

Mobilizing EPA's CompTox Chemistry Dashboard Data on Mobile Devices ¹Antony Williams^{*}, ²Kirill Blinov, ³Andrew McEachran and ¹Chris Grulke

¹U.S. EPA, Office of Research and Development, National Center for Computational Toxicology (NCCT), Research Triangle Park, NC, ²Molecules Apps, OR and ³Oak Ridge Institute for Science and Education (ORISE) Participant, Research Triangle Park, NC,

Project Description and Goals

Problem: Browser-based access via the internet can be limiting in terms of having "always accessible" chemical datasets. How can the ~720,000 chemical structures available via the web-based CompTox Chemistry dashboard be made available on a mobile device? Is it possible to integrate additional capabilities to support structure identification above and beyond simply mass-and formula-based searching available on the dashboard.

Goals: To deliver access to the chemical structure set associated with the CompTox Chemistry Dashboard on a mobile device and enhance the data using analytical spectroscopy parameters of value to support structure identification.

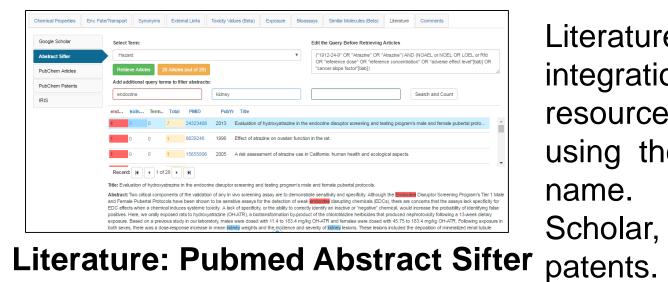
Abstract

The EPA's National Center of Computational Toxicology (NCCT) CompTox Chemistry Dashboard provides access to chemistry data for about 750,000 chemical substances, ~720,000 of these are represented by chemical structures. The application is used to source data such as physicochemical property data, bioassay screening data, functional use, and product composition data. The dashboard allows flexible searches based on chemical names, CAS numbers and mass and formula to support non-targeted screening research for mass spectrometry.

Since NCCT provides open data it is therefore available for reuse and redistribution in other forms of software. Free access to these data has resulted in the development of mobile applications supporting identifier and mass-based searching of the data. These software applications provide access to hundreds of thousands of chemical structures on an iOS mobile device (i.e. iPhone and iPad) to support compound identification. The CompTox Mobile app provides access to chemistry data via a chemical search on a handheld device and directly connects to the web-based dashboard.

Open Data

Large collections of the data accessed via the CompTox Chemistry dashboard can be downloaded and reused. The dashboard provides access to a number of identifier mapping files (e.g. CAS Number-Name-InChIKeys and to the SDF file used in this work. Data are available from http://comptox.ag.epa.gov/dashboard/downloads.



The CompTox Chemistry Dashboard



Dashboard Entry Page

For those records with associated chemical structure representations various inherent properties (for example, formula and mass) as well as predicted physicochemical properties (logP, water solubility etc.) are provided.

Chemical Properties	Synon	yms	External L	inks	Env. Fate	/Transport	Toxicity	Values	Bioassays	Expos	ure	Literature	Simil	ar Molecules	Comments
Summary									LogP	: Octano	I-Water	r			
LogP: Octanol-Wa	ter							Average			Media	in		Range	
Water Solubility				Experime	ental			2.61 (1)			2.61			2.61	
Water Solubility				Predicte	d			2.76 (4)			2.76			2.50 to 3.05	
Density		Do	wnload as:	TSV	Excel	SDF									
Melting Point															
Boiling Point									E	kperimen	tal				
Boiling Point		S	ource				Result								
Surface Tension		P	hysPropNCC	т			2.61								
Vapor Pressure										Predicted	1				
LogKoa: Octanol-Air		S	ource				Result		Calculat	ion Detai	ils				QMRF
Henry's Law		E	PISUITE				2.82		Not Avail	able					Not Available
		C	PERA				3.05		OPERA	Model Rep	port				Available
Index of Refraction		N	IICEATM				2.50		Not Avail	able					Available
Molar Refractivity		А	CD/Labs				2.67		Not Avail	able					Not Available
pKa Basic Apparent															

Chemical Properties Panel

The Toxicity Values tab provides access to data assembled from a public resources including EPA data (i.e. IRIS and PPRTV reports, ToxRef DB). Data can be downloaded as TSV and Excel

The landing the page for dashboard is a simple text entry box allowing a type-ahead search for systematic, trade and trivial names, CAS Registry Numbers and InChls.

