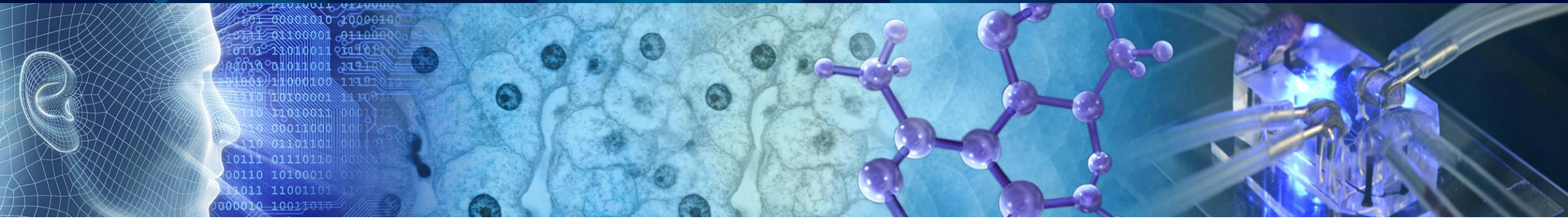
National Institute of Environmental Health Sciences
Division of Translational Toxicology



Curating Chemical Use Categories and Exposure Predictions to Inform Chemical Assessment

Victoria Hull, M.S.

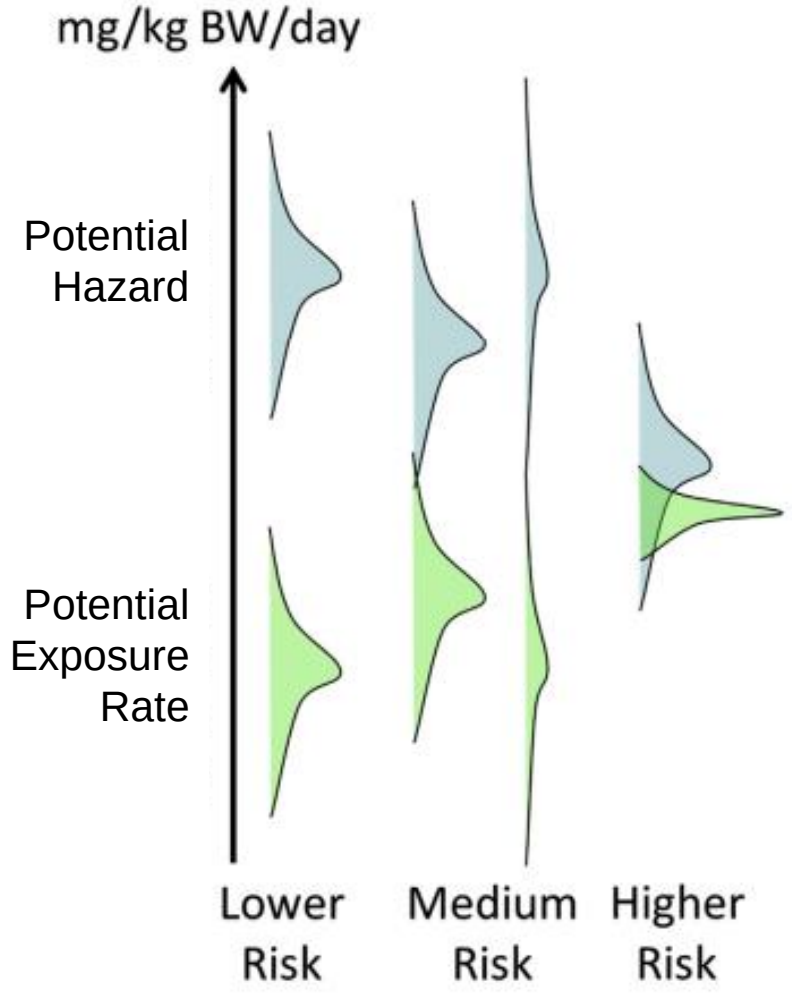
Inotiv, Inc., contractor supporting the NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

ASCCT-ESTIV Webinar Series
Webinar 1, February 2024

Disclaimer: Inotiv staff provide technical support for NICEATM, but do not represent NIEHS, NTP, or the official positions of any federal agency.

Background

- Understanding how human populations interact with and are exposed to chemical sources is essential for contextualizing chemical hazard and understanding chemical risk
  - Many chemicals lack measured estimates of human exposure
- High-throughput exposure simulations and structure-based chemical use models can help inform exposure scenarios for data-poor chemicals
  - The results of high-throughput methods can be difficult to navigate, especially for those unfamiliar with computational methods



Adapted from *Wambaugh et al. 2019, Current Opinion in Toxicology*

National Institute of Environmental Health Sciences
Division of Translational Toxicology

ICE: The Integrated Chemical Environment

- In 2023, we integrated exposure predictions from the Environmental Protection Agency's (EPA's) Systematic Empirical Evaluation of Models (SEEM3) and functional use categories from EPA's Chemical and Product Database (CPDat) into ICE

Integrated Chemical Environment

https://ice.ntp.niehs.nih.gov/

EPA's SEEM3 Model

- SEEM3 was developed by the ExpoCast group at EPA's Center for Computational Toxicology and Exposure

