A Pilot Database of UVCB Tier 0 Information

Erin Maloney[1], Marc Fernandez[2], Lilly Cesh[2], Antony Williams[3], Charles Lowe[3], Julie Krzykwa[4] and Sandrine Déglin[4]

[1] Shell Global Solutions., [2]Environment and Climate Change Canada, [3]US Environmental Protection Agency, [4]Health and Environmental Sciences Institute

Issue:

Tier 0 information* is important for the prioritization and preliminary assessment of UVCBs. However, to date there is not a public-facing database that contains tier 0 data for UVCBs. In addition, the data available for UVCBs in online large chemical property databases and environmental risk profiling may be based on incorrect or overly simplified substance characterization.

Goal:

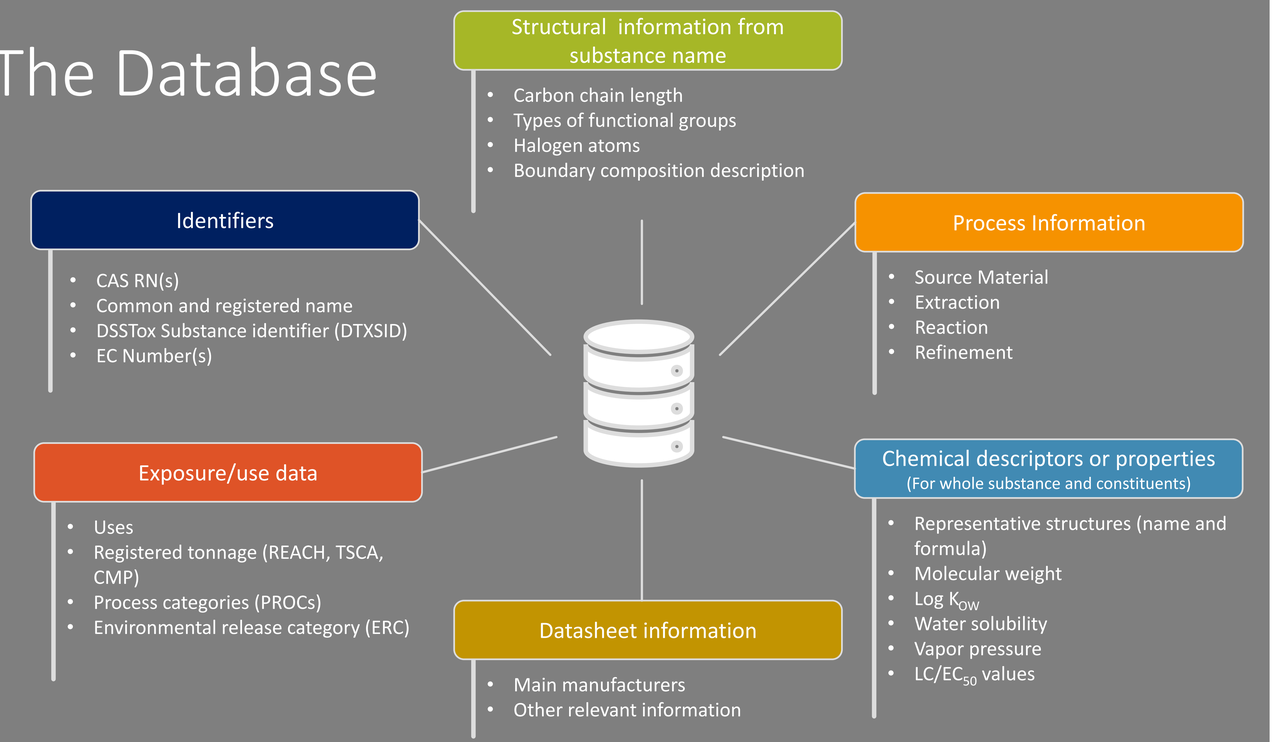
Provide a consistent source of undisputed information regarding commercially relevant UVCBs to help characterize a UVCB substance for the purpose of prioritization and, if required, risk assessment, but also read-across and the filling of data gaps.

Proposal:

Develop a pilot database of Tier 0 information for 109 UVCBs selected based on representation of the categories of significance to regulators in North America, and expanded to include a wide range of chemistries presenting challenges for testing and risk assessment.

*Tier 0 information is data that includes basic compositional, structural, exposure and hazard information, collected from simple analyses (e.g., Elemental analysis, HPLC, GC), substance specifications, QA/QC data, starting material and process information.

The Database



The compiled information is publicly available and may sometimes be based on incorrect, overly simplified, or incomplete substance characterization.

Workflow Approaches and Pilot Database Creation

- Routine workflows were developed for each data type that involved batch searching of publicly available databases with various chemical identifiers (commencing with the EPA's DSSTox Substance Identifier (DTXSID) and synonyms so such routines could potentially be automated
- Manual curation steps will be added following the automated search to correct the data, discard erroneous information, and flag uncertainty with an emphasis on identity and structural information upon which all other information may be based
- The initial pilot database was developed by the sub-working group of HESI's UVCB Committee for approximately 100 UVCBs of regulatory significance. The relevant data were compiled in an excel document which was shared via SharePoint.

Obtaining buy-in from industrial and regulatory partners will be key to facilitate the data collection process, and ultimately, realize the full potential of such a repository of data.

Summary of the availability of key parameters in the database

Parameters	Boundary description	Types of Functional groups	Process information	Component based information	Use Volume	Use Category
Percentage available (n = 100)	14%	84-100%	4-16%	requires advances in cheminformatics approaches	28-40%	76%

Data Gathering Challenges and Summary

- Although the information relevant to this repository is not particularly sensitive and should be publicly available, key data are not readily accessible, meaning that the mining process can be difficult and time consuming, leading to incomplete entries.
- The time necessary for data gathering was one of the biggest hurdles. For example, finding whole substance structural information for an entry can involve searching the US-EPA CompTox Chemicals Dashboard, CAS' Common Chemistry database, PubChem, ECHA dossiers, and broader internet searches.

For information about this project or to become involved, please contact Sandrine Deglin (sdeglin@hesiglobal.org) or Julie Krzykwa (jkrzykwa@hesiglobal.org). The views expressed in this poster are those of the authors and do not necessarily reflect the views or policies of their respective institutions.