

US-EPA Chemicals Dashboard – an integrated data hub for environmental science

Antony John Williams

Center for Computational Toxicology and Exposure, US-EPA, RTP, NC

May 5th 2021 : Lehman College



- Short overview of the CompTox Chemicals Dashboard
- Live demo please ask questions as we go!!!

SLIDES WILL BE MADE AVAILABLE AT THE END

CompTox Chemicals Dashboard https://comptox.epa.gov/dashboard

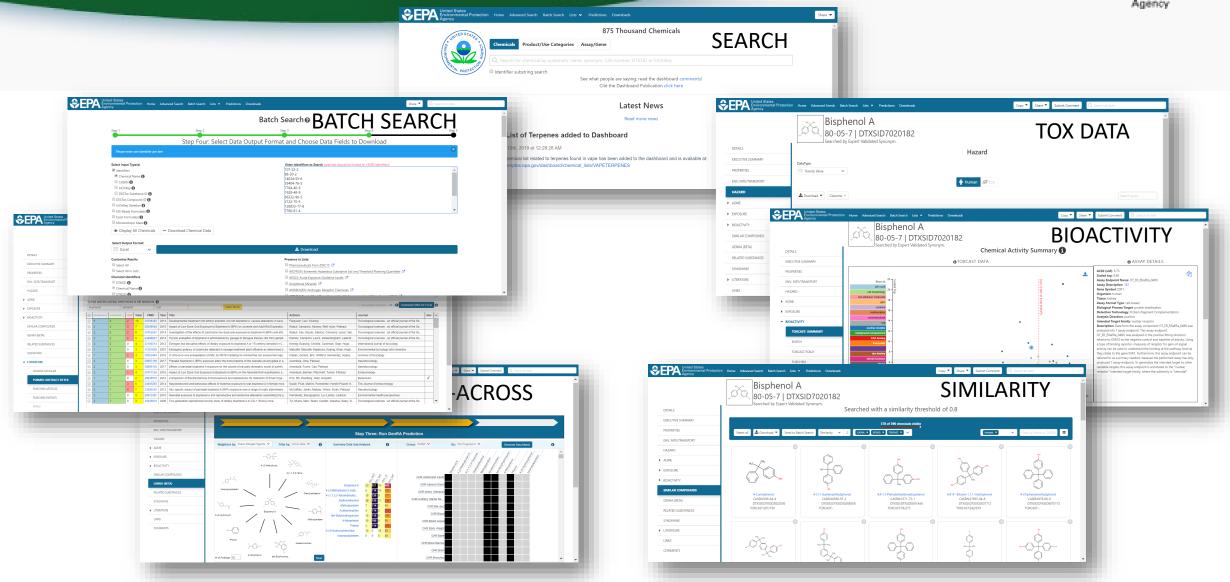


A publicly accessible website delivering access:

- ~883,000 chemicals with related property data
- Experimental and predicted physicochemical property data
- Experimental Human and Ecological hazard 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

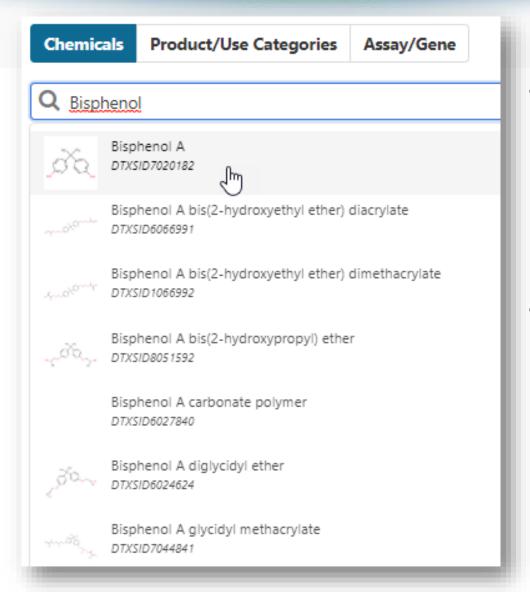
CompTox Chemicals Dashboard https://comptox.epa.gov/dashboard