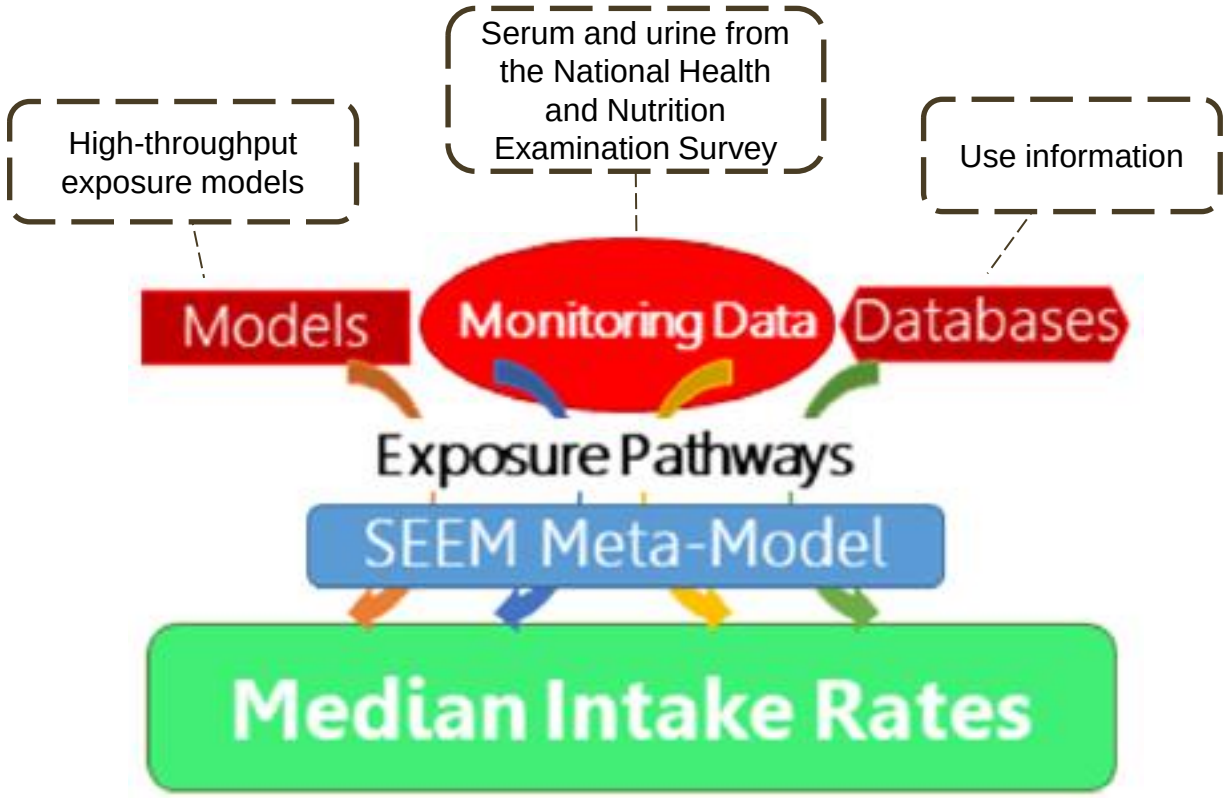
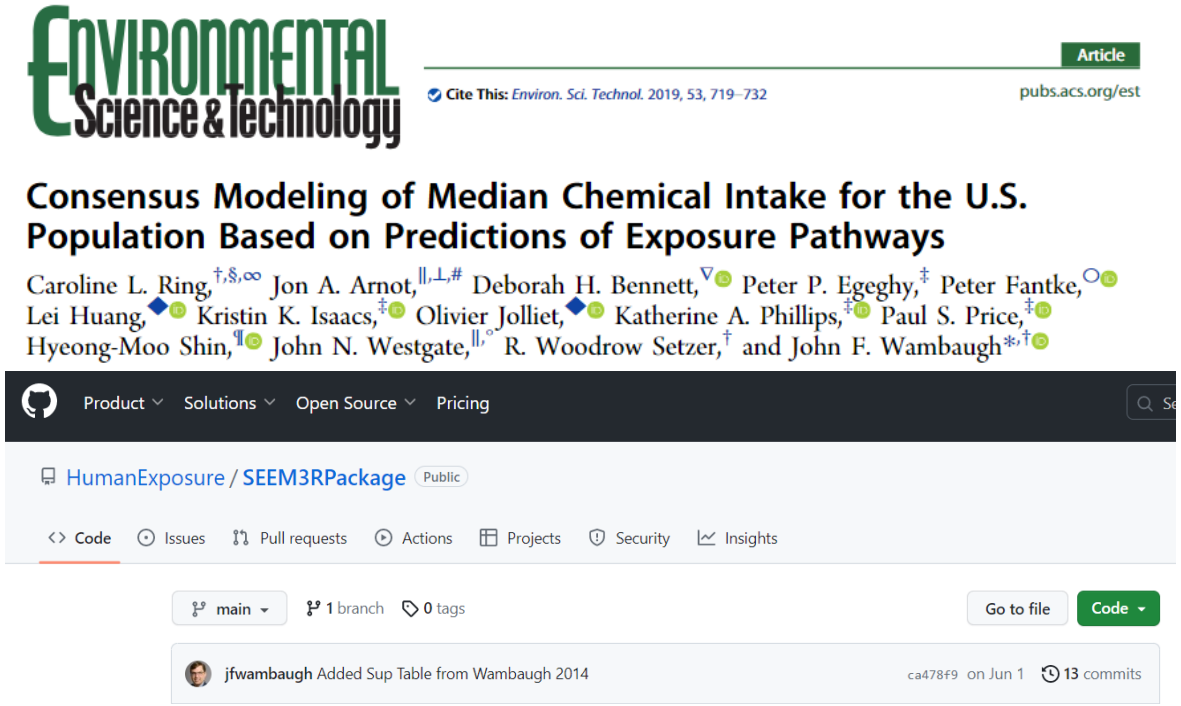


