The EPA Comptox Chemistry Dashboard: Two Years in as an Integration Hub for Environmental Toxicology Data

Antony Williams

National Center for Computational Toxicology, U.S. Environmental Protection Agency, RTP, NC

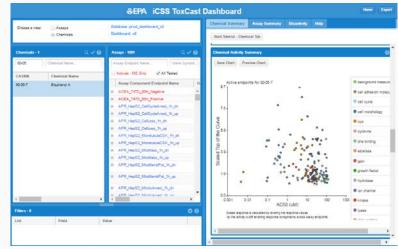
The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA

April 30th 2018 NIEHS

Early Dashboard Applications

United States Environmental Protection





EPA United States	Environmental Protect	NA AGAINST			
Cat: Chemical and a are here: D'A.Homa - C		gories cology Research - Chemica	l Use		EEContact Us
Pitome Pisarch	* Results	5 Dictionary & Downlo	ad Allelp		
Chemical: BISPHENOL	A				
Hyd Control of the Data					
Jae Information:					
CPCat Descrip	ption o	Source Description o	ACTOR Data SetUat ::	Scarce ¢	Class of Chemical Category :
consumer_use_ACToRt	JseOB	Consumer Use		ACTOR UseDB	Use Categories
personal_care_ACToRU	lse08	Personal Care Product		ACTOR UseDB	Use Categories
industrial_manufacturing	ACTORUSeD6	Chemical Industrial		ACTOR UseDB	Use Categories
child_use detected		Consumer Products	The Danish EPA Exposure of 2-year-olds to chemical substances in Consumer Products. This project included a survey of the products as well as chemical analyses and rail, assessments of a number of selected products that 2 year-old children come into context with	ACToR Data Sets and Lists	Use Categories

DODTHENTP BOLUML DODTHE MID DUDY DUMINANES

\$	EDSP21 Dashboard Endecrine Disruption Screening Program for the 21st Century	
Denial Surmary Public	ematos Disartivity Summary Disartivity High-Throughput Exposure Assay Definitions Terminity	
EDSP Dashboard Overview		
EDSP Dashboard Overvie		
Congress requires EPA's Ended Screening Program for the 21st	Edisautize Streaming Program to evaluate chemicals for potential endocrine disruption, and there are thousands of chemicals of interest to the program. El ntury Dearboard (EDG P21 Dearboard) to provide access to new chemical data on over 1,800 chemicals of interest.	PA researchers developed the Endocrine Disruption
The purpose of the EDSP21 Da	coard is to help the Endoorne Disruptor Screening Program evaluate chemicals for endoorine-related activity.	
The data for this version of the 0	Ploand nomes from various sources -	
 Chemical exposure data 	http://bioacyclu/demoid surrends/data.generated by the ERA's Tourishy Foresistie (TouCard) project and the federal Tourishy Testing in the 25st certury prediction modes: Expo2exDED). uses and emotiones (DES Tris).	/ (Tar21) sollaboration.
ToxCast Data Use Consid-	ations	
Careful review is required	a specific assay does not necessarily meen that it will asses taxionly or an adverse health autome. There are many factors that betwrine whether a the determine the use of the data in a particular declarat correct. In an expected to hange over time about the solance and analytical methods improve.	mical wil cause a specific adverse health outcome.
EPA will continuously add functi	aity and reprove overall usability and performance.	
To get the best possible experies	using the EDSP Dashboard application we recommend using MacEa Foelow or Google Onome.	
0		



- Chemistry data mashed together based on CAS Number/Names
- Chemistry data quality issues
- Multiple applications requiring maintenance

 April 2016 – beta release of the CompTox Chemistry Dashboard as an integration hub

The CompTox Chemistry Dashboard

A publicly accessible website delivering access:

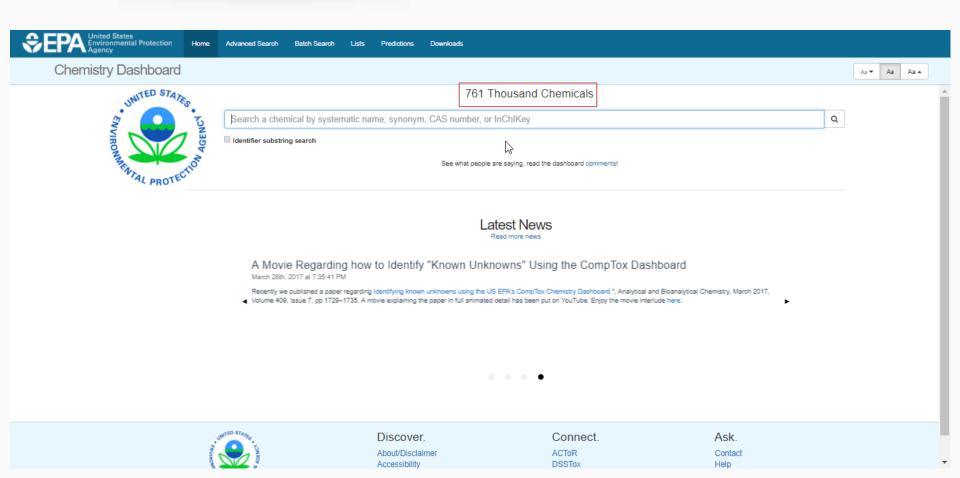
- ~760,000 chemicals with related property data
- Experimental and predicted physicochemical property data
- Integration to "biological assay data" for 1000s of chemicals
- Information regarding consumer products containing chemicals
- Links to other agency websites and public data resources
- "Literature" searches for chemicals using public resources
- "Batch searching" for thousands of chemicals
- DOWNLOADABLE Open Data for reuse and repurposing

mental Protection

CompTox Chemistry Dashboard

SEPA United States Environmental Protection Agency

https://comptox.epa.gov/dashboard



Detailed Chemical Pages



SEPA United States Environmental Protection Home Advanced S Agency	arch Batch Se	arch Lists	Predictions	Downloads				s	earch All Data	Q
Chemistry Dashboard EPAHFR						Submit Comm	ent Copy	·• A	a▼ Aa	Aa 🔺
1,2-Propylene glyco 57-55-6 DTXSID0021206					2					
Searched by DSSTox_Substance_Id: Found 1 resul	for 'DTXSID002120	6'.								
	Wikiped	ia								
ОН	C3H8O2 is classe	It is a viscous of as a diol and it	colorless liquid wh s miscible with a b	ich is nearly odd road range of so	hetic organic compou orless but possesses olvents, including wat uction of polymers, b	a faintly sweet tas ter, acetone, and c	ste. Chemically i chloroform. It is	t		
HO HO	Intrinsic	Properties								
СН3	Structur	al Identifiers								
	Linked	Substances								
	Present	e in Lists								
	Record	Information								
	Quality	Control Notes								
Chemical Properties Env. Fate/Transport Hazard ADME (Be	a) Exposure	Bioassays	Similar Con	pounds R	elated Substances	Synonyms	Literature	Links	Comment	ts

Presence in Lists

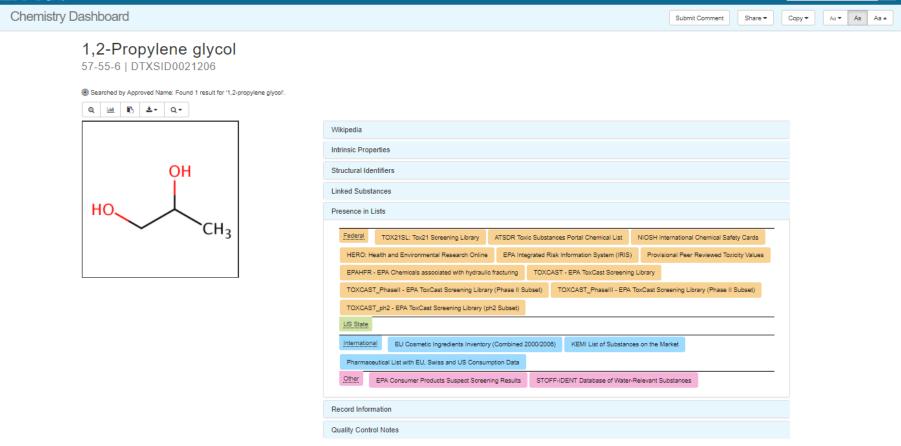


Search All Data

Q

Sepan United States Environmental Protection Agency

Home Advanced Search Batch Search Lists Predictions



Downloads

Executive Summary (Beta)	Chemical Properties Env. Fate/Transport	Hazard	ADME (Beta)	Exposure	Bioassays	Similar Compounds	Related Substances	Synonyms	Literature	Links	Comments
--------------------------	-----------------------------------------	--------	-------------	----------	-----------	-------------------	--------------------	----------	------------	-------	----------

Data Quality



Control Control Protection Home Advanced Search Batch Search Lists Pr	edictions Downloads		Search	All Data	Q
Chemistry Dashboard	s	Submit Comment Share 🕶	Сору 🕶	Aa•	Aa Aa 🔺
1,2-Propylene glycol 57-55-6 DTXSID0021206					
Searched by Approved Name: Found 1 result for '1,2-propylene glycol'. Q Lett C C C					
	Wikipedia				
	Intrinsic Properties				
ОН	Structural Identifiers				
	Linked Substances				
HO	Presence in Lists				
СН3	Record Information				
	Citation U.S. Environmental Protection Agency. Chemistry Dashboard. https://comptox.epa.gov/dashboard/DTXSID0021208 (accessed glycol		в		
	Data Quality				
	Level 1:Expert curated, highest confidence in accuracy and consistency of unique chemical identifiers				
	Level 2:Expert curated, unique chemical identifiers using multiple sources Level 3:Programmatically curated from high quality EPA source, unique chemical identifiers have no conflicts in ChemiD and Pu	ubChem			
	Level 4:Programmatically curated from ChemID, unique chemical identifiers have no conflicts in PubChem				
	Level 5: Programmatically curated from ACTOR or PubChem, unique chemical identifiers with low confidence, single public soun	ice .			
	Quality Control Notes				
Executive Summary (Beta) Chemical Properties Env. Fate/Transport Hazard ADME (Beta)	Exposure Bioassays Similar Compounds Related Substances Synonyms Literature Links	Comments			