BASIC Search

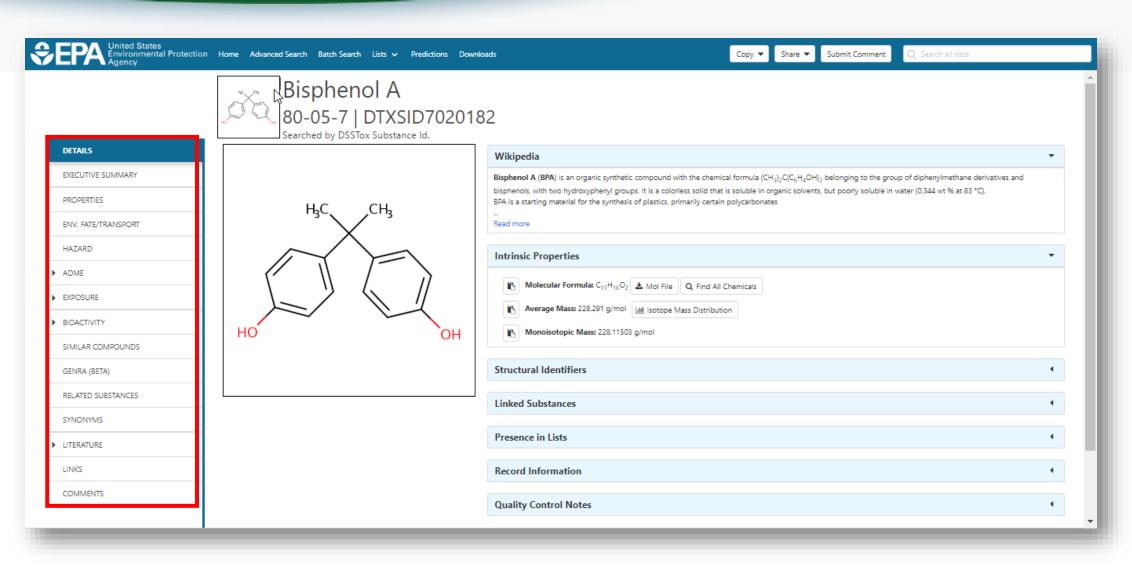




- Three searches from the home page
 - Chemicals (Names, CASRN, Substring)
 - Product/Use Categories
 - Assay Gene
- Searching millions of synonyms and includes Active, Deleted and Alternate CASRNs

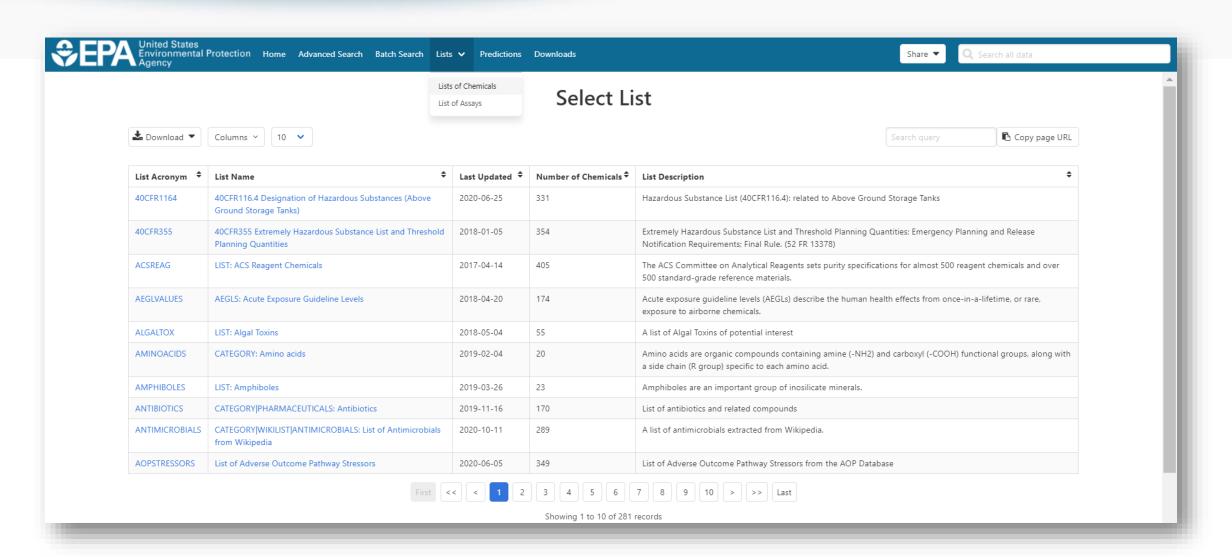
Detailed Chemical Pages





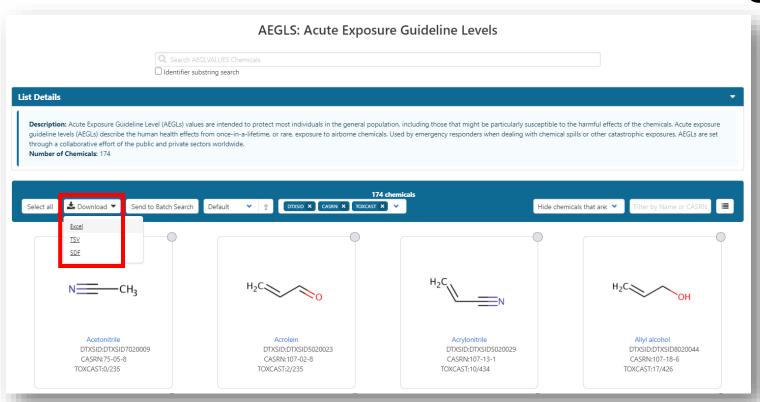
Lists of Lists of Chemicals >290 lists and growing





