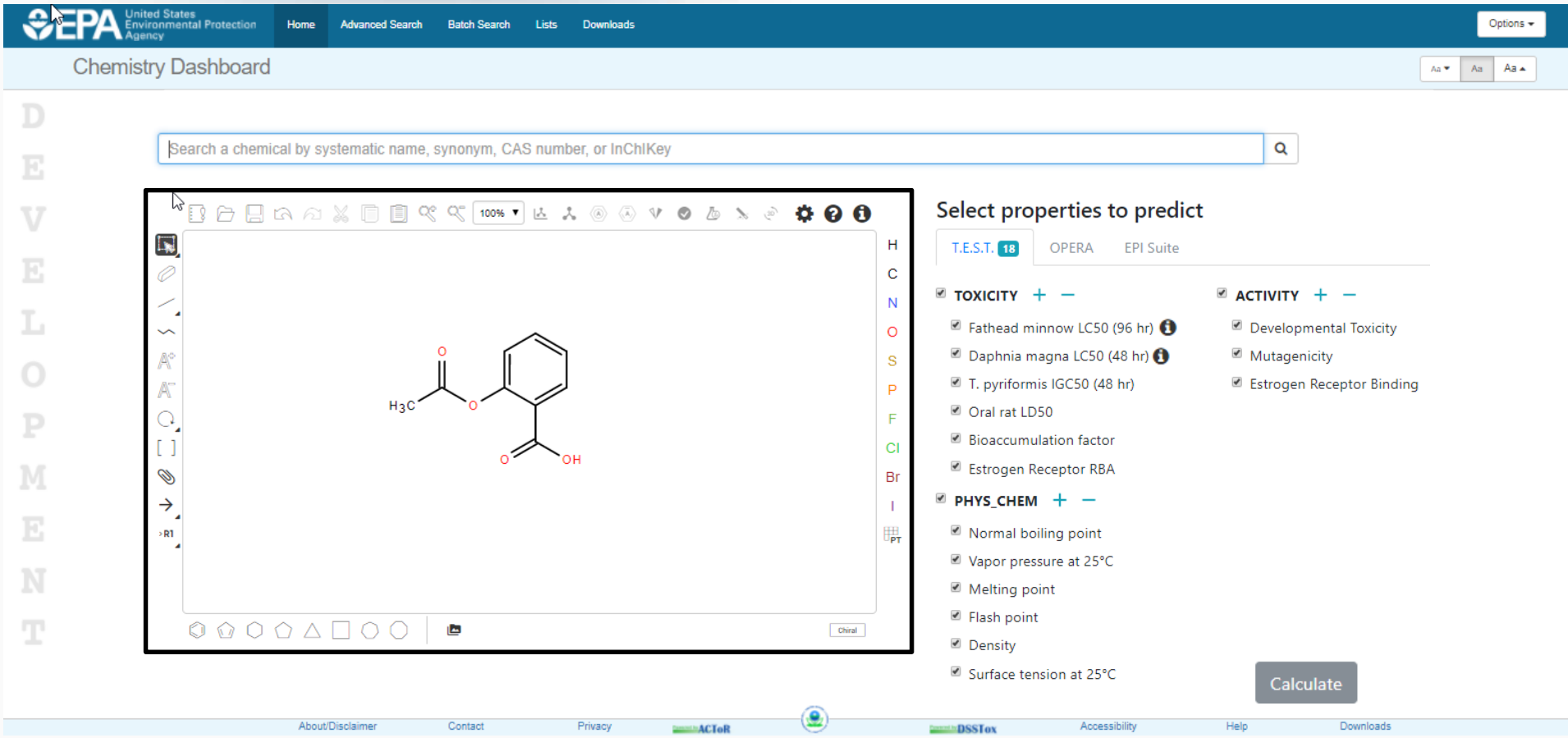
Save | Copy


```
uuid:              "919325ac-14e5-45b5-b932-e401b79edf8c"
predictionTime:    1502741333903
software:          "T.E.S.T (Toxicity Estimation Software Tool)"
softwareVersion:   "4.2"
condition:         "25°C"
endpoint:          "Water solubility at 25°C"
method:            "Consensus"
predictions:
  0:
    id:            "C_1502741333903"
    smiles:        "ClC(Cl)(Cl)Cl"
    expVal:        "2.288"
    expValMass:    "792.473"
    predVal:       "2.186"
    predValMass:   "1003.411"
```