Properties, Fate and Transport



Chemistry Da	shboard					Submit	Comment Share -	Copy 🗸 🛛 Aa 🔻 Aa
Summary	Download as: TSV -	Event CDE						
LogP: Octanol-Water	Download as: 15V -	Excert SDF +						
Water Solubility	Property	Ave	erage	Med	ian		Range	Unit
Density		Experimental	Predicted	Experimental	Predicted	Experimental	Predicted	
Donoky	LogP: Octanol-Water	-0.920 (1)	-0.983 (5)	-	-0.966	-0.920	-1.34 to -0.780	-
Flash Point	Water Solubility	13.1 (1)	12.6 (4)	-	11.2	13.1	4.32 to 23.6	mol/L
Melting Point	Density	-	1.02 (2)	-	1.02	-	1.01 to 1.04	g/cm^3
-	Flash Point	-	88.8 (2)	-	88.8	-	70.4 to 107	°C
Boiling Point	Melting Point	-60.0 (6)	-30.6 (4)	-60.0	-29.3	-60.0	-42.4 to -21.6	°C
Surface Tension	Boiling Point	187 (6)	180 (5)	187	185	187 to 188	155 to 200	°C
Thermal Conductivity	Surface Tension	-	35.5 (2)	-	35.5	-	33.1 to 38.0	dyn/cm
Thermal Conductivity	Thermal Conductivity	-	185 (1)	-	-	-	-	mW/(m*K)
Vapor Pressure	Vapor Pressure	1.29e-01 (1)	1.91e-01 (4)	-	2.08e-01	1.29e-01	5.37e-02 to 2.95e-01	mmHg
Viscosity	Viscosity	-	12.6 (1)	-	•	-	-	cP
	LogKoa: Octanol-Air	-	6.74 (1)	-	-	-	-	-
LogKoa: Octanol-Air	Henry's Law	-	6.02e-08 (1)	-	-	-	-	atm-m3/mole
Henry's Law	Index of Refraction	-	1.43 (1)	-	-	-	-	-
	Molar Refractivity	-	19.0 (1)	-	•	-	-	cm^3
Index of Refraction	Molar Volume	-	73.4 (1)	-	•	-	•	cm^3
Molar Refractivity	Polarizability	-	7.52 (1)	-	•	-	-	Å^3

Model Performance Details

(Redesigned display)

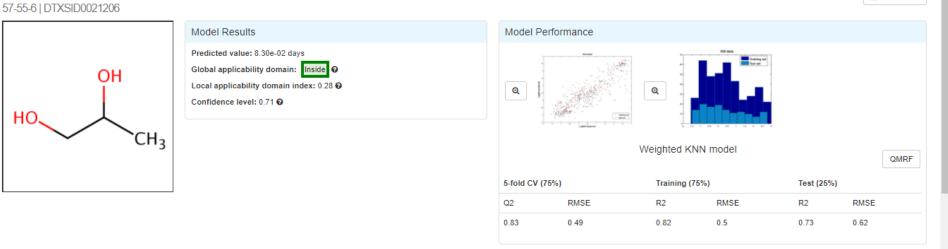
United States Environmental Protection Agency

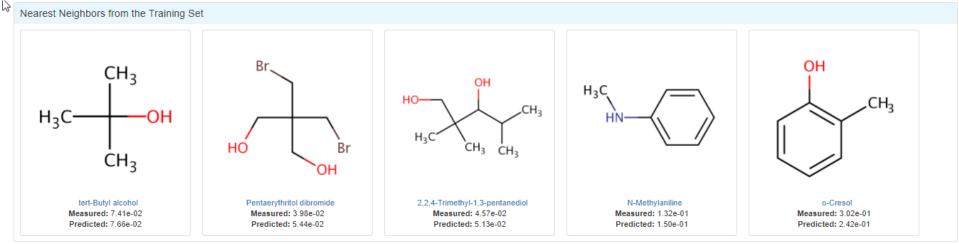
Save PDF



OPERA Models: Fish Biotrans. Half-Life (Km)

1,2-Propylene glycol





OPERA: OPEN Data and OPEN Models



Mansouri et al. J Cheminform (2018) 10:10 https://doi.org/10.1186/s13321-018-0263-1

Journal of Cheminformatics

RESEARCH ARTICLE



Open Access

OPERA models for predicting physicochemical properties and environmental fate endpoints

Kamel Mansouri^{1,2,3*}⁽¹⁾, Chris M. Grulke¹, Richard S. Judson¹ and Antony J. Williams¹

T.E.S.T Models



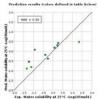


- <u>QSAR Methodologies</u>
- <u>What's New in Version 4.2.1?</u>
- Predictions made for the whole database
- Includes detailed calculation reports

T.E.S.T Calculation Reports



Chemistry Dashboard	March 12			Doveload FDF As *
Predicted Water solubility at 25°C for 57-54 Predicter results	-6 from Consensus method			
Endpoint	Experimental value (CAS+ 57-55-5) Source: EPI Soite v 4.00		Predicted value ⁴	
Viale: solubility at 25°C-Log(D);vall(.)	Scerce: EPI Scree V 4.00		-0.04	
Viater solubility at 25°C mg/L	1001017 82		329661.67	
Write the test chemical was present in the having set. The prediction do individual Productions	es notrepresent an external production.			
	er ochryssent in eithers prodon. Predicted vision 			
Individual Productions	Predicted value			
velviduil Predictions Method	Produced value -inghtfered13 -1.12 -0.01	and the second		
Individual Productions Method Herarchical cluttering	Predicted value .ingitigrad () .4.12	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		
vdvidual Productores Nethod Herarchide Cuthining dirace contribution	Produced value -inghtfered13 -1.12 -0.01			



emicals	MAE'
nåre set	0.58
imilarity coefficient = 0.5	0.39
fean absolute error in -Log10(mol/L)	
Aean aboolute error in -Lag10(molit.)	
lean atoolula emor in -Log10ymolit.)	

CAS	Structure	Similarity Coefficient	Experimental value -Log10(mol/L)	Predicted value -Log10(mol/L)
57-55-6 (test chemical)	И СН		-1.12	-0.64
78-92-2	H ₁ C CH ₂	0.96	-0.39	-0.18
64-17-5	но СН3	0.95	-1.34	-0.91
584-02-1	HSC HO	0.92	0.23	0.30
353-36-6	H ₃ C F	0.91	1.35	0.48
109-86-4	110	0.90	-1.12	-0.82
75_50_5		0.00	_0.26	_n 70

Access to Chemical Hazard Data



EPA United Stat Environme Agency	es ntal Protection	Home Adva	nced Search	Batch Sea	rch Lists	Predicti	ons Do	wnloads						Sea	arch All Data	0
Chemistry D	ashboard E	EPAHFR									Submit Com	nment	Сору	Aa	▼ Aa	Aa 🔺
Chemical Properties En	/. Fate/Transport	lazard AD	ME (Beta)	Exposure	Bioassays	Simila	r Compound	s Rela	ted Substa	noes	Synonyms		Literature	Links	Comments	5
Exposure Limit	Download t	able as: TSV	Excel				Human	Eco								
Lethality Effect Level Point of Departure	•	r .	÷	÷	Risk	.+	÷	. C	fuela i	Experie	-			÷	÷	, ×
Toxicity Value																^
	•	Downloa	reference	Excel	referenc	e										
	+	Study ID	Value S	screenign refe /alue valu	rence	study	issue	endpoint	POD		benchmark response		uncertainty factor A	uncertainty factor D	factor H	y un fac
	•	1600000	Sub- chronic 9 Reference (Dose (s- RfD)) 20	mg/kg bw-day ADD	Sub- chronic	Hematologic	Erythrocyte count (RBC)	5200 mg/k 5200 day ADD	LOAEL		0	10	1	10	3
	+	1600000	Chronic) 20	mg/kg bw-day ADD	Sub- chronic	Hematologic	Erythrocyte count (RBC)	5200 mg/k 5200 day ADD	LOAEL		0	10	1	10	3
	+	3														×
	+	Close														

In Vitro Bioassay Screening ToxCast and Tox21





14

Detailed Assay Description as PDF (limited number of assays)



NVS_NR_hFXR_Antagonist

Assay Title: NovaScreen Human Farnesoid x Receptor Alpha (FXR) Ligand-Binding Antagonist Screening Assay

- 1. Assay Descriptions
- 1.1. Overview

Assay Summary:

High-throughput screening of in vitro chemical-target interactions across a wide variety of compounds through a broad range of biochemical interactions will help describe the chemicalassay bioactivity space for chemicals with limited available information. There exists a large number of environmental chemicals for which there is little information about the potential for bioactivity. The NVS NR human farnesoid x receptor (FXR, NR1H4) agonist assay format allows for an efficient screening of thousands of chemicals for the ability to competitively bind to the ligandbinding domain of a xenobiotic sensing nuclear receptor. This assay was developed to screen the ToxCast chemical library for potential farnesoid x receptor ligand-binding activity using a TR-FRET competitive displacement assay and a known FXR receptor agonist (Chenodeoxycholic Acid, CDCA) as a reference compound. Biochemical high-throughput screening offers preliminary evidence for chemical targets in a cell or tissues which, when combined with information from literature or targeted in vivo studies, can indicate potential pathways for toxicity. This assay was run for a test duration of 1 hour in a 384-well plate.