List of Chemicals Example: Acute Exposure Guideline Levels





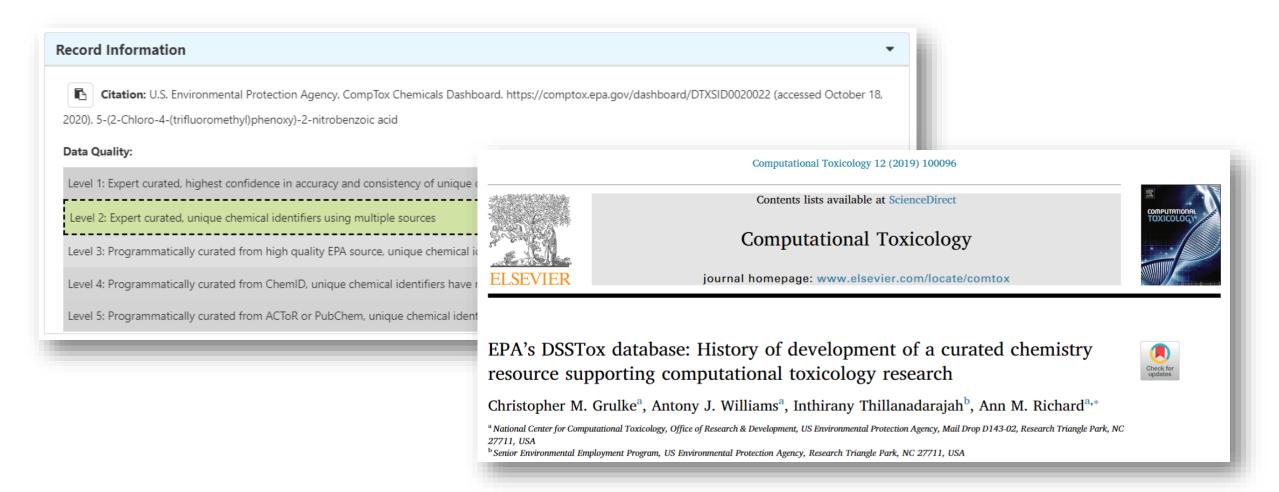
Chemical List Examples

- Disinfection by-products
- EPA Pesticide Search DB
- Consumer Products DB
- Chemicals in Biosolids
- ATSDR Toxicological Profiles
- Canadian Domestic Substances
- Chemical Inventory for ToxCast
- EFSA OpenFoodTox
- EDSP Universe of Chemicals
- Hydraulic Fracturing Chemicals

Quality



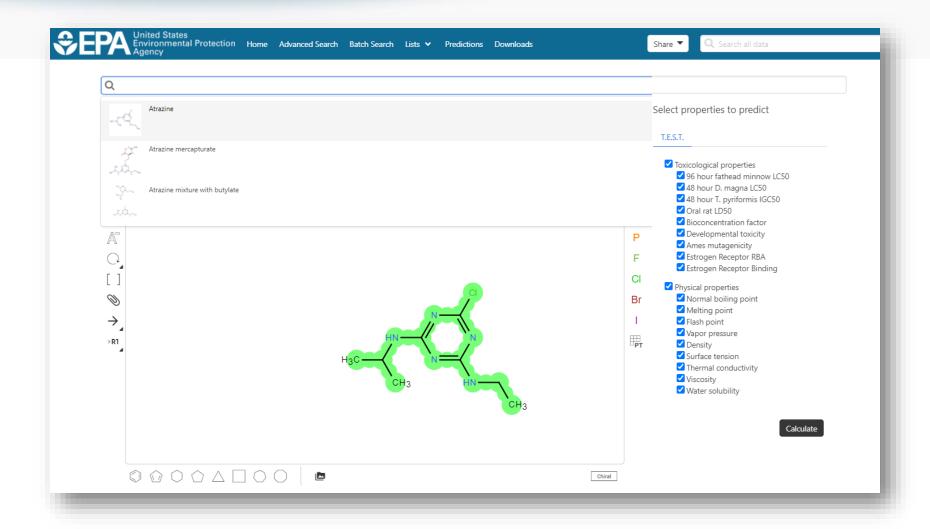
Five full time curators register and curate data to elevate quality



