**Problem Definition and Goals**

**Problem:** There are few sources of curated data, and integrated workflows, available online to support structure identification using mass spectrometry approaches.

**Goals:** Deliver online access to hundreds of thousands of chemicals of interest to environmental science and computational toxicology. Provide lists of suspect screening chemicals that have been, or could be detected in water via a simple to use web-based interface. Deliver online access to hundreds of thousands of chemicals of interest to environmental science and computational toxicology. Provide lists of suspect screening chemicals that have been, or could be detected in water via a simple to use web-based interface. Deliver online access to hundreds of thousands of chemicals of interest to environmental science and computational toxicology. Provide lists of suspect screening chemicals that have been, or could be detected in water via a simple to use web-based interface. Deliver online access to hundreds of thousands of chemicals of interest to environmental science and computational toxicology. Provide lists of suspect screening chemicals that have been, or could be detected in water via a simple to use web-based interface.

**Abstract**

Non-targeted, targeted and suspect screening, as well as “Known Unknowns” and “Unknown Unknowns” are now common terms in the field of water analysis. While data processing can be highly automated, the identification of chemicals from extracted masses, formulae or fragmentation utilizes reference spectral libraries or metadata. The US EPA CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard) provides access to data for ~675,000 substances, searchable by mass and formula and then ranked using associated metadata. Cheminformatics approaches are also utilized to provide mapped relationships between individual substances and their “MS-Ready” (desalted, non-stereospecific) forms. This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

**MS-Ready Structures for Database Searching**

To facilitate searching, structures are processed into “MS-Ready” forms [4]. This removes salts and stereochemistry and separates mixture components while retaining linkages to the original structures. This enables the form of a structure observed via MS to be related to all variants of a structure.

**Advanced Searching for Chemical Identification Using MS Data**

Advanced searching includes mass or formula. The user can select from a set of potential adducts and for formula searching can perform either MS-ready formula or Exact formula searches. It is also possible to generate matching formulae in the database from mass.

**References**


**Future Work**

- The prediction of mass spectral fragmentation data (LC-MS positive and negative ion mode (10/20/40eV) and GS-MS data allowing for searching of experimental vs predicted data is in testing [4]
- Searching based on structure, substructure and similarity, including filtering based on mass, formula and presence/absence of specific elements is already available as an internal prototype.

**Acknowledgements**

The authors thank the chemical curation team for their rigorous work and the software development team for the development of the dashboard.