1.2. Assay Definition

Assay Throughput:

Human FXR ligand-binding domain (LBD) incubated in 384-well microtiter plates for 1 hour prior to measuring ligand dependent binding of cofactor to the receptor using TR-FRET.

Sources of Exposure to Chemicals



Product & Use Categories	osure	Bioassays	Similar Compounds	Related Substances	Synonyms	Literature	Links	Commen	nts
Chemical Weight Fraction		Proc	duct & Use Categorie	es (PUCs) 🕄					
Chemical Functional Use		Categorizat	<u>ion type</u>		nber of Unique P	roducts		*	<u>_</u>
Monitoring Data		PUC PUC		208					
		PUC PUC		107					
Exposure Predictions		PUC PUC		101					
Production Volume		PUC		90					
		PUC		89					•

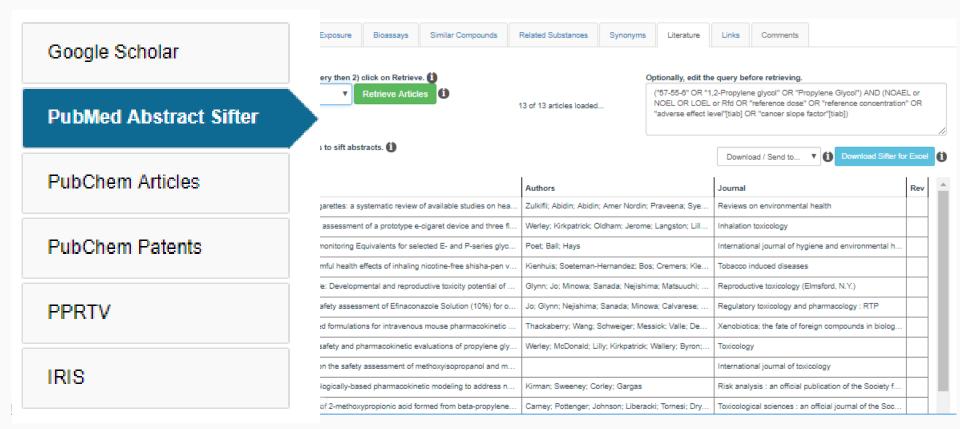
Identifiers to Support Searches



Chemical Properties	Env. Fate/Transport	Hazard	ADME (Be	eta) Exposu	re Bioassays	Similar Compounds	Related Substances	Sy
					Found 78 s	synonyms		
			Legend: V	alid Synonyms	Good Synonyms	Other Synonyms	Copy all Synonyms	
1,2-Propylene glycol								
Propane-1,2-diol								
1,2-Propanediol								
57-55-6 Active CAS-RN								
alpha-Propylene glycol								
(+/-) 1,2-Propanediol								
(RS)-1,2-Propanediol								
dl-Propylene glycol								
3-01-00-02142 Belictein R	egistry Number							
1,2-Propanediol								
(.+)-1,2-Propanediol								
(.+)-Propylene glycol								
1,2-(RS)-Propanediol								
1,2-DIHYDROXYPROPA	NE							
1,2-PROPANDIOL								

Literature Searches and Links





Abstract Sifter for Excel



F1000Research

F1000Research 2017, 6(Chem Inf Sci):2164 Last updated: 06 FEB 2018

Karolinska Institutet, Sweden



SOFTWARE TOOL ARTICLE

Abstract Sifter: a comprehensive front-end system to PubMed

[version 1; referees: 2 approved]

Nancy Baker ¹, Thomas Knudsen², Antony Williams ²

¹Leidos, Research Triangle Park, NC, USA

²National Center for Computational Toxicology, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

First published: 21 Dec 2017, 6(Chem Inf Sci):2164 (doi: **Open Peer Review** 10.12688/f1000research.12865.1) Latest published: 21 Dec 2017, 6(Chem Inf Sci):2164 (doi: Referee Status: 🗹 🗸 10.12688/f1000research.12865.1) Abstract Invited Referees The Abstract Sifter is a Microsoft Excel based application that enhances 1 2 existing search capabilities of PubMed. The Abstract Sifter assists researchers to search effectively, triage results, and keep track of articles of interest. The ~ version 1 tool implements an innovative "sifter" functionality for relevance ranking, giving published report report the researcher a way to find articles of interest guickly. The tool also gives 21 Dec 2017 researchers a view of the literature landscape for a set of entities such as chemicals or genes. The Abstract Sifter is available as a Microsoft Excel Pauliina Damdimopoulou 🕕

macro-enabled workbook application.

Embedded Toxics Release Inventory

CONTRACTOR United Sta



Chemistry Da	1			- 1												
		ċ				Presence in	n Lists									
		C				Record Info	rmation									
						Quality Cor	itrol Notes									
						Quality Col										
Chemical Properties E	Env. Fate/Transport	Hazard	ADME (Beta)	Exposure	Bioassays	Similar Compounds	Related Substances	Synonyms Lit	Literature Links	Comments						
Product & Use Categories	es						т.	uias Delessa								
							IC	xics Release	se inventory						Prin	t Page
Chemical Weight Fraction Chemical Functional Use Toxics Release Inventor	e		Data Source: 20 The Toxics Releas	16 Dataset (r e Inventory (TF	eleased March 20	agement of certain tox	c chemicals that may pose a t						1 of each chemica	al is		
Chemical Functional Use	e		Data Source: 20 The Toxics Releas recycled, combust	16 Dataset (r e Inventory (Tf ed for energy r	eleased March 20	018) agement of certain toxi or destruction, and disp		on- and off-site. This		vely referred to as production			h of each chemic	al is		
Chemical Functional Use Toxics Release Inventor Monitoring Data	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (Tf ed for energy r	eleased March 20 RI) tracks the mana ecovery, treated fo	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a t	on- and off-site. This	is information is collecti	vely referred to as production		naged.	n of each chemic	al is		Î
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (TF ed for energy r ities Reportin	eleased March 20 RI) tracks the mana ecovery, treated fo	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a t	on- and off-site. This Qu	is information is collecti	vely referred to as productio	on-related waste ma	naged.	n of each chemic	al is		Î
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (TF ed for energy r ities Reportin	eleased March 24 RI) tracks the mana ecovery, treated fo ng CHLOROPICRI	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a t	on- and off-site. This Qu N Fi Tr R	is information is collecti Quick Facts for 20 Number of TRI Facilities:	vely referred to as production	on-related waste ma	naged. tes	h of each chemic	al is		
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (TF ed for energy r ities Reportin	eleased March 24 RI) tracks the mana ecovery, treated fo ng CHLOROPICRI	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a t	on- and off-site. This Qu N Fi T T Q Q	is information is collecti Quick Facts for 20 Number of TRI Facilities: Total Production. Related Waste Managed:	Vely referred to as production IS Chemical 10 - 96.3 thousand lbs 4 6.3 thousand lbs	United Star 22,130	tes Ibs	h of each chemic	il is		
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (TF ed for energy r ities Reportin	eleased March 2/ (1) tracks the mana eccovery, treated for ng CHLOROPICAL CHLOROPICAL NITED	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a t	on- and off-site. This Qu F, F, T T R M M J T Q Q R	s information is collecti quick Facts for 20 Number of TRI Facilities: Total Production Related Waste Managed: Total On-site and Off-site Disposal or Other Releases; Total On-site:	Chemical Description Official Official	United Sta 22,130 27.1 billion l 3.4 billion lb 2.9 billion lb	tes Ibs Is	h of each chemic	al is		
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (TF ed for energy r ities Reportin	eleased March 2/ (1) tracks the mana eccovery, treated for ng CHLOROPICAL CHLOROPICAL NITED	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a to seed of or otherwise released	on- and off-site. This Qu F, F, T T R M M J T Q Q R	s information is collecti quick Facts for 2C Number of TRI Facilities: Total Production Related Waste Managed: Total On-site Releases: Total On-site: • <u>Air</u> :	Plant Part of the second	United Sta 22,130 27.1 billion I 3.4 billion Ib 686.4 million	tes Ibs vs n Ibs	h of each chemic	al is		
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r e Inventory (TF ed for energy r ities Reportin	eleased March 2/ (1) tracks the mana eccovery, treated for ng CHLOROPICAL CHLOROPICAL NITED	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a t	on- and off-site. This Qu F, F, T T R M M J T Q Q R	s information is collecti quick Facts for 2C Number of TRI facilities: Total Production Related Waste Managed: Total On-site an Off-site Disposal of Other Releases; Total On-site: • <u>Air</u> : • <u>Water</u> :	Prevent of the second sec	United Sta 22,130 27.1 billion l 3.4 billion lb 686.4 millio 198.2 millio	tes bs ss n lbs n lbs	h of each chemic	il is		
Chemical Functional Use Toxics Release Inventor Monitoring Data Exposure Predictions	e		Data Source: 20 The Toxics Releas recycled, combust Map of TRI Faci	16 Dataset (r E Inventory (Tf ed for energy r Ities Reportin	eleased March 2/ (1) tracks the mana eccovery, treated for ng CHLOROPICAL CHLOROPICAL NITED	018) agement of certain toxi or destruction, and disp	c chemicals that may pose a to osed of or otherwise released	on- and off-site. This Qu N F, F, T T R M M T I I I I	s information is collecti quick Facts for 2C Number of TRI Facilities: Total Production Related Waste Managed: Total On-site Releases: Total On-site: • <u>Air</u> :	Plant Part of the second	United Sta 22,130 27.1 billion I 3.4 billion Ib 686.4 million	tes bs n lbs n lbs os s	h of each chemic	il is		