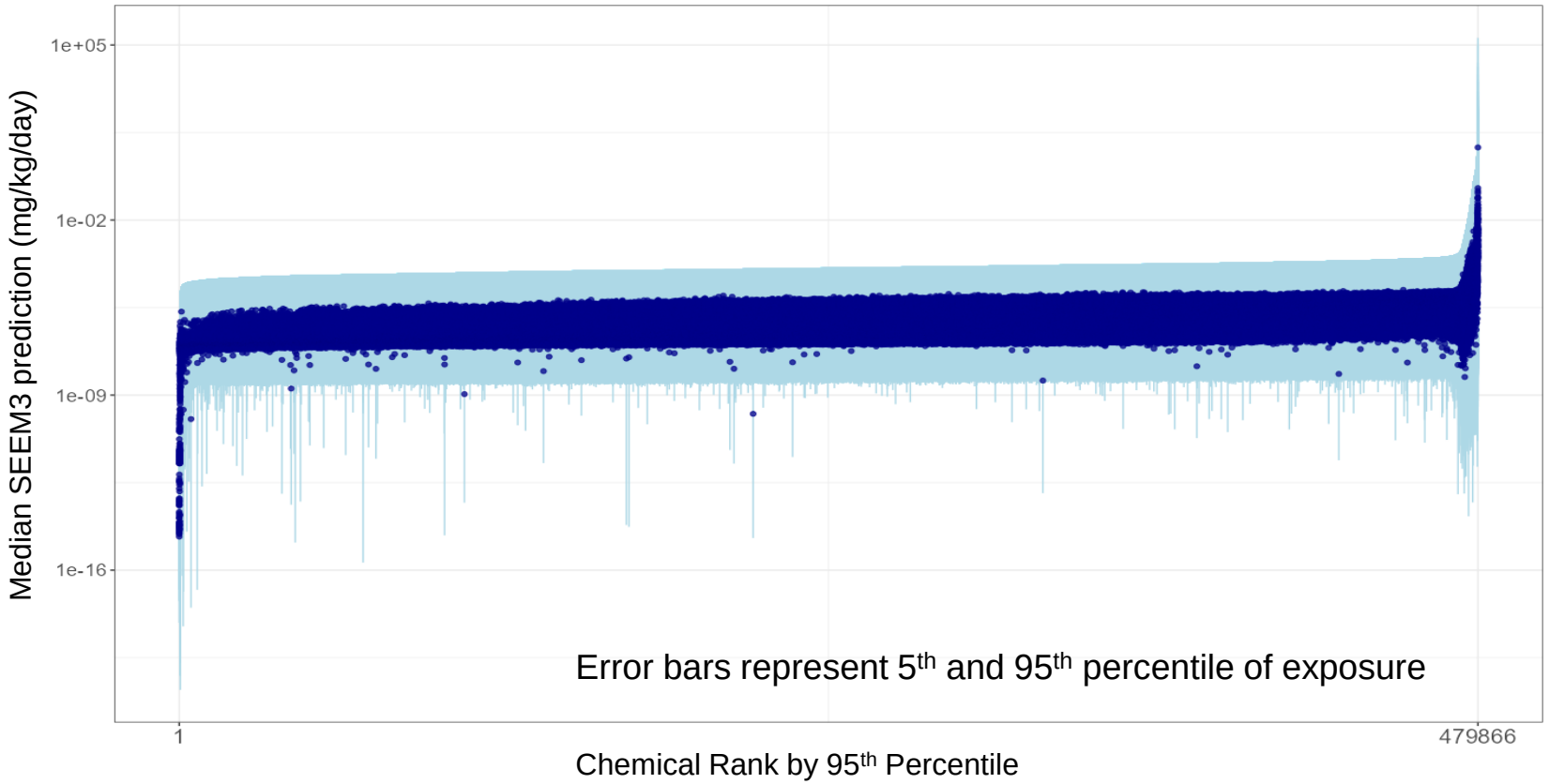
Figure adapted from *Ring et al. 2019 Env. Sci. Tech.*