Predictions for New Chemicals

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





Predictions for New Chemicals

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



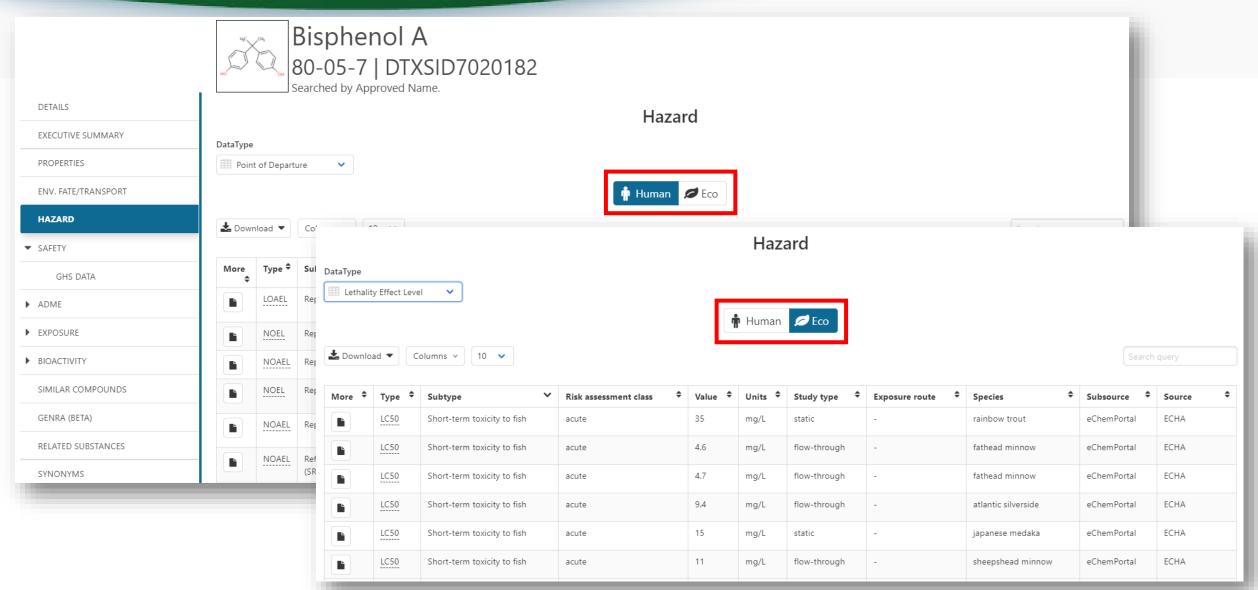
Provider: T.E.S.T.

♣ Download Summary	•
--------------------	---

Property	Experimental Value	Consensus	Hierarchical clustering	Single model	Group contribution	Nearest neighbor
96 hour fathead minnow LC50		3.969 -Log10(mol/L) 22.469 mg/L	3.936 -Log10(mol/L) 24.234 mg/L	3.742 -Log10(mol/L) 37.882 mg/L	4.135 -Log10(mol/L) 15.355 mg/L	4.064 -Log10(mol/L) 18.079 mg/L
48 hour D. magna LC50		3.791 -Log10(mol/L) 33.887 mg/L	3.568 -Log10(mol/L) 56.592 mg/L	4.243 -Log10(mol/L) 11.976 mg/L	3.832 -Log10(mol/L) 30.824 mg/L	3.521 -Log10(mol/L) 63.123 mg/L
48 hour T. pyriformis IGC50			2.986 -Log10(mol/L) 216.359 mg/L			
Oral rat LD50		2.439 -Log10(mol/kg) 761.182 mg/kg	2.355 -Log10(mol/kg) 924.254 mg/kg			2.524 -Log10(mol/kg) 626.882 mg/kg
Bioconcentration factor		0.847 Log10 7.038	0.906 Log10 8.051	0.690 Log10 4.903	0.820 Log10 6.607	0.974 Log10 9.409
Developmental toxicity		true	true	true		
Ames mutagenicity		false	true			false
Estrogen Receptor RBA						
Estrogen Receptor Binding		false	false	false	false	
Normal boiling point		315.9 ℃	318.0 °C		345.4 °C	284.3 °C
Melting point		114.0 °C	105.0 °C		79.4 °C	157.7 °C
Flash point		176.1 °C	180.1 °C		172.5 °C	175.7 °C
Vapor pressure		-6.072 Log10(mmHg) 8.464*10^-7 mmHg	-6.164 Log10(mmHg) 6.851*10^-7 mmHg		-5.679 Log10(mmHg) 2.093*10^-6 mmHg	-6.374 Log10(mmHg) 4.23*10^-7 mmHg

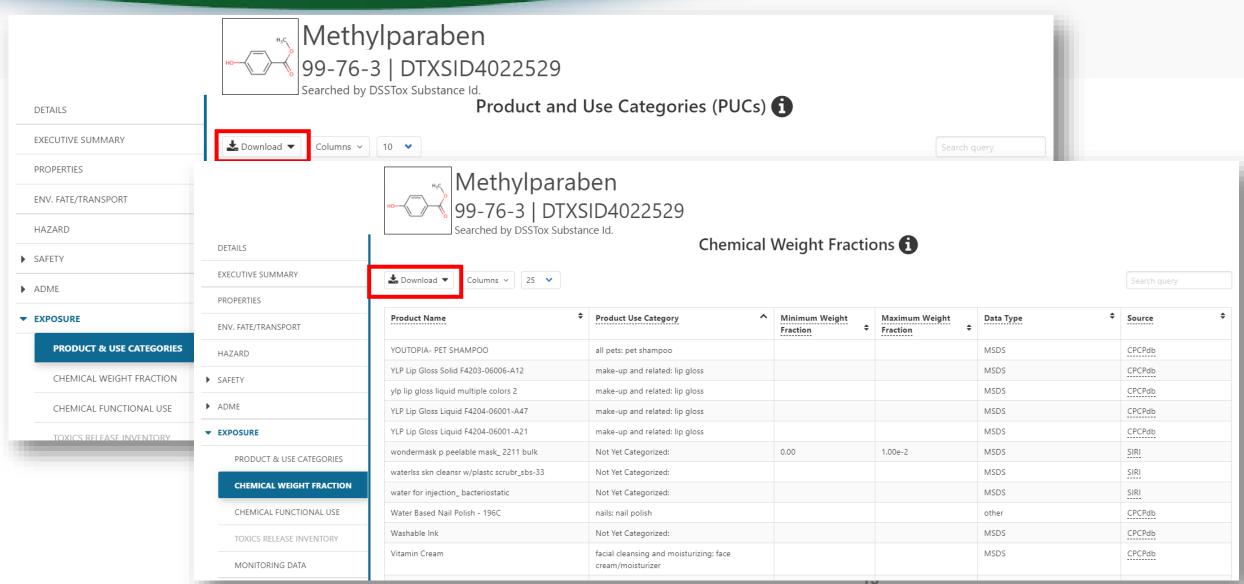
Hazard Data (Human and Eco)





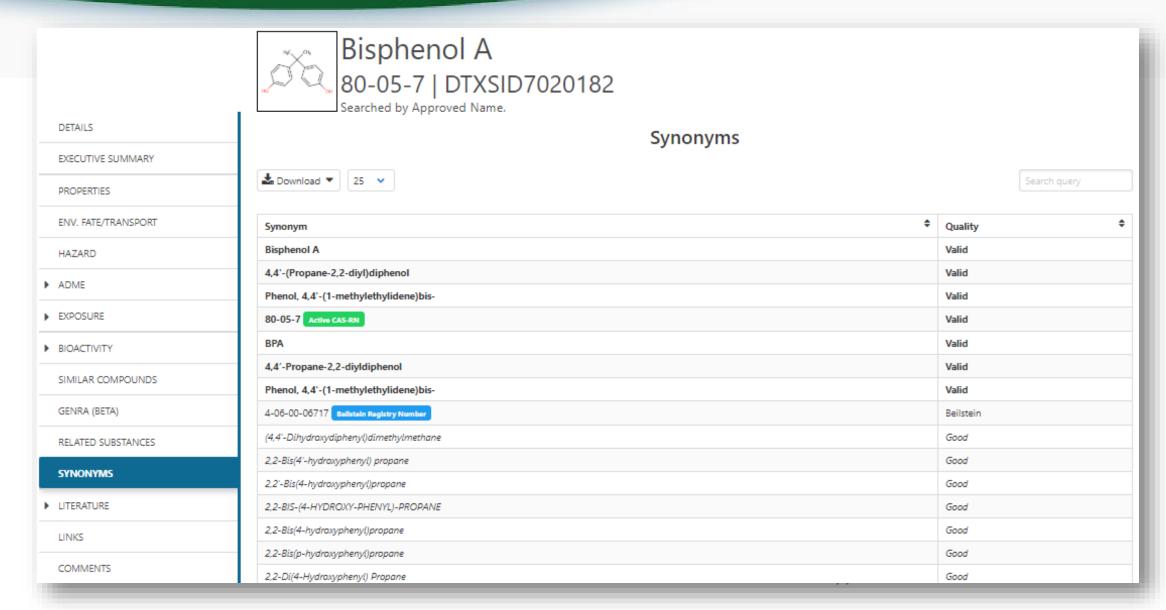
Sources of Exposure to Chemicals





Identifiers to Support Searches

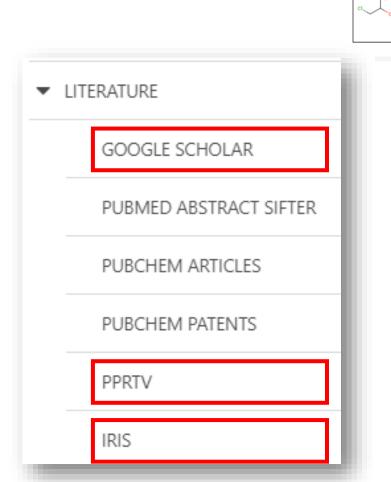




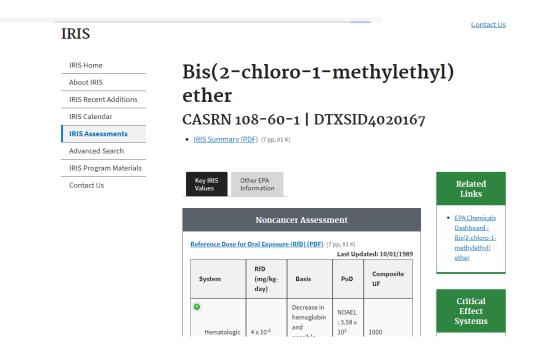
Identifiers are used in the app



Identifiers are used to feed and link into "Literature"



bis(2-Chloro-1-methylethyl) ether 108-60-1 | DTXSID4020167



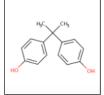
External Links – Also use Identifiers Names, CASRN, PubChem IDs...