External Links to Data and Services

🐞 Wolfram Alpha



Links

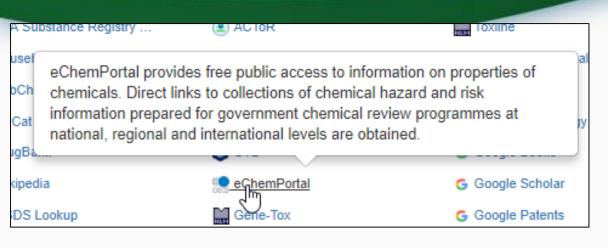
V									
Chemical Properties	Env. Fate/Transport	Hazard	ADME (Beta)	Exposure	Bioassays	Similar Compounds	Related Substances	Synonyms	Literature
General	Toxicol	ogy		Publications		Analytical	P	rediction	
EPA Substance Regist	ry 🛞 ACT	ACToR		Toxline		C RSC Analytic	al Abstracts	2D NMR HSQC/HMBC Pr	
or the second se	ata 🐂 Drug	gPortal		Environmental Health Per		A Tox21 Analyti	cal Data 🧯	Carbon-13 NMF	R Prediction
PubChem	CCF	રાડ		NIEHS		MONA: Mass	Bank North 🧯	Proton NMR Pr	ediction
CPCat	() Che	mView		National Toxi	icology Progr	NET NIST IR Spec	trum 1	ChemRTP Pred	lictor
🤌 DrugBank	С СТС)		G Google Book	s	NET NIST MS Spe	ctrum	LSERD	
W Wikipedia	🧶 eCh	emPortal		G Google Scho	blar				
Q MSDS Lookup	Gen	e-Tox		G Google Pate	nts				
ChEMBL	HSI	ЭB		PPRTVWEB					
Q Chemical Vendors	(E) Tox	Cast Dashboar	rd 2	NII) PubMed					
NIOSH Chemical Safet	ty 🔛 Lac	Med		IRIS Assessi	ments				
ToxPlanet	🔛 Inte	rnational Toxici	ity Esti	🖲 EPA HERO					
< ACS Reagent Chemica	als 🕜 ATS	DR Toxic Subs	stances	C RSC Publica	tions				
W Wikidata	ACT	OR PDF Repo	rt	🚮 BioCaddie D	ataMed				
🍄 ChemHat: Hazards and	d A 🛶 CRE	EST		Springer Mat	terials				

Federal Register

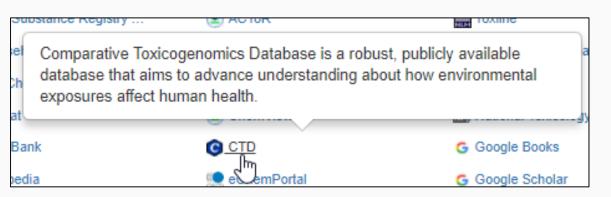
21

Integrated Linkouts









Integrated Linkouts Comparative Toxicogenomics DB



ctc	Illuminating	how chemicals	affect hum	an health							Y	OUR QUERIES C	ONTACT US
			anecthan	ian nealun.						Chemicals	▼ Name, CAS RN, ID		Search ?
Comparative T	oxicogenomics	Database				20							
Home 🔻	Search 📼	Analyze 👻	Down	iload 👻	Help								
🕗 Propyl	ene Glyco												
Basics Ge	ne Interactions	Genes	seases	Phenotypes	5 Com	Pathways GO	Exposure Studies	Exposure Details	References				

These diseases are associated with *Propylene Glycol* or its descendants. Each association is *curated* (M *marker/mechanism* and/or T *therapeutic*) and/or *inferred* (via a curated gene interaction).

Disease categories [Show chart]



Filt	ter by	Association type ALL • Filter					
	0 of 240 results.						
I¶ ₽	First Previous 1	2 3 4 5 Next HLast					
	Chemical 💠	Disease +	Direct Evidence	Enrichment Analysis	Inference Network +	Inference Score	References#
1.	Propylene Glycol	Drug-Related Side Effects and Adverse Reactions	м	100	😵 2 genes: ABCC2 ABCC4	4.09	5
2.	Propylene Glycol	Acute Kidney Injury	Μ	100	😵 2 genes: IL6 TGFB1	3.78	3
3.	Propylene Glycol	Chemical and Drug Induced Liver Injury	м	100	😵 2 genes: ABCC2 IL6	2.82	5
4.	Propylene Glycol	Kidney Diseases	м		1 gene: TGFB1	2.54	4

Submit Comments for Review



Chemistry Dashboard Submit Comment New Comment X Magnesium ascorbyl phospha Comment 113170-55-1 | DTXSID00150312 That dang structure is soooooo ugly (i) Searched by DSSTox Substance Id: Found 1 result for 'DTXSID00150312'. Email address fussyguy@gmail.com Q 💷 🗈 🚣 Q 🗸 Submit Q Find All Chemicals L. Monoisotopic Mass: 577.905013 g/mol Structural Identifiers Linked Substances Må Presence in Lists Record Information

Submit Comments for Review



ie I	New Comment	Synonyms	Literature	Links Comments
e	Details to be submitted with your comment: Text selected: parrot feather watermilfoil Found On: April 26th 2018, 12:59:54 pm Original Query: /dsstoxdb/results? utf8=%E2%9C%93&search=atrazine#eco_ecotox_effect_wel Browser: Chrome 65	/ Туре	Exposure Route	Species
1	Comment	growth	renewal	duckweed
	Is this parrot feather from a Norwegian Blue by any chance? Was it not nailed to the perch properly? Did it fly away looking for the fjords? The FJORDS???	growth	renewal	duckweed
	Email addross	rowth	renewal	waterweed
	Major_Tom_to_Ground_Control@jonsobus.com	growth	renewal	waterweed
		growth	renewal	duckweed
	Submit	growth	renew	duckweed
IC	25 - growth:acute 35.2 mg/m3	growth	renewal	parrot feather w

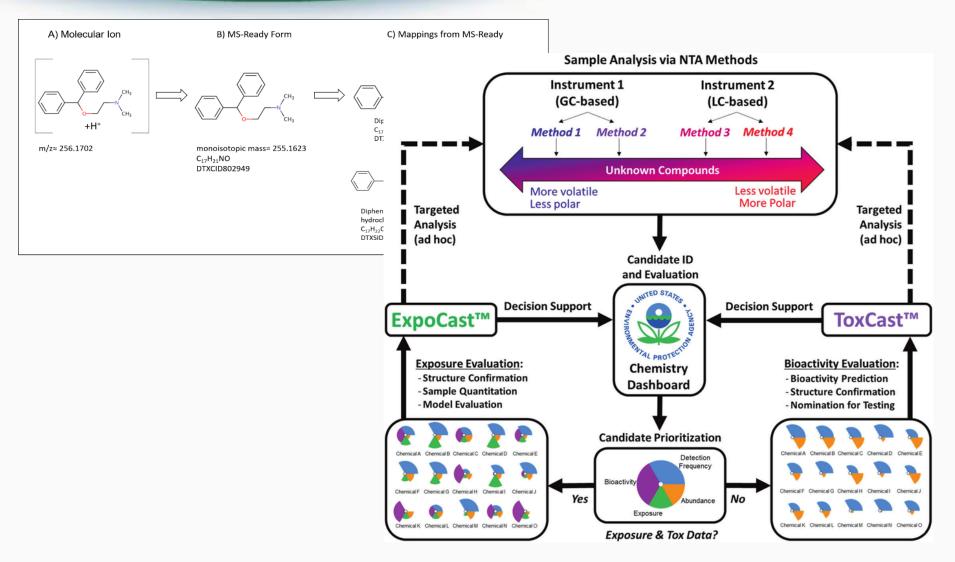
Advanced Searches



	Advance	d Search@		
Mass Search () ± Min/Max M • Mass Da	± Error		Da ppm	Search Q
Molecular Formula Search () Molecular Formula		MS Ready Formu Exact Formula	-	Search Q
Generate Molecular Formula(e)				
Mass Da Default Options: C[1-50] H[0-100] O[0-20] N[0-20] Include Halogens: F[0-20] Cl[0-20] Br[0-20] I[0-			Da ppm	Search Q
Options 👻				

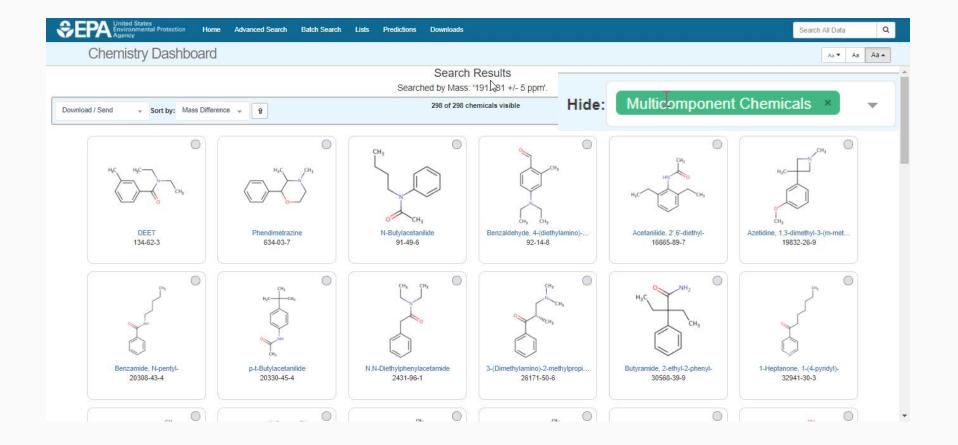
Specific Data-Mappings "MS-Ready Structures"