https://github.com/HumanExposure/SEEM3RPackage

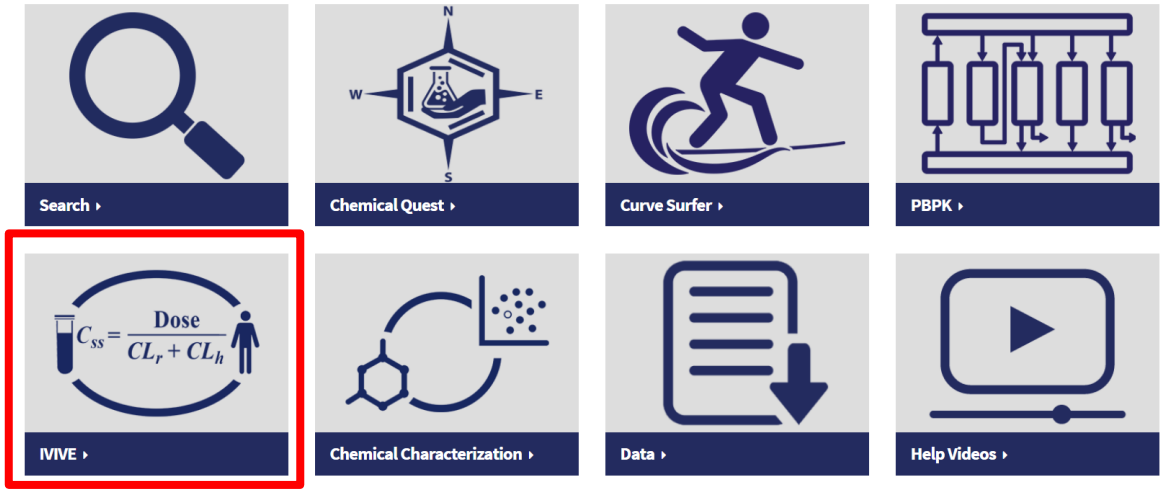
ICE Exposure: Curation and Summary

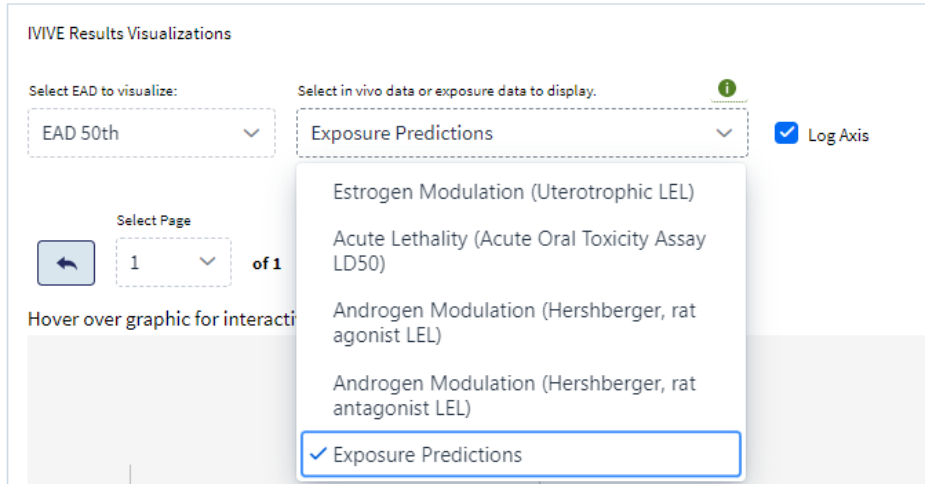
- We pulled SEEM3 predictions for over 600,000 chemicals from EPA's github page (https://github.com/HumanExposure/SEEM3RPackage) in November 2022
  - 5th, 50th, and 95th percentile of exposure in mg/kg/day
  - Limited dataset to **~480,000 chemicals** that were within the model's domain of applicability

