POSTER in ENVR: Applications of Cheminformatics & Computational Chemistry in Environmental Health

Delivering an Informational Hub for Data at the National Center for Computational Toxicology

Authors: Antony Williams, Chris Grulke, Jennifer Smith, Kamel Mansouri, Andrew McEachran, Grace Patlewicz, Jeremy Fitzpatrick, Ann Richard, and Jeff Edwards

The U.S. Environmental Protection Agency (EPA) Computational Toxicology Program integrates advances in biology, chemistry, and computer science to help prioritize chemicals for further research based on potential human health risks. This work involves computational and data driven approaches that integrate chemistry, exposure and biological data. These efforts have produced rich data collections of chemical and bioassay data and resulted in the development of numerous software tools for delivering the data to the community. Much has been learned from the development of a disparate suite of software applications and recent work has focused on the integration of the various data sources into a new software architecture. This is intended to reduce the learning curve for multiple applications, uses curated data sources to improve data integration and recall and ultimately delivers better data in a more consumable form for both the user visiting a website and to computers visiting web services. The resulting application has been used to deliver the CompTox chemistry dashboard. This application provides access to ~720,000 chemicals and associated experimental and predicted properties, high-throughput screening data from the ToxCast project and product and functional use data for thousands of chemicals. Flexible searching supports simple chemical identifier look-up based on chemical name and CAS registry number (CASRN) and structure identification is feasible using mass and formula based searching to support mass spectrometrists performing non-targeted analysis. Batch-based searching provides the user with the ability to look up large collections of chemical data using inputs based on name, CASRN, InChIkeys and other identifiers and to export associated information in a series of standard file formats. The CompTox Chemistry Dashboard architecture and development approach has delivered a foundation architecture on which to build new applications for use within the Agency and for use by the research community. This poster will review the software architecture, data streams and the present capabilities of the CompTox Chemistry Dashboard.
This abstract does not reflect U.S. EPA policy.