Advanced Searches



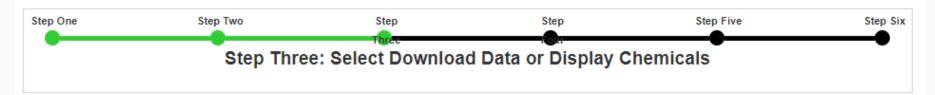


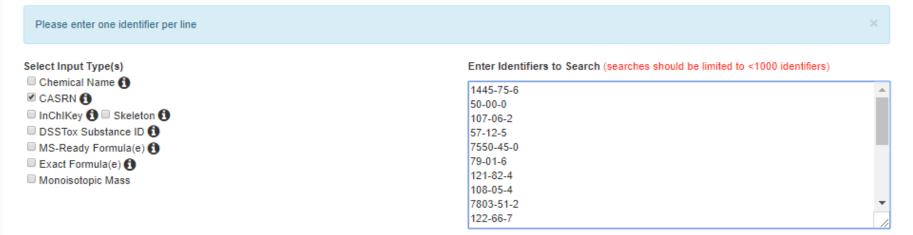
Batch Searches

2





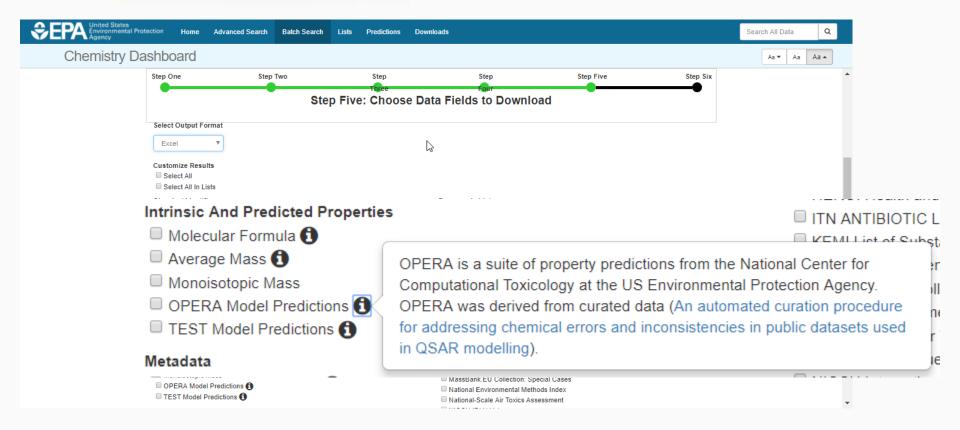




Display All Chemicals Download Chemical Data

OPERA and TEST in Batch

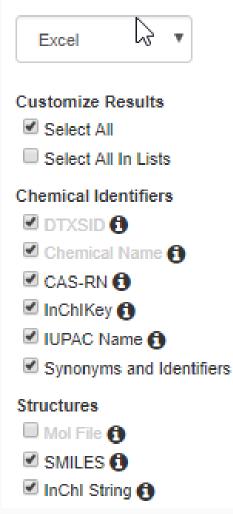




Batch Search



Select Output Format



Intrinsic And Predicted Properties Molecular Formula 6 🗹 Average Mass 🚯 🗹 Monoisotopic Mass 🚯 OPERA Model Predictions f TEST Model Predictions f Metadata. Curation Level Details (1) Data Sources 6 🗹 Assay Hit Count 🚯 Include links to ACToR reports - SLOW! (BETA) 1 NHANES/Predicted Exposure 6 🗹 Include ToxVal Data Availability 🚯 Number of PubMed Articles 6 🗹 Abstract Sifter Input File (Beta) 🖪 MetFrag Input File(Beta) IRIS PPRTV PubChem Data Sources.

Excel Output



	А	В	С	D	E	F	G	Н
1	INPUT	FOUND_BY	DTXSID	PREFERRED_NAME	EXPOCAS	EXPOCAS	NHANES	TOXVAL_D
2	1445-75-6	CAS-RN	DTXSID5024051	Diisopropyl methylpho:	2.09e-08	Y	-	Y
3	50-00-0	CAS-RN	DTXSID7020637	Formaldehyde	1.32e-06	Y	-	Y
4	107-06-2	CAS-RN	DTXSID6020438	1,2-Dichloroethane	4.9e-06	Y	-	Y
5	57-12-5	CAS-RN	DTXSID6023991	Cyanide	-	-	-	Y
6	7550-45-0	CAS-RN	DTXSID8042476	Titanium tetrachloride	-	-	-	Y
7	79-01-6	CAS-RN	DTXSID0021383	Trichloroethylene	7.27e-06	Y	-	Y
8	121-82-4	CAS-RN	DTXSID9024142	Cyclonite	6.72e-08	Y	-	Υ
9	108-05-4	CAS-RN	DTXSID3021431	Vinyl acetate	8.3e-05	Y	-	Y
10	7803-51-2	CAS-RN	DTXSID2021157	Phosphine	-	-	-	Y
11	122-66-7	CAS-RN	DTXSID7020710	1,2-Diphenylhydrazine	1.49e-07	Y	-	Y
12	101-77-9	CAS-RN	DTXSID6022422	4,4'-Methylenedianiline	6.08e-06	Y	-	Y
13	14017-34-6	CAS-RN	DTXSID90161250	Selenium difluoride	-	-	-	-
14	75-44-5	CAS-RN	DTXSID0024260	Phosgene	-	-	-	Υ
15	621-64-7	CAS-RN	DTXSID6021032	N-Nitrosodipropylamine	4.55e-07	Y	-	Υ
16	75-09-2	CAS-RN	DTXSID0020868	Dichloromethane	2.02e-06	Y	-	Υ
17	100-41-4	CAS-RN	DTXSID3020596	Ethylbenzene	8.32e-05	Y	-	Y
18	7440-28-0	CAS-RN	DTXSID2036035	Thallium	-	-	-	Υ
19	108-88-3	CAS-RN	DTXSID7021360	Toluene	8.61e-05	Y	-	Y
20	111-44-4	CAS-RN	DTXSID9020168	Bis(2-chloroethyl) ethe	2.82e-07	Y	-	Υ
21	7440-42-8	CAS-RN	DTXSID3023922	Boron	-	-	-	Υ
22	7440-29-1	CAS-RN	DTXSID6049800	Thorium	-	-	-	Y

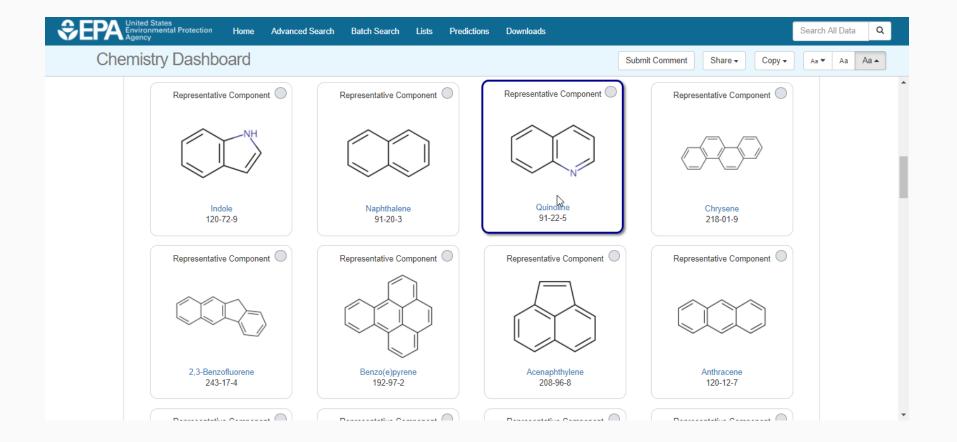
Lists and Families



- Growing chemical categories of interest
 - Polychlorinated biphenyls (PCBs)
 - Polybrominated diphenyl ethers (PBDEs)
 - Polyaromatic hydrocarbons (PAHs)

Searching Families of chemicals Polyaromatic Hydrocarbons





A List of Lists of Chemicals

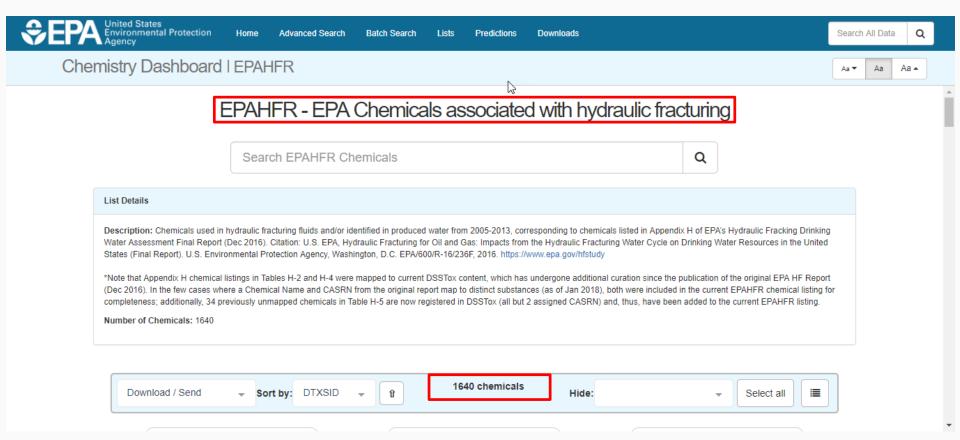
https://comptox.epa.gov/dashboard/chemical_lists



