Mobilizing EPA's CompTox Chemistry Dashboard Data on Mobile Devices

[1]Antony Williams*, [2]Kirill Blinov, [3]Andrew McEachran and [1]Chris Grulke

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

ORCID: 0000-0002-2668-4821
Antony Williams I williams.antony@epa.gov I 919-541-1033

ACS Meeting,
San Francisco
April 2-6, 2017

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.

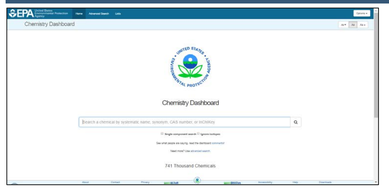
Abstract

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.

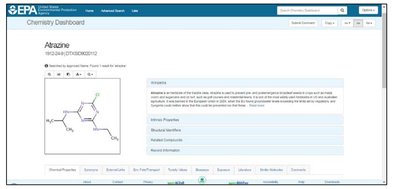
The CompTox Chemistry Dashboard



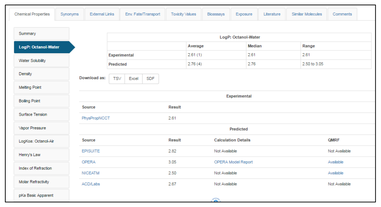
Dashboard Entry Page

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

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.



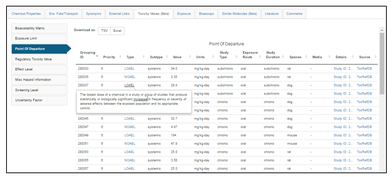
Chemical Record Page: Atrazine



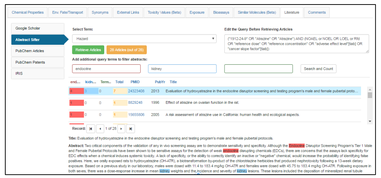
Chemical Properties Panel

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.

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 files.



Toxicity Values Panel



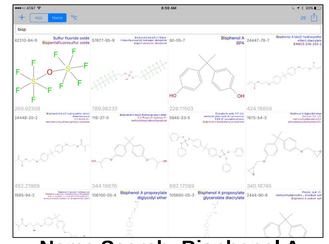
Literature: Pubmed Abstract Sifter

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

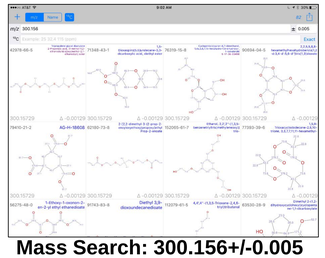
The CompTox iOS Mobile App

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 Number (CASRN) and by chemical name substring. The hit list changes in real time as the name is entered,



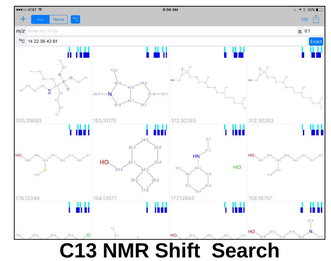
Name Search: Bisphenol A



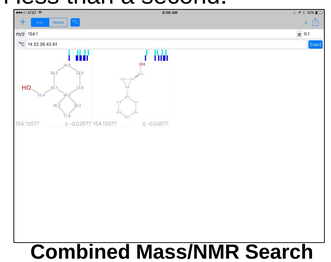
Mass Search: 300.156+/-0.005

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

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.



C13 NMR Shift Search



Combined Mass/NMR Search

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.

Download the App

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