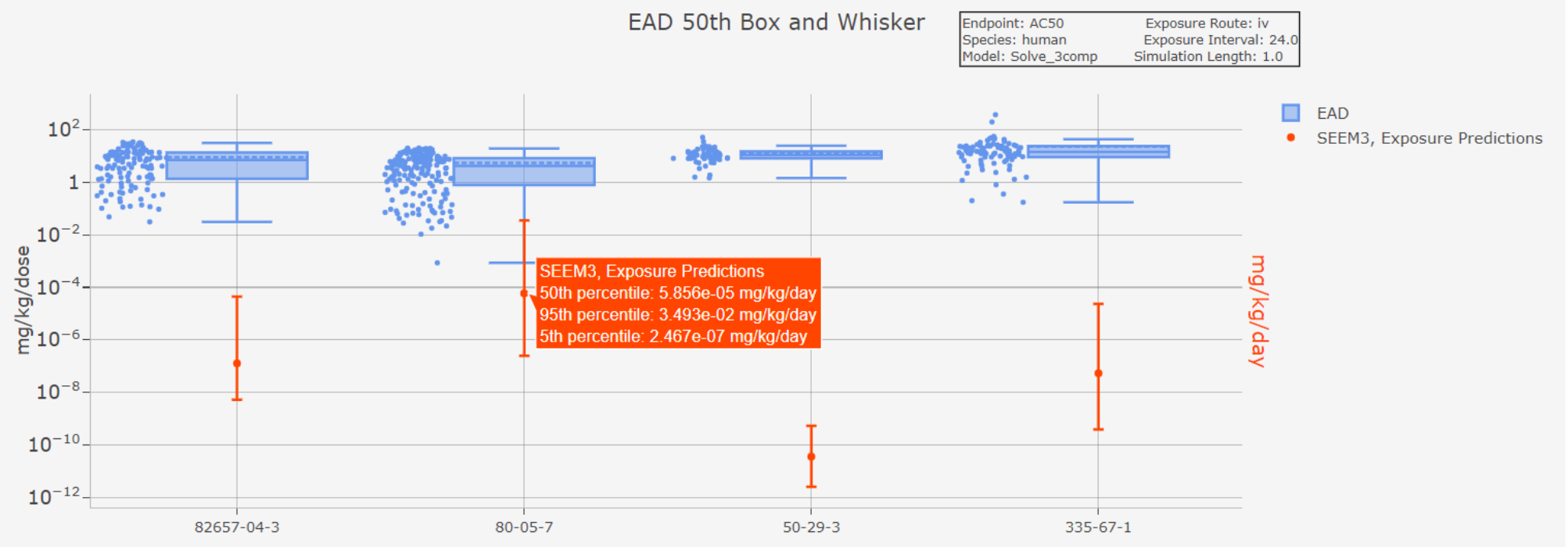


National Institute of Environmental Health Sciences
Division of Translational Toxicology

Exposure in ICE: In Vitro to In Vivo Extrapolation (IVIVE)







SEEM3 Exposure Pathways

- Exposure pathway is often not known