	80-05-7 D' Searched by Approved	TXSID7020182 I Name.			
DETAILS	General	Toxicology	Publications	Analytical	Prediction
EXECUTIVE SUMMARY	● EPA Substance Registry Service	ACTOR	Toxline	☑ FOR-IDENT	3 2D NMR HSQC/HMBC Predict
PROPERTIES	♣ Household Products Database ♣ Chemical Entities of Biological Interest ♣ Chemical Entitle Entit Entitle Entit	OH, DrugPortal ☐ CCRIS	Environmental Health Perspectives NIEHS	NEMI: National Environmental Methods Index RSC Analytical Abstracts	Carbon-13 NMR Prediction Proton NMR Prediction
ENV. FATE/TRANSPORT	(ChEBI)	(a) ChemView	National Toxicology Program	♠ Tox21 Analytical Data	ChemRTP Predictor
HAZARD	PubChem Chemspider	CTD CTD	G Google Books G Google Scholar	MONA: MassBank North America mzCloud	
ADME	(B) CPCat	Gene-Tox	G Google Patents	NIST IR Spectrum	
EXPOSURE	DrugBank Amp HMDB	HSDB (a) ToxCast Dashboard 2	PPRTVWEB PubMed	NST NIST MS Spectrum	
BIOACTIVITY	W Wikipedia Q MSDS Lookup	LactMed	IRIS Assessments		
SIMILAR COMPOUNDS	© ChEMBL	International Toxicity Estimates for Risk ATSDR Toxic Substances Portal	● EPA HERO ■ NIOSH Skin Notation Profiles		
GENRA (BETA)	Q Chemical Vendors CalEPA Office of Environmental Health	Superfund Chemical Data matrix NIOSH IDLH Values	NIOSH Skin Notation Profiles NIOSH Pocket Guide RSC Publications		
RELATED SUBSTANCES	Hazard Assessment NIOSH Chemical Safety Cards	ACTOR PDF Report	■ RSC Publications ■ BioCaddie DataMed		
SYNONYMS	© ToxPlanet	(B) Toxics Release Inventory	Springer Materials		
LITERATURE	ACS Reagent Chemicals W Wikidata	CREST National Air Toxics Assessment	Federal Register Regulations.gov		
LINKS	ChemHat: Hazards and Alternatives Toolbox		Bielefeld Academic Search Engine		
COMMENTS	Wolfram Alpha ScrubChem		CORE Literature Search		

External Links





Bisphenol A 80-05-7 | DTXSID7020182

Searched by DSSTox Substance Id.

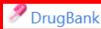
General

EPA Substance Registry Service



Chemspider

CPCat



W Wikipedia

Q MSDS Lookup

ChEMBI

ToxPlanet

ACS Reagent Chemicals

ChemHat: Hazards and Alternatives Toolbox

Toxicology

ACToR

оң DrugPortal

CCRIS

ChemView





eChemPortal



HSDB

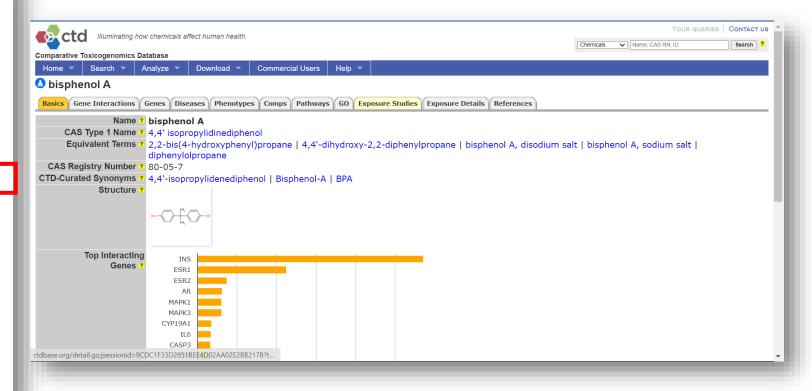
ToxCast Dashboard 2

● ACToR PDF Report

CREST

National Air Toxics
Assessment

 Links to ~90 websites providing access to additional data on the chemical of interest