EPA United States Environmental Protection Agency	Home	Advanced Search	Batch Search	Lists	Predictions	Downloads	Search All D:
Chemistry Dashboa	rd			•			Aa▼ Aa Aa▲
			Sele	ect Lis	st		
				R			
List Name	Number of Chemicals	List Descr	iption	Ū			
40CFR355	354		azardous Substand 52 FR 13378)	e List and	Threshold Plann	ing Quantities; Emergency Planning and Relea	ase Notification Requirements;
Algal Toxins	54	A set of alga	I toxins of interest				
Androgen Receptor Chemicals	110		nemicals used to id acs.org/doi/abs/10.1			o AR binding . From Kleinstrauer et al 47	
ATSDR Toxic Substances Portal Chemical List	200	The Agency Human Serv		es and Dis	ease Registry (A	TSDR)is a federal public health agency of the l	U.S. Department of Health and
Bisphenol Compounds	52	This list repr	esents a collection	of Bisphen	ol Compounds		
California Office of Environmental Health Hazard Assessment	972					hazard information including reference exposu or numbers, soil-screening levels, and fish adv	
Chemicals with interesting names	17	This is a list	of chemicals with i	nteresting a	and fun names		
EPA Integrated Risk Information System (IRIS)	510		Program identifies a mical, a group of re			hazards of chemicals found in the environment ex mixture.	nt. Each IRIS assessment can
EPAHFR - EPA Chemicals associated with hydraulic fracturing	1640		ts chemicals associ t Final Report (Dec		ydraulic fracturin	ig from 2005-20013, as reported in EPA's Hydr	aulic Fracturing Drinking Water
EU Cosmetic Ingredients Inventory (Combined 2000/2006)	2878					dients Employed in Cosmetic Products (2000, d for NORMAN by P. von der Ohe (UBA) and F	
EU Toxrisk Dataset	230	Compounds	of interest to the E	U-ToxRisk	Case Studies.		
French Monitoring List	1171					ring activities in France, developed in cooperat o, INERIS, France. Further details on the webs	

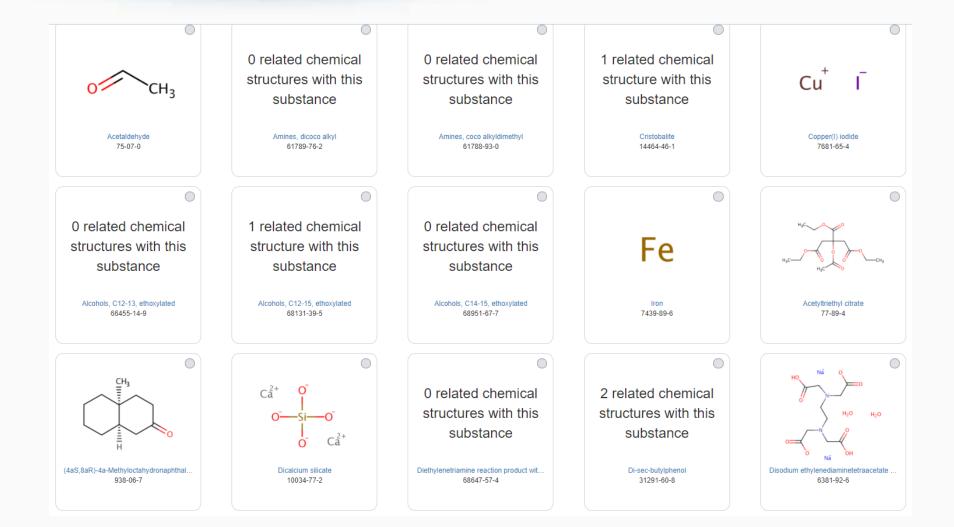
The EPA List of Hydraulic Fracturing Chemicals





Many Hydraulic Fracturing Chemicals are "Complex"



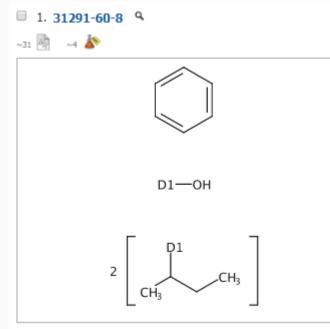


Di-sec-butylphenol



Dashboard Representation

CAS Representation



C14 H22 O Phenol, bis(1-methylpropyl)-

Di-sec-butylphenol

31291-60-8 | DTXSID5049574

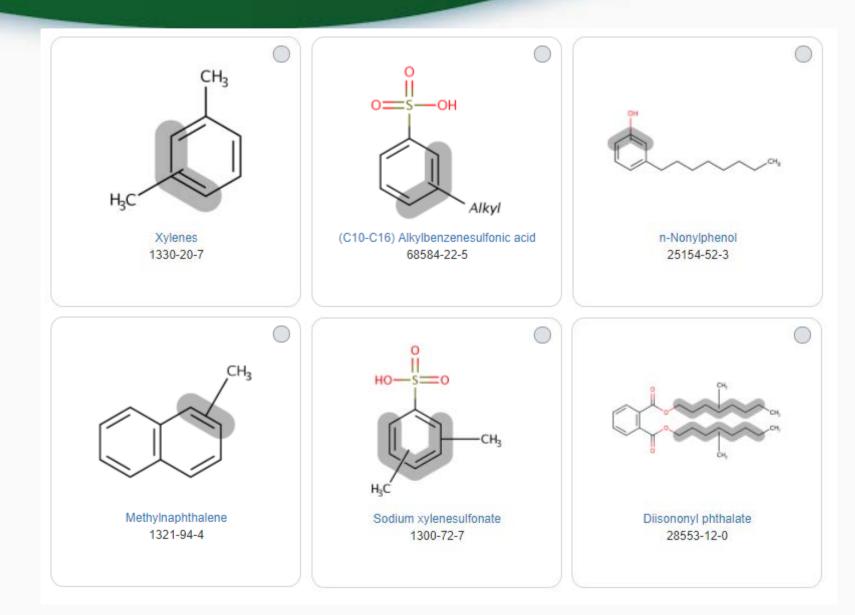
Searched by DSSTox_Substance_Id: Found 1 result for 'DTXSID5049574'.

ecord Information								
auality Control Note	es							
elated Substances	Synonyms	Links	Bioassays	Exposure	Hazard	Comments	Chemical Properties	Literature
Download / Se	end 👻	Sort by:	Relationship	- Î			:	3 chemicals
Download / Se	end 👻	Sort by:	Relationship	• Î			;	3 chemicals
	end 🚽		Relationship		ntative Isomer	0		3 chemicals
			Relationship		_СН3	0		
Se		0	Relationship	Represen	CH ₃			
2 rela	earched Chemical	nical	Relationship	Represen	_СН3	0		
2 rela	earched Chemical ated chem cures with	nical this	Relationship	Represen	CH ₃	•		
2 rela	earched Chemical	nical this	Relationship	Represen	CH ₃			
2 rela struct s	earched Chemical ated chem cures with	nical this	Relationship	Represer	CH ₃		Represent	

"Markush Structures"