- SEEM3 predicts pathway of exposure through chemical structure-based and property-based machine learning models
  - Consumer
  - Dietary
  - Far-field pesticides
  - Far-field industrial

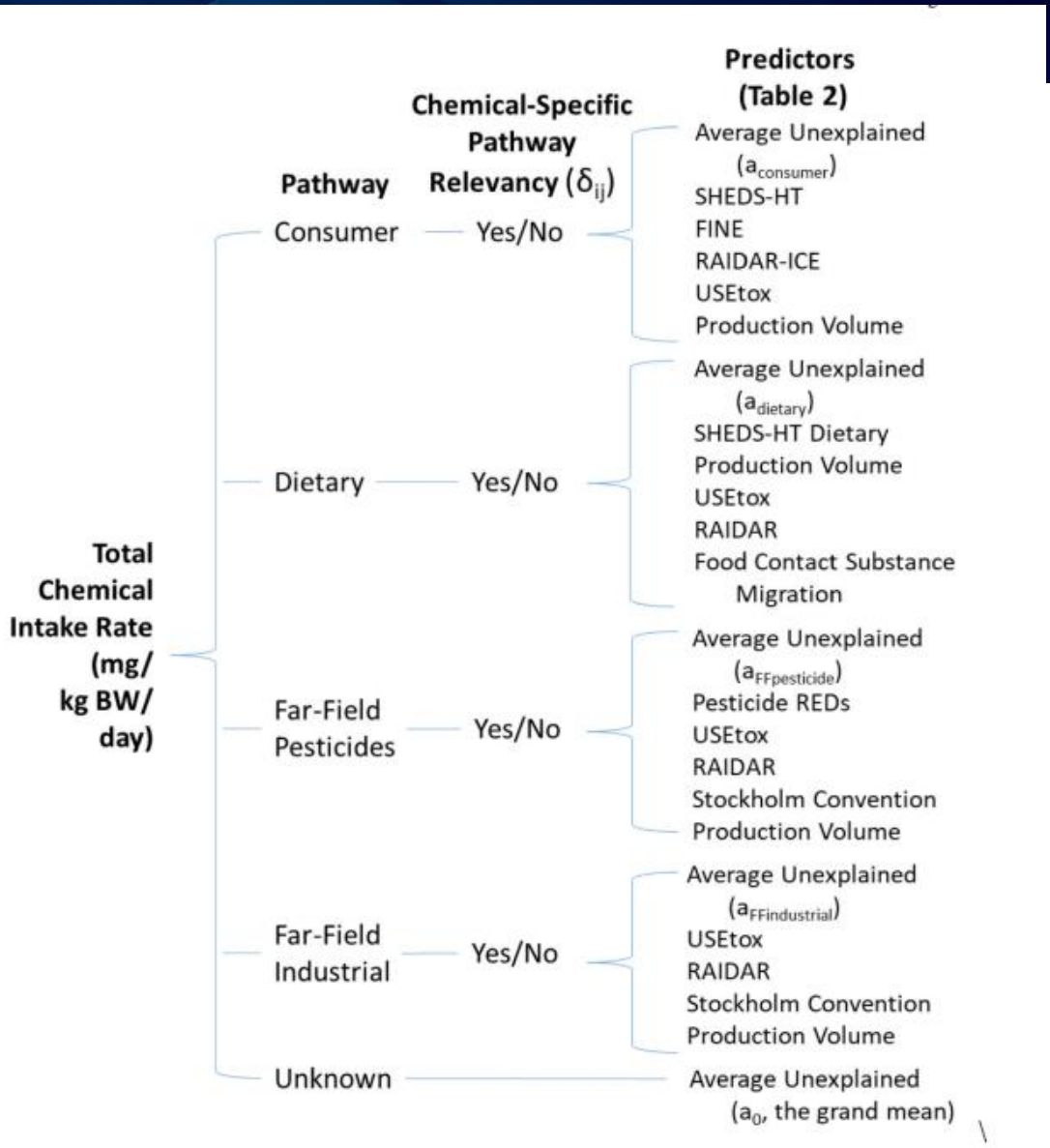
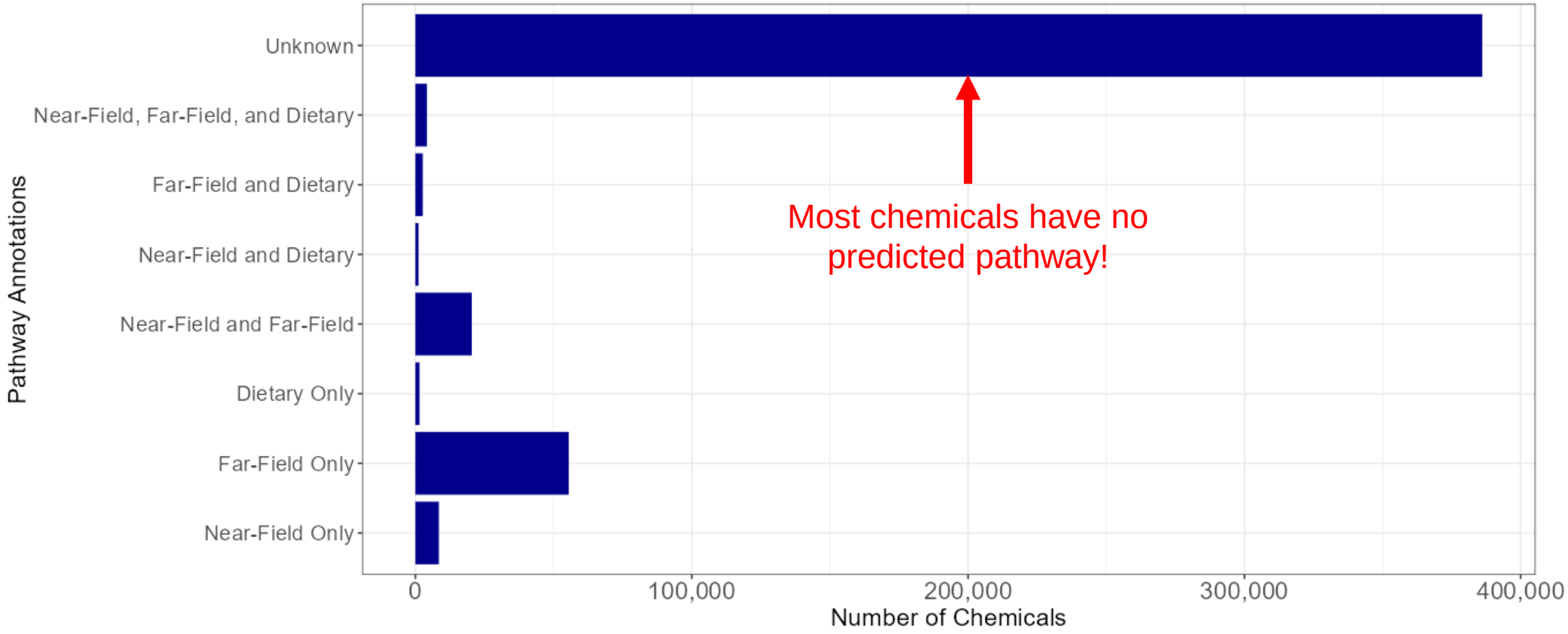