United States Environmental Protection Agency	Home	Advanced Search	Line					500	sh Chemistry Dashboa	rd .	Q	Option
hemistry Dashboa	rd								Submit Comment	Copy •	A4 *	Aa Aa
Atrazine 1912-249 (DTXSII O Searches by Approve a m n a		result for 'attractine'.	(com) ar agricultur	is an herbicide of the id sugarcane and on re. It was banned in th	turf, such as golf he European Uni	courses and re- ion in 2004, whe	idential lawns. It in the EU found (d posternerigence broa is one of the most wide	ly used herbicides in US	and Australian		
нзс		СНа	Intrinsic	Properties al Identifiera	het this could be	prevented nor t	hal these Kee	d more				
				Information								
Chemical Properties	Synonyma	External Links	Env. Fate/Transport	Toxicity Values	Bioeseeys	Exposure	Literature	Similar Molecules	Commenta			
	About	Contact	Privacy	international and the second s			DSSTex	Accessibility	Heb	Download		

Chemical Record Page: Atrazine

Where possible, links are provided to related Wikipedia articles. An associated mol file is available for download to the desktop, and a summary report containing record data can be provided as a PDF file.

hemical Properties Env. Fa	te/Transport Synor	nyms Exte	ernal Links	oxicity Values (Beta	a) Exposure	Bioassays	Similar Mo	olecules (Beta)	Literature	Comments			
Bioavailability Metric	Download as	TSV E	xcel										
Exposure Limit						_							
Point Of Departure						Po	int Of Depai						
Regulatory Toxicity Value	Grouping	Priority	0 Туре	Subtype	Value	Units 0	Study Type	Exposure Route	Study Duration	Species	Media	0 Details 0	Source
Effect Level	280333	6	LOAEL	systemic	34.0	mg/kg-day	subchronic	oral	subchronic	rat	-	Study ID: 2	ToxRefD
Misc Hazard Information	280335	6	NOAEL	systemic	3.35	mg/kg-day	subchronic	oral	subchronic	rat	-	Study ID: 2	ToxRefDE
	280337	6	LOAEL	systemic	28.9	mg/kg-day	subchronic	oral	subchronic	dog		Study ID: 2	ToxRefD
Screening Level				or group of studies		mg/kg-day	subchronic	oral	subchronic	dog		Study ID: 2	ToxRefD
Uncertainty Factor				eases in frequency opulation and its app		mg/kg-day	chronic	oral	chronic	dog	-	Study ID: 2	ToxRefD
	control.					mg/kg-day	chronic	oral	chronic	dog	-	Study ID: 2	ToxRefD
	280345	6	LOAEL	systemic	33.7	mg/kg-day	chronic	oral	chronic	dog	-	Study ID: 2	ToxRefD
	280347	6	NOAEL	systemic	4.97	mg/kg-day	chronic	oral	chronic	dog	-	Study ID: 2	ToxRefD
	280349	6	LOAEL	systemic	194	mg/kg-day	chronic	oral	chronic	mouse	-	Study ID: 2	ToxRefD
	280351	8	NOAEL	systemic	47.9	mg/kg-day	chronic	oral	chronic	mouse		Study ID: 2	ToxRefD
	280353	8	LOAEL	systemic	25.0	mg/kg-day	chronic	oral	chronic	rat		Study ID: 2	ToxRefD
	280355	6	NOAEL	systemic	3.50	mg/kg-day	chronic	oral	chronic	rat	-	Study ID: 2	ToxRefD
	280357	6	LOAEL	systemic	25.0	mg/kg-day	chronic	oral	chronic	rat	-	Study ID: 2	ToxRefD