https://en.wikipedia.org/wiki/Markush_structure

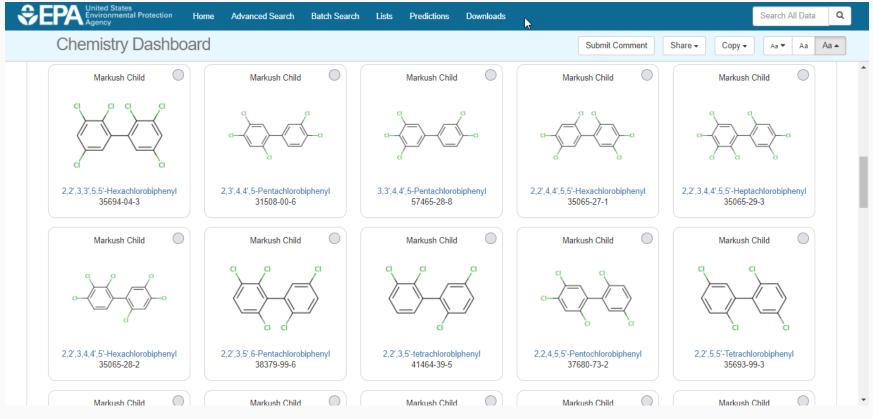




Enumeration of Markush

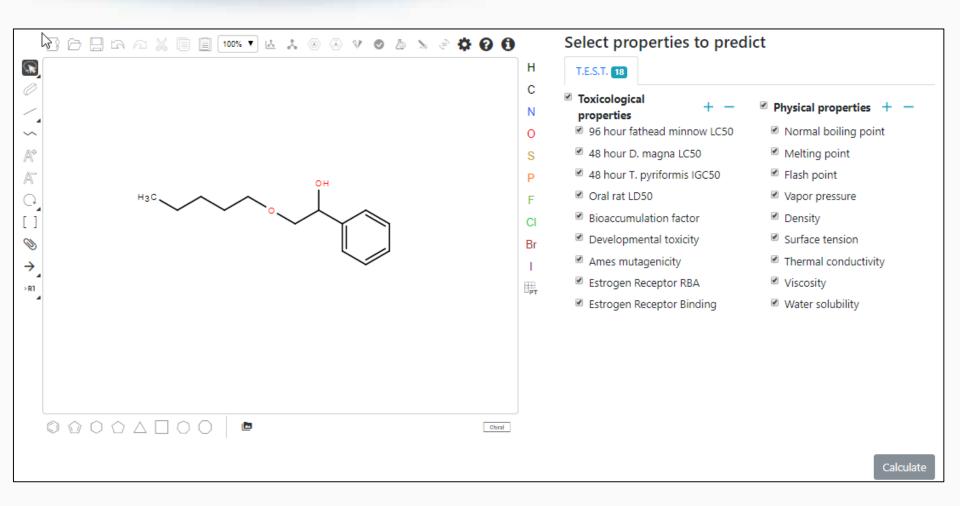


Markush structures can be enumerated into chemical families



Real-Time Predictions





Real-Time Predictions

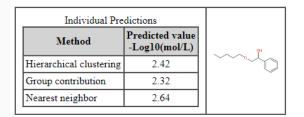


		Experimental			Prediction		
	Property	Value	Consensus	Hierarchical clustering	Single model	Group contribution	Nearest neighbor
6	96 hour fathead minnow LC50		4.477 -Log10(mol/L) 6.954 mg/L	4.195 -Log10(mol/L) 13.288 mg/L	3.994 -Log10(mol/L) 21.110 mg/L	3.478 -Log10(mol/L) 69.224 mg/L	6.238 -Log10(mol/L) 0.120 mg/L
	48 hour D. magna LC50		4.398 -Log10(mol/L) 8.328 mg/L	3.877 -Log10(mol/L) 27.677 mg/L	4.039 -Log10(mol/L) 19.026 mg/L	4.084 -Log10(mol/L) 17.173 mg/L	5.593 -Log10(mol/L) 0.532 mg/L
	48 hour T. pyriformis IGC50		4.063 -Log10(mol/L) 18.039 mg/L	3.731 -Log10(mol/L) 38.668 mg/L		3.386 -Log10(mol/L) 85.610 mg/L	5.070 -Log10(mol/L) 1.773 mg/L
	Oral rat LD50		1.758 -Log10(mol/kg) 3640.950 mg/kg	1.982 -Log10(mol/kg) 2172.756 mg/kg			1.533 -Log10(mol/kg) 6101.245 mg/kg
	Bioaccumulation factor		1.797 Log10 62.700	2.202 Log10 159.310	1.287 Log10 19.346	1.181 Log10 15.157	2.520 Log10 330.834
	Developmental toxicity		false	false	false		true
	Ames mutagenicity		false	false			false
	Estrogen Receptor RBA		-3.075 Log10 8.418*10 ⁻⁴	-3.078 Log10 8.356*10 ⁻⁴	-3.720 Log10 1.907*10 ⁻⁴		-2.427 Log10 0.004
	Estrogen Receptor Binding		true	true	true	false	true

Real-Time Predictions

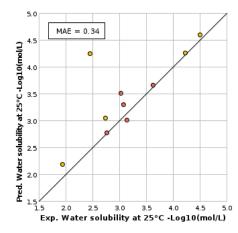
Predicted Water solubility at 25°C for OC(C=1C=CC=CC1)COCCCCC from Consensus method

Prediction results							
Endpoint	Experimental value	Predicted value					
Water solubility at 25°C -Log10(mol/L)	N/A	2.46					
Water solubility at 25°C mg/L	N/A	723.26					



Predictions for the test chemical and for the most similar chemicals

k Prediction results (colors defined in table below)



Chemicals	MAE*
Entire set	0.58
Similarity coefficient ≥ 0.5	0.34
*Mean absolute error in -Log1	0(mol/L)

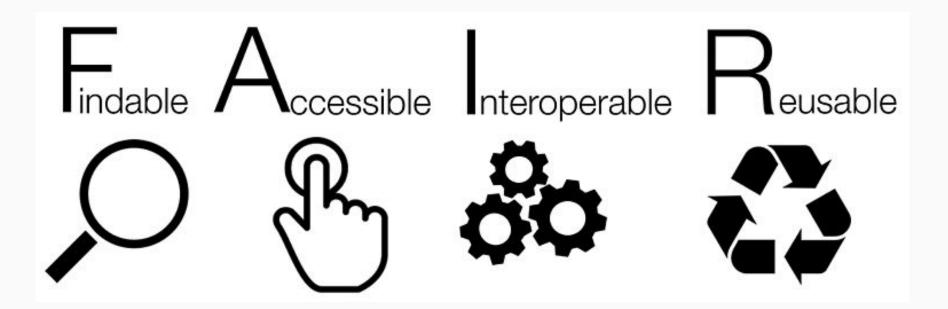
Invironmental Protection

Agency

CAS	Structure	Similarity Coefficient	Experimental value -Log10(mol/L)	Predicted value -Log10(mol/L)
OC(C=1C=CC=CC1)COCCCCC (test chemical)			N/A	2.46
<u>104-40-5</u>	Û	0.68	4.50	4.60
<u>1219-38-1</u>		0.67	4.22	4.26

United States Environmental Protection Agency

Our support for FAIR Data



Downloadable Data

8 75-05-8

9 127-06-0

10 65734-38-5

DTXSID7020009

DTXSID6020010

DTXSID6020012



mistry Dash	board									Aa 🔻	Aa
				D	ownloa	ds					
DSSTox Identifier	to PubChem Ide	ntifier Mapping File							Posted: 11/14/2016		
SID 31638889 31638889 31638888 31638888 31638888 31638888	CID 1 20404 0 10142816 9 50742127 8 19073841 7 11505215 6 25021861 5 2784427	DTXSID DTXSID3087314 DTXSID7087313 DTXSID4087313 DTXSID2087313 DTXSID0087313 DTXSID8087313 DTXSID6087313 DTXSID6087313	13 12 19 17 15 13 11			-	and DSSTox substance id	. ,			
		Numbers and Nam		ber DSST	ox substance ide	ntifier (DTXSI	and the Preferred Name	-	Posted: 11/14/2016		
1 casrn	dsstox_substa	nce_id preferred			on case and in		, and the restored Humo				
2 26148-68-5	DTXSID7020001										
3 107-29-9 4 60-35-5	DTXSID2020004 DTXSID7020005		je oxime								
5 103-90-2	DTXSID702000		hen								
		noecumino									
	DTXSID7020007	Acetohexan	nide								
6 968-81-0 7 18523-69-8	DTXSID7020007 DTXSID2020008		nide (5-nitro-2-furyl)-2-ti	hiazolvl] hv	drazone						

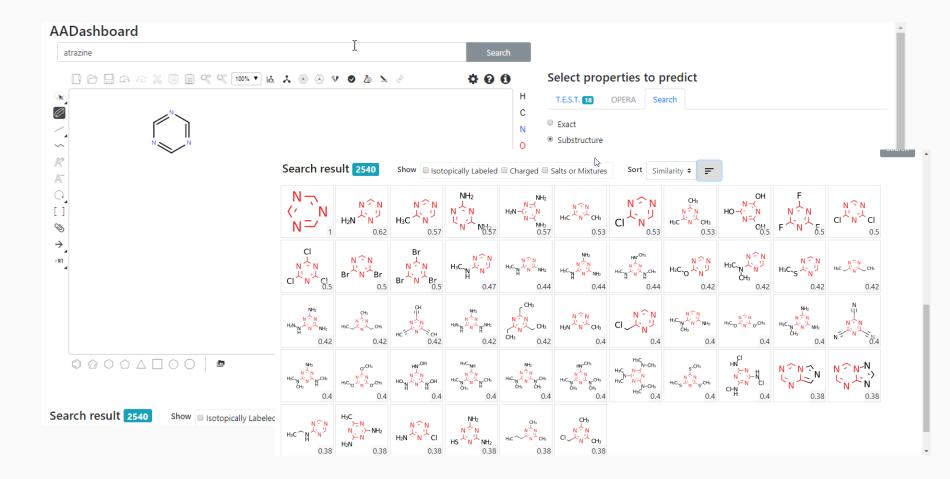
N'-Acetyl-4-(hydroxymethyl) phenylhydrazine

Acetonitrile

Acetoxime

Prototype Development





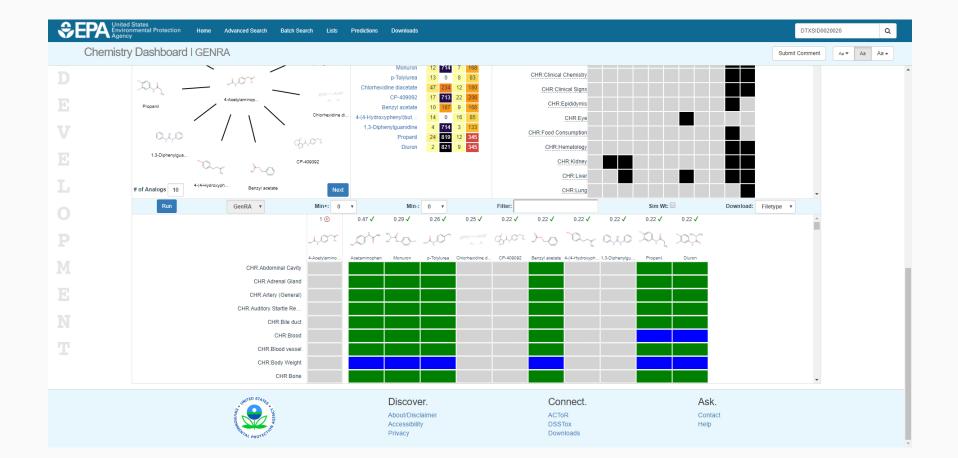
Work in Progress – ToxCast Data Early start transferring EDSP