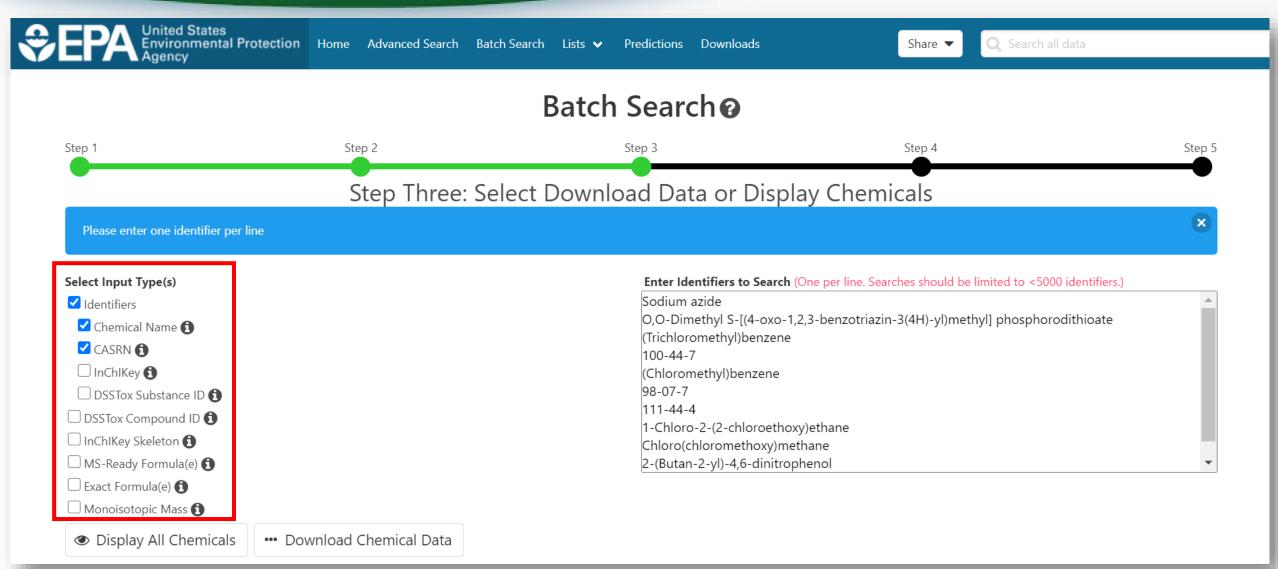




Batch Searching

Access data *en masse* for thousands of chemicals....





Select Output Format and Content



Step Four: Select Data Output Format and Choose Data Fields to Download

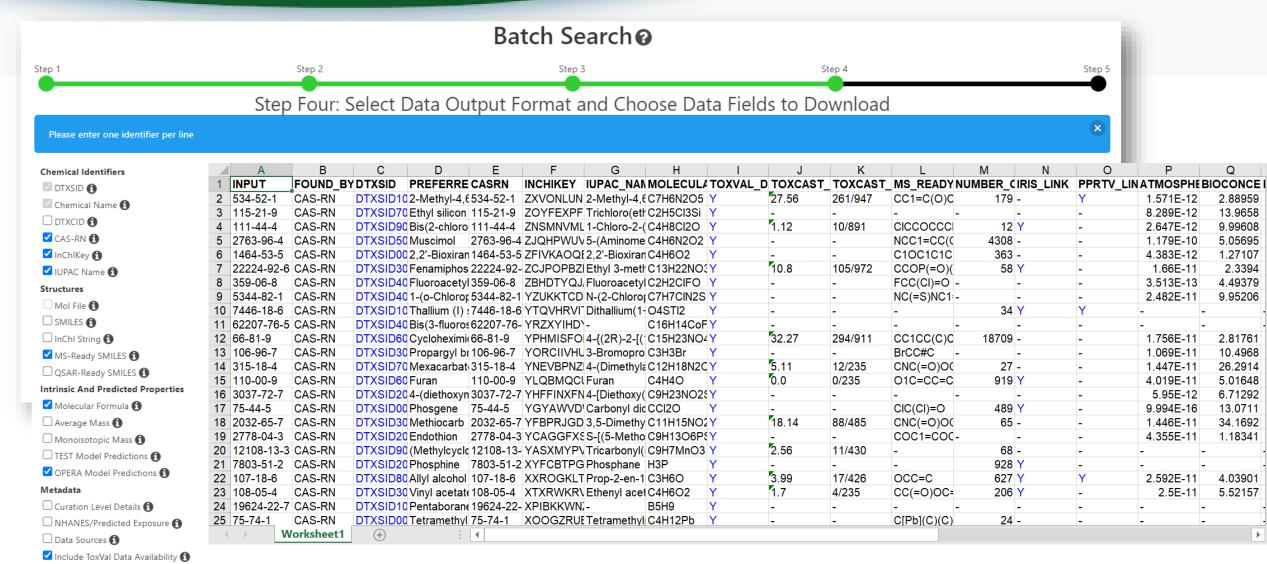
		▲ Download
Customize Results	Intrinsic And Predicted Properties	Presence in Lists:
Select All	☐ Molecular Formula 1	☐ 40CFR116.4 Designation of Hazardous Substances (Above Ground Storage Tanks) 🗹
Select All in Lists	Average Mass 1	☐ 40CFR355 Extremely Hazardous Substance List and Threshold Planning Quantities 🗹
Chemical Identifiers	☐ Monoisotopic Mass 1	☐ AEGLS: Acute Exposure Guideline Levels ☑
✓ DTXSID (1)	☐ TEST Model Predictions 1	ANDROGEN: Androgen Receptor Chemicals 🖸
Chemical Name (1)	OPERA Model Predictions (1)	ARTICLE; Bench-Mark Dose Human Health Assessment List (Wignall et al., 2014)
□ CAS-RN	Metadata ☐ Curation Level Details ① ☐ NHANES/Predicted Exposure ①	ARTICLE: Collaborative Estrogen Receptor Activity Prediction Project (CERAPP) ARTICLE: Collaborative Estrogen Receptor Activity Prediction Project (COMPARA)
	☐ Include ToxVal Data Availability ☐ Assay Hit Count ☐ Number of PubMed Articles ☐ PubChem Data Sources ☐ CPDat Product Occurrence Count ☐ IRIS ☐ PPRTV ☐ PPRTV ☐ PPRTV	ATSDR: Minimal Risk Levels (MRLs) for Hazardous Substances ATSDR: Minimal Risk Levels (MRLs) for Hazardous Substances ATSDR: Toxic Substances Portal Chemical List California Office of Environmental Health Hazard Assessment Canadian Domestic Substances List 2019 CATEGORY: Amino acids CATEGORY: Color Index dyes CATEGORY: Flame Retardants CATEGORY: Flame Reta
	☐ Wikipedia Article☐ QC Notes 1☐ Include links to ACToR reports - SLOW! (BETA) 1	