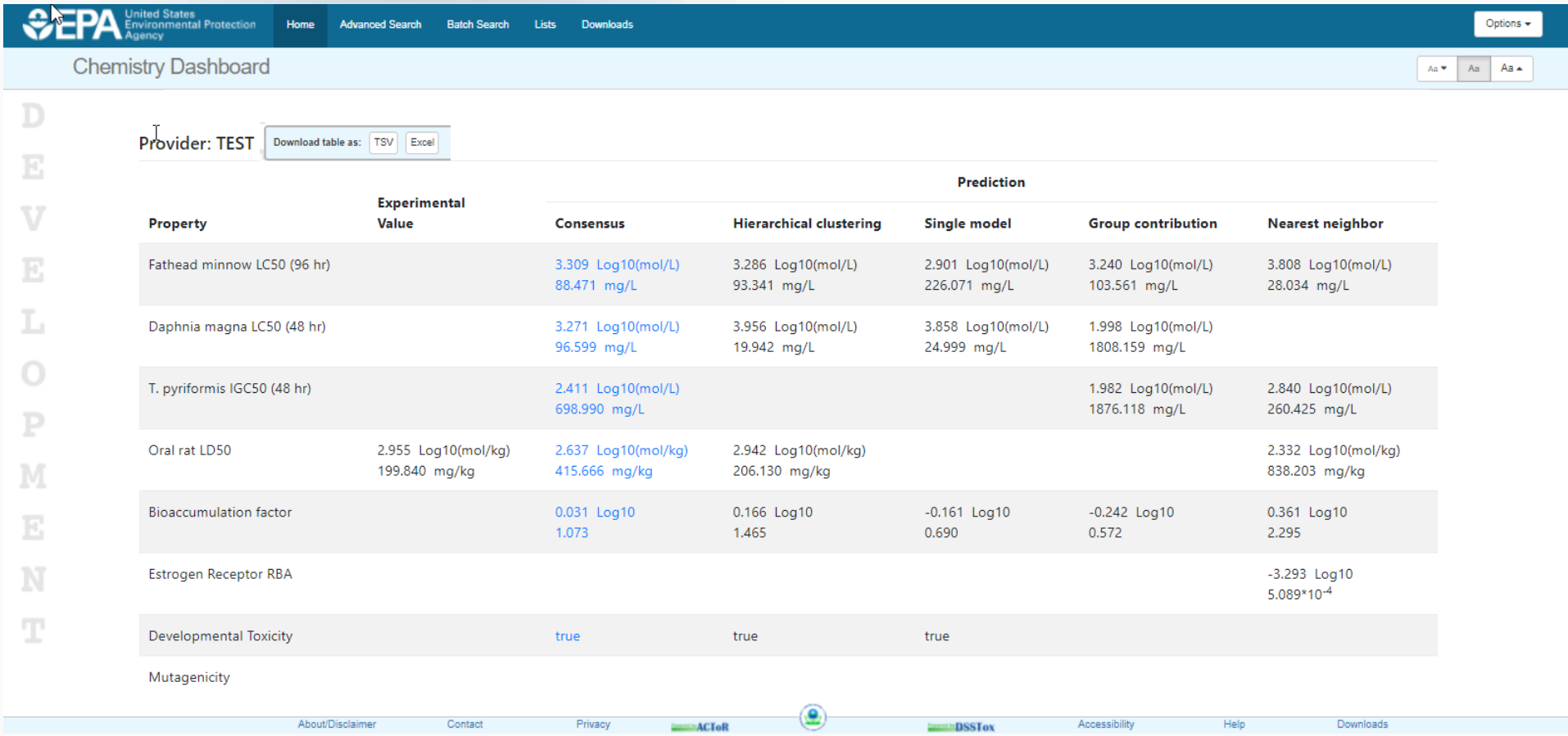
T.E.S.T services (ALPHA)

- 96hr fathead minnow 50% lethal concentration (LC50)
- 48hr daphnia magna 50% lethal concentration (LC50)
- Tetrahymena pyriformis 50% growth inhibition conc. (IGC50)
- Oral rat 50% lethal dose (LD50)
- Bioconcentration Factor (BCF)
- Developmental Toxicity (DevTox)
- Ames Mutagenicity (Mutagenicity)
- Normal boiling point, Flash point, Melting point
- Surface tension, Viscosity, Water Solubility
- Thermal Conductivity, Vapor Pressure, Density

Real time predictions



Real time predictions



Property	Experimental Value	Prediction				
		Consensus	Hierarchical clustering	Single model	Group contribution	Nearest neighbor
Fathead minnow LC50 (96 hr)		3.309 Log10(mol/L) 88.471 mg/L	3.286 Log10(mol/L) 93.341 mg/L	2.901 Log10(mol/L) 226.071 mg/L	3.240 Log10(mol/L) 103.561 mg/L	3.808 Log10(mol/L) 28.034 mg/L
Daphnia magna LC50 (48 hr)		3.271 Log10(mol/L) 96.599 mg/L	3.956 Log10(mol/L) 19.942 mg/L	3.858 Log10(mol/L) 24.999 mg/L	1.998 Log10(mol/L) 1808.159 mg/L	
T. pyriformis IGC50 (48 hr)		2.411 Log10(mol/L) 698.990 mg/L			1.982 Log10(mol/L) 1876.118 mg/L	2.840 Log10(mol/L) 260.425 mg/L
Oral rat LD50	2.955 Log10(mol/kg) 199.840 mg/kg	2.637 Log10(mol/kg) 415.666 mg/kg	2.942 Log10(mol/kg) 206.130 mg/kg			2.332 Log10(mol/kg) 838.203 mg/kg
Bioaccumulation factor		0.031 Log10 1.073	0.166 Log10 1.465	-0.161 Log10 0.690	-0.242 Log10 0.572	0.361 Log10 2.295
Estrogen Receptor RBA						-3.293 Log10 $5.089*10^{-4}$
Developmental Toxicity		true	true	true		
Mutagenicity						

Future Work

- Continuous updating of lists and ongoing curation
- Integration to other agency databases – ECOTOX, ChemView
- Structure and substructure searching in development
- Release of Analytical QC data for ToxCast
- Further Non-Targeted Analysis support – spectral search
- Open API and web services

- "CompTox ~~Chemistry~~ Dashboard" will integrate other dashboards over time (i.e. EDSP21 and ToxCast)

Acknowledgements

EPA-RTP

An enormous team of contributors from NCCT

and collaborators from
NERL
NHERL
NRMRL

Contact

Antony Williams
US EPA Office of Research and Development
National Center for Computational Toxicology (NCCT)
Williams.Antony@epa.gov
ORCID: https://orcid.org/0000-0002-2668-4821