Figure from *Ring et al. 2019 Env. Sci. Tech.*

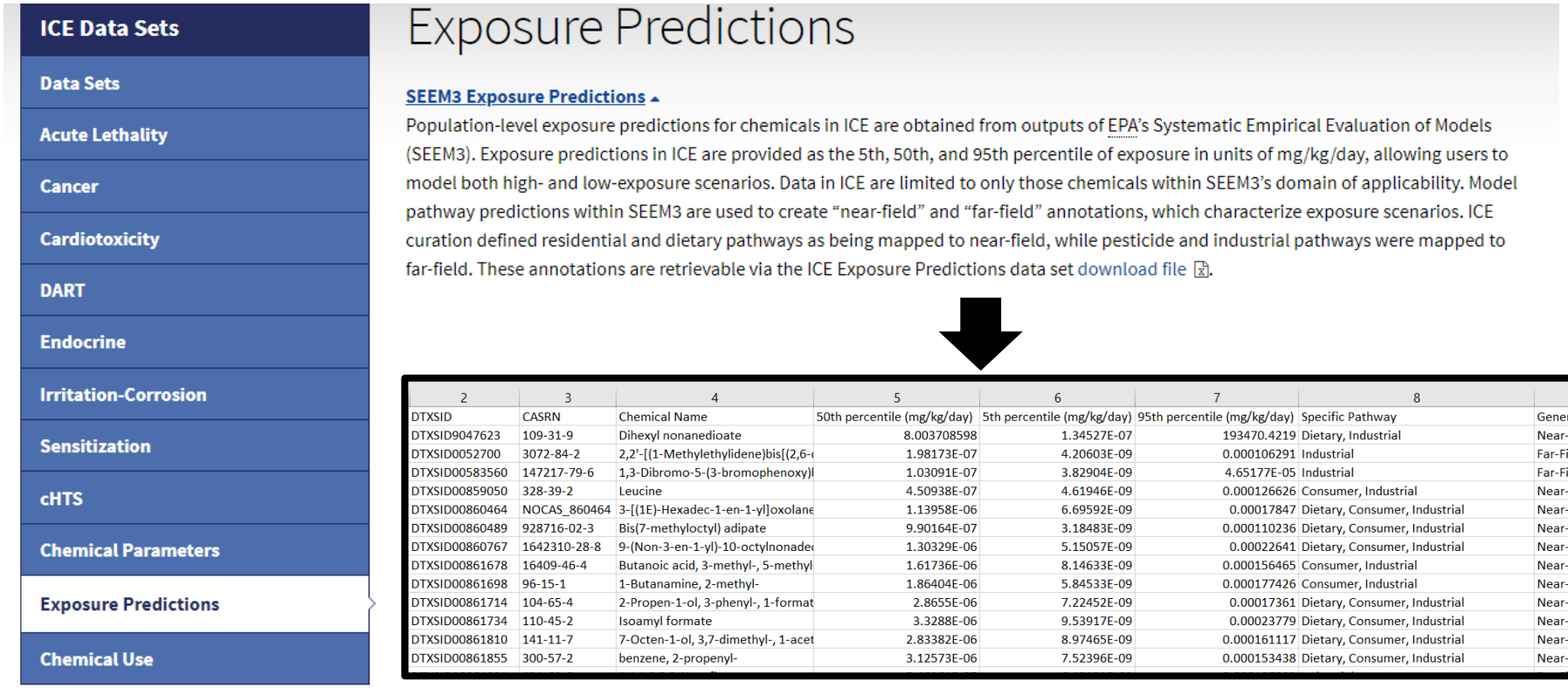
ICE Exposure: Curation and Summary

- For the **~480,000 chemicals** we also pulled the pathway predictions from SEEM3
- We created our near-field and far-field annotations based on the predicted pathway



Direct Access of Exposure Predictions Data Set

- ICE Data Sets page contains downloadable data set with SEEM3 exposure estimates, SEEM3 predicted pathways, and ICE-curated near-field and far-field annotations



2	3	4	5	6	7	8	9
DTXSID	CASRN	Chemical Name	50th percentile (mg/kg/day)	5th percentile (mg/kg/day)	95th percentile (mg/kg/day)	Specific Pathway	General Pathway
DTXSID9047623	109-31-9	Dihexyl nonanedioate	8.003708598	1.34527E-07	193470.4219	Dietary, Industrial	Near-Field and Far-Field
DTXSID0052700	3072-84-2	2,2'-[(1-Methylethylidene)bis[(2,6-(	1.98173E-07	4.20603E-09	0.000106291	Industrial	Far-Field
DTXSID00583560	147217-79-6	1,3-Dibromo-5-(3-bromophenoxy)l	1.03091E-07	3.82904E-09	4.65177E-05	Industrial	Far-Field
DTXSID00859050	328-39-2	Leucine	4.50938E-07	4.61946E-09	0.000126626	Consumer, Industrial	Near-Field and Far-Field
DTXSID00860464	NOCAS_860464	3-[(1E)-Hexadec-1-en-1-yl]oxolane	1.13958E-06	6.69592E-09	0.00017847	Dietary, Consumer, Industrial	Near-Field and Far-Field
DTXSID00860489	928716-02-3	Bis(7-methyloctyl) adipate	9.90164E-07	3.18483E-09	0.000110236	Dietary, Consumer, Industrial	Near-Field and Far-Field
DTXSID00860767	1642310-28-8	9-(Non-3-en-1-yl)-10-octylnonade	1.30329E-06	5.15057E-09	0.00022641	Dietary, Consumer, Industrial	Near-Field and Far-Field
DTXSID00861678	16409-46-4	Butanoic acid, 3-methyl-, 5-methyl	1.61736E-06	8.14633E-09	0.000156465	Consumer, Industrial	Near-Field and Far-Field
DTXSID00861698	96-15-1	1-Butanamine, 2-methyl-	1.86404E-06	5.84533E-09	0.000177426	Consumer, Industrial	Near-Field and Far-Field
DTXSID00861714	104-65-4	2-Propen-1-ol, 3-phenyl-, 1-format	2.8655E-06	7.22452E-09	0.00017361	Dietary, Consumer, Industrial	Near-Field and Far-Field
DTXSID00861734	110-45-2	Isoamyl formate	3.3288E-06	9.53917E-09	0.00023779	Dietary, Consumer, Industrial	Near-Field and Far-Field
DTXSID00861810	141-11-7	7-Octen-1-ol, 3,7-dimethyl-, 1-acet	2.83382E-06	8.97465E-09	0.000161117	Dietary, Consumer, Industrial	Near-Field and Far-Field
DTXSID00861855	300-57-2	benzene, 2-propenyl-	3.12573E-06	7.52396E-09	0.000153438	Dietary, Consumer, Industrial	Near-Field and Far-Field