Batch Search CASRNs

✓ Assay Hit Count <a>6

✓ Number of PubMed Articles ♠

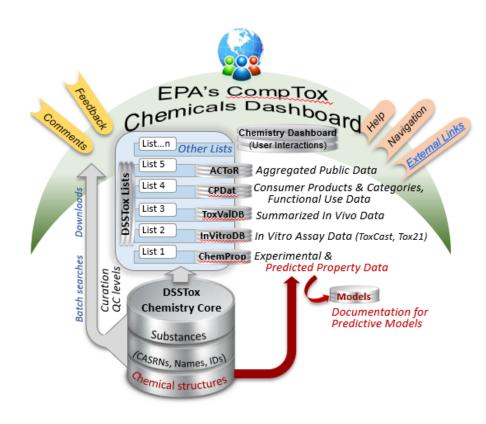




Conclusion



- Dashboard access to data for ~883,000 chemicals (and growing)
- Data aggregation for ALL data sources continues unabated
- Data curation is essential. Bigger does NOT mean better
- Flexible search capabilities continue to expand release-to-release but our efforts are in a rearchitecture effort



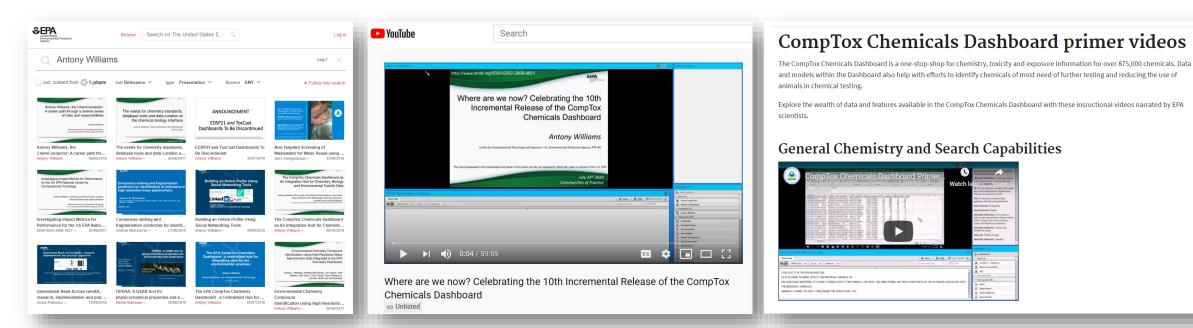


Online Materials

You want to know more...



- Lots of resources available
 - Presentations: https://tinyurl.com/w5hqs55
 - Communities of Practice Videos: https://rb.gy/qsbno1
 - Manual: https://rb.gy/4fgydc
 - Latest News: https://comptox.epa.gov/dashboard/news info



Acknowledgments



- Contact: Williams.Antony@epa.gov
- Feedback and follow-up is welcomed!
- The dashboard is based on the efforts of many team members. Many collaborators provide data also.



EPA's Center for Computational Toxicology and Exposure