EPA Environmental Protection Home	Advanced Search Batch Search Lists Predictions Downloads	bisph	nenol A Q
Chemistry Dashboard		Submit C	Comment Aa Aa Aa
	Quality C	ontrol Notes	
Executive Summary (Beta) Chemical Properties	Env. Fate/Transport Hazard ADME (Beta) Exposure Bioassays	GenRA (Beta) Similar Compounds Related Substances Synonyms Literature Links Comments	
ToxCast: Summary			
ToxCast: Data	Assay Selection 1 Selected (A Single Assay Can Have Multiple Charts	Number of Charts: 6
PubChem	Show: Active Inactive Search Search assays	Submit Comment :	Save Chart
	Assay Set: ER (1 of 18 Selected)	150 ACEA_T47D_80hr_Positive	
	ACEA_T47D_80hr_Positive	HITCALL-CATIVE Bisphenol A (80-05-77) F DISXID7020182	
	ATG_ERE_CIS_up		
	ATG_ERa_TRANS_up	Page Page Page Page Page Page Page Page	
	NVS_NR_bER	50-	
	NVS_NR_hER	Cur Off	
	VVS_NR_mERa		
	OT_ER_ER#ER#_1440		
	OT_ER_ERaERb_0480		
	OT_ER_ERaERb_1440	-80 - 1 - 1 - 1 - 2 - 3 Log Concentration (ult)	
	OT_ER_ERbERb_0480	Constant Model Cain-Loss Model Hill Model	
	OT_ER_ERbERb_1440	Witning Model Model AIC PMISE PROE Top AC50 AC50 Stope Stope Error	
	OT_ERa_EREGFP_0120	Constant 185.05 80.61 0.00	
	OT_ERa_EREGFP_0480	√ Gan-Lass 154.83 159.2 1.00 139.50 0.42 21.98 1.48 13.42 2.85 H0 172.42 44.58 0.00 105.19 0.31 - 2.16 - 3.44	
	Tox21_ERa_BLA_Agonist_ratio		
	Tox21_ERa_BLA_Antagonist_ratio		

Work in Progress – GenRA





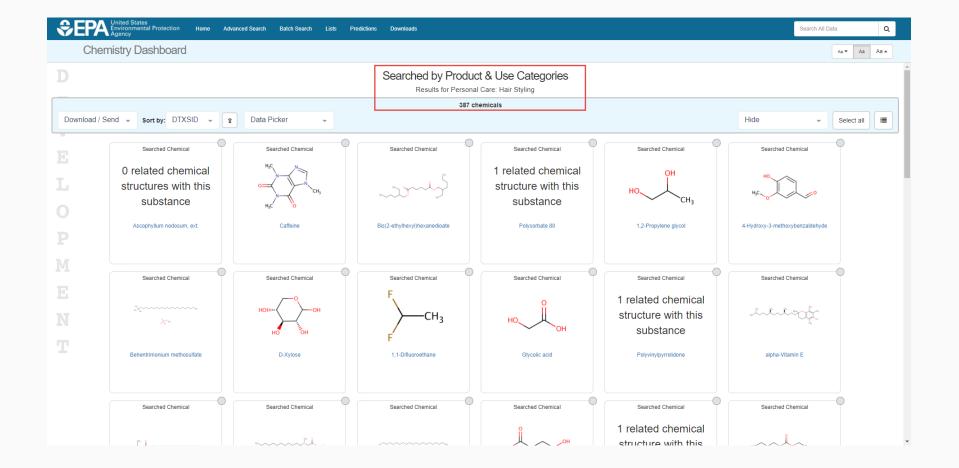
What chemicals in What Products?



\$	EPA United States Environmenta Agency	I Protection Hom	e Advanced Search	Batch Search	h Lists P	redictions	Downloads								METHY	L PARABEN	۵	
	Chemistry Da	shboard													Submit Commen	t Aa 🕶	Aa Aa 🔺	
D		Methylpa	araben XSID4022529															
Е			m from Valid Source: Found	1 result for 'MF	THYL PARABEN													
V			±- Q-															
						Wi	kipedia											
E	H ₃ C						thylparaben, also r lead more	nethyl paraben, one o	of the parabens, is a prese	rvative with the chemical fo	ormula CH3(C6H	4(OH)COO). It is	the methyl	ester of p-hydrox	ybenzoic acid.			
1.1.1						Int	rinsic Properties											
		но				Str	Structural Identifiers											
						Lin	Linked Substances											
			(Pre	esence in Lists											
Μ						Re	cord Information											
E						Qu	ality Control Not	es										
	Executive Summary (Beta)	Chemical Properties	Env. Fate/Transport	Hazard	ADME (Beta)	Exposure	Bioassays	GenRA (Beta)	Similar Compounds	Related Substances	Synonyms	Literature	Links	Comments				
N	Product & Use Categories								Product & Use Cat									
	Chemical Weight Fraction		Download as: TSV	Excel					Product & Use Cat	egones (POCs)								
	Chemical Functional Use		Broduct or Lieo Cotor	orization				Categorization for	20			Number of Ur	ique Dredu	icte				
	Monitoring Data		Product or Use Categ	UTZALION	1		\$	Categorization typ	些		÷	Number of Un	iique Produ	icis			•	
	Exposure Predictions		personal care: face crea					PUC				147						
	Production Volume		personal care: hand/bo personal care: hair styli					PUC				142						
					1													

What chemicals in What Products?





Coming Soon





Env	viro	nm	enta	l To	pics

Laws & Regulations

About EPA

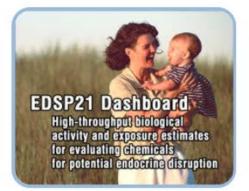
Search EPA.gov





ACTOR Appropriated Publicly Available Chamfeel Date S ACTOR ady. -





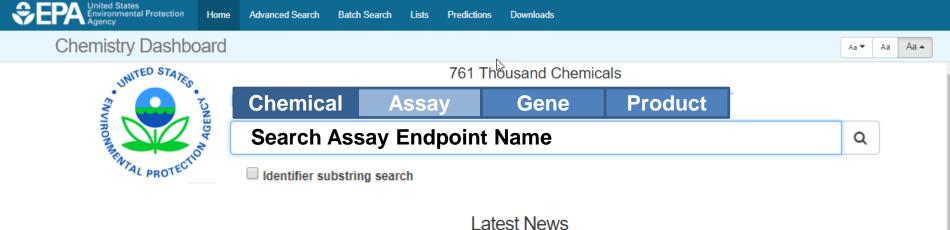




Future Search Possibilities



►



Read more news

An article regarding an Excel Version of the Abstract Sifter is published.

March 7th, 2018 at 9:21:27 AM

The abstract sifter that is integrated into the Dashboard (for example here for Atrazine) is available as an Excel add-in. Our recent article on the Abstract Sifter for Excel has been published.

. . . .

April 1st was the 2nd birthday



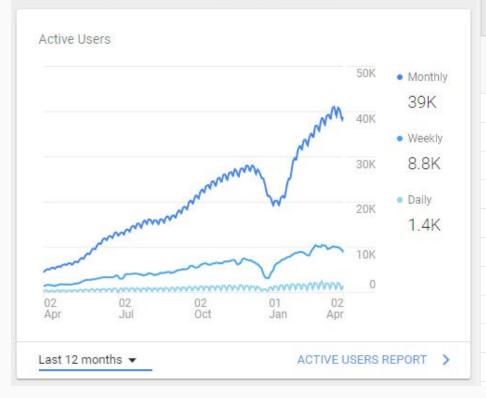


Take up





How are your active users trending over time?



	ountry ?	Acquisition						
	ountry T	Users ?	New Users					
		8,766 % of Total: 100.00% (8,766)	7,268 % of Total: 100.06% (7,264)					
1.	United States	3,153 (35.97%)	2,611 (35.92%)					
2.	💶 India	910 (10.38%)	741 (10.20%)					
3.	📧 South Korea	450 (5.13%)	370 (5.09%)					
4.	💽 Canada	334 (3.81%)	267 (3.67%)					
5.	United Kingdom	287 (3.27%)	241 (3.32%)					
6.	Japan	285 (3.25%)	221 (3.04%)					
7.	🔲 Germany	239 (2.73%)	188 (2.59%)					
8.	France	227 (2.59%)	183 (2.52%)					
9.	China	177 (2.02%)	136 (1.87%)					
10.	😒 Brazil	140 (1.60%)	117 (1.61%)					

Conclusion



- The CompTox Chemistry Dashboard provides access to data for ~760,000 chemicals
- High quality data from ongoing curation efforts
- An integration hub for multiple "modules"
 - Experimental and predicted properties
 - Human and Ecological Hazard data
 - Exposure data products, data in the environment
 - In vitro bioassay data ToxCast/Tox21
 - Literature searching Google Scholar and PubMed
 - Specialized searches mass/formula for analytical support
 - Batch searching and Real Time Predictions
- The primary architecture for NCCT data



- The NCCT CompTox Chemistry Dashboard Development Team
- The NCCT Team
- NERL scientists MS, Exposure, CPDat
- Kamel Mansouri OPERA models
- Todd Martin TEST predictions





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