Download Data from ICE.
https://ice.ntp.niehs.nih.gov/DATASETDESCRIPTION

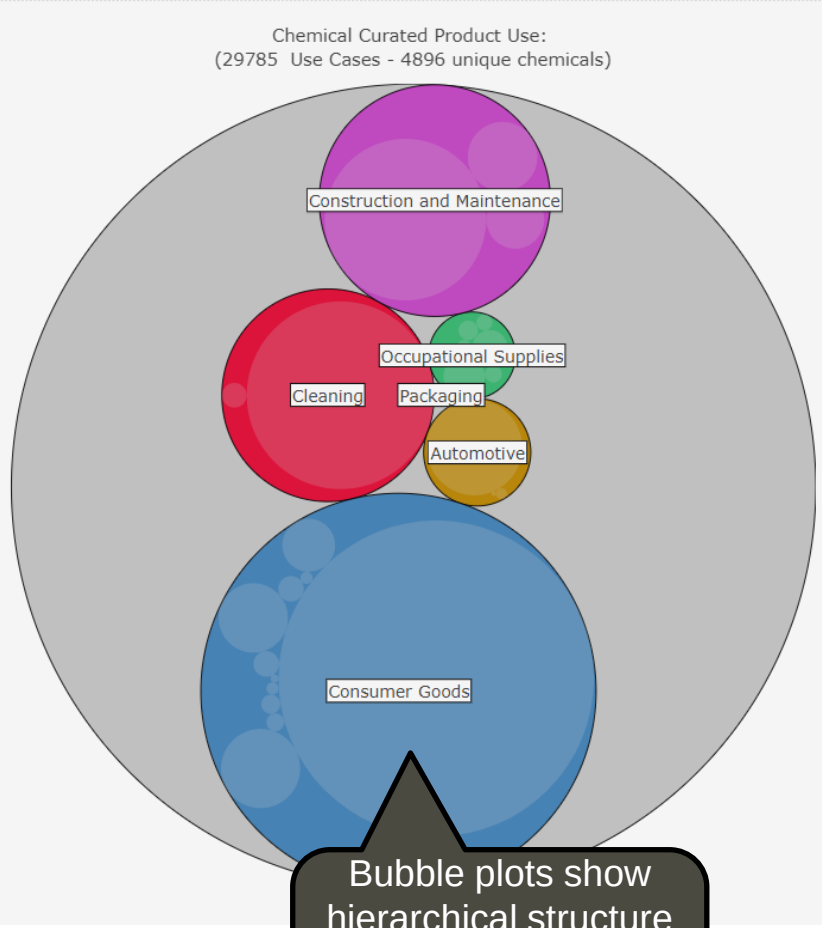
Chemical Use Categories in ICE

- Chemical use categories explain the various ways in which a chemical can be used
  - In what sectors is a chemical used? What role does the chemical play? What products a chemical can be found in? etc.
- Can help inform pathway for exposure models
- In ICE, our chemical use categories are derived from CPDat (v3 2021, Williams et al. 2017)

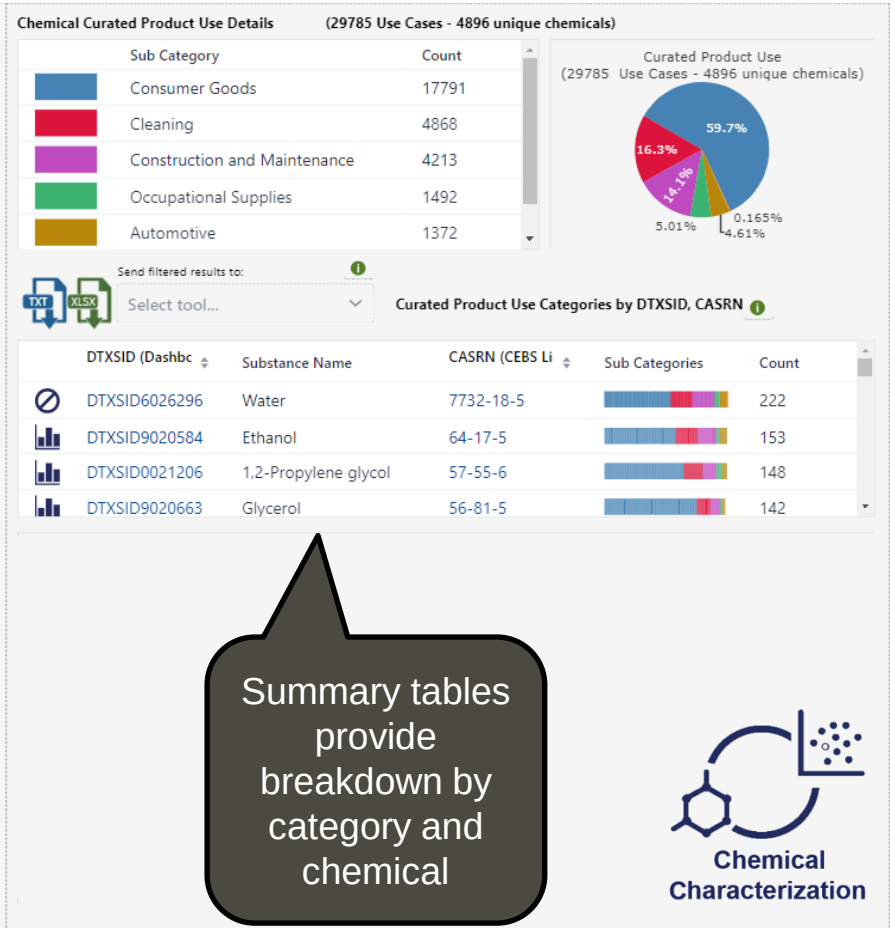
https://www.epa.gov/chemical-research/chemical-and-products-database-cpdat

Curated Product Use Categories

- Derived from EPA's Product Use Categories (*Isaacs et al. 2019, JESEE*)
- Curