Division: Division of Chemical Information

Session: Successful Projects Fueled by Open-Source Tools (Oral)

Facilitating community-based chemical curation by providing an Open Source version of the DSSTox chemical and list registration software that supports the EPA CompTox Chemicals Dashboard

Authors: Christopher Grulke, Antony Williams, Amar Singh, Jeremy Dunne, Jeff Edwards, Ann Richard

The Distributed Structure Searchable Toxicity (DSSTox) database serves as the chemical substance foothold that allows for the collection, integration, and surfacing of data in US-EPA's CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard). Whereas the DSSTox project has always had as a primary focus the sharing of data linked to curated chemical information, the tools built to support our data-structure curation process have not been sufficiently documented and bug-free to provide to the community. As a result, this closed source, open data model has prevented others from being able to directly employ our processes for curating and managing their own sets of chemistry, thus inhibiting a larger community-based curation effort. Hence, we have launched a Chemical Registration Open Source project to provide a public version of the application that has supported DSSTox chemical registration for the past 4 years. This project is crafting a set of publicly available microservices for integrating cheminformatics support functions using Open Source toolkits. These components will provide the underpinning for a user interface to support curator tasks that will handle all aspects of structure normalization, substance mapping, conflict-resolution, substance registration and storage. Database population functions allowing a new user to initialize the database with publicly downloadable DSSTox content will be included. This first step in publishing an Open Source application that will allow users to incorporate and manage the DSSTox database and associated data is expected to be augmented with additional capabilities covering our other data domains being surfaced on the Comptox Chemicals Dashboard. The goal will be to enable efficient integration of data resources to facilitate the management and transfer of environmentally relevant data between stakeholders. This presentation will provide an overview of the Chemical Registration Open Source project and status update for delivering enhanced solutions to the community. *This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.*