Toxicity Values Panel

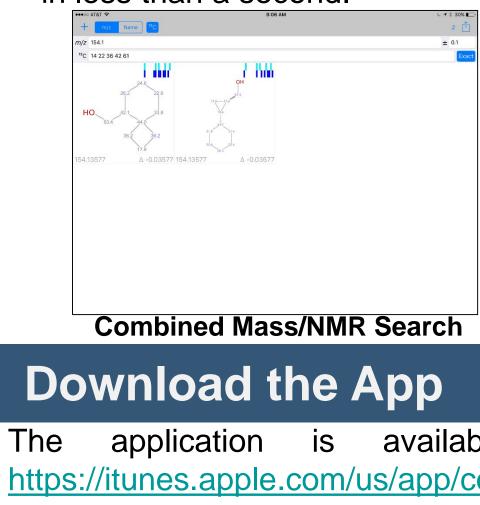
Literature searching using integration to a series of online can be performed resources using the CASRN and chemical This includes Google name. Scholar, PubMed and PubChem

The SDF file associated with the dashboard was reformatted for display in the CompTox Mobile app for iOS: both iPhone and iPad.

The chemical content contained within the application can be searched by Chemical Abstract (CASRN) and by Number chemical name substring. The hit list changes in real time as the name is entered,

T	m/z
m/z	300.156
¹³ C	Example
42978	3-66-5
2	a t
	15729
/ 5410	-21-2
8	17-0 1708- 0 7
300.	15729
56275	5-48-0
u, n	

A user can also search based on a list of carbon-13 NMR chemical shifts that performs a fast search based on looking for similar shifts within a 3 ppm binning window. The search is optimized and can search almost 100 million records in less than a second.



Innovative Research for a Sustainable Future

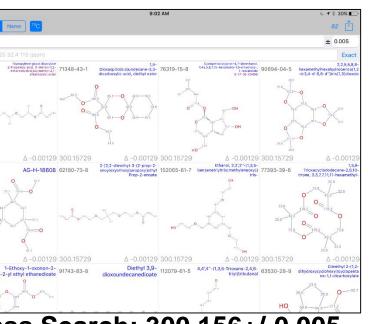
ACS Meeting, San Francisco April 2-6, 2017



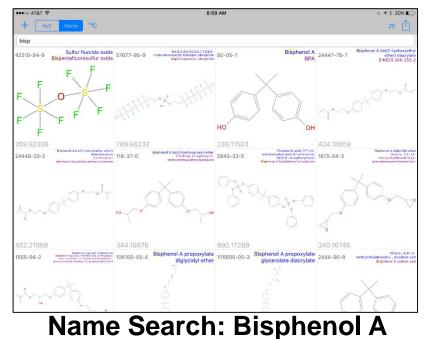
ORCID: 0000-0002-2668-4821

Antony Williams I williams.antony@epa.gov I 919-541-1033

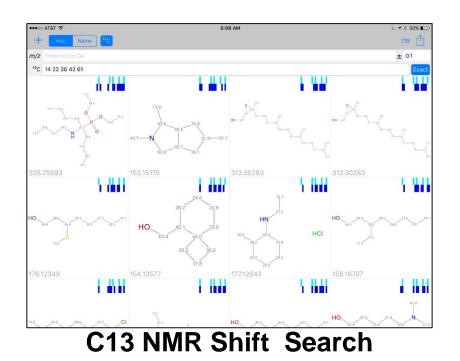
The CompTox iOS Mobile App



iss Search: 300.156+/-0.005



Since each chemical has an accurate "monoisotopic mass" associated searching by mass +/error always for quickly an filtering down to а set of browsing. The chemicals for mass is for the neutral compound.



Combining a mass filter with a list of chemical shifts can result in suggested structures that can then be analyzed and confirmed further with additional analytical spectroscopy studies.

Store available the App at: on https://itunes.apple.com/us/app/comptox-mobile/id1179517689.

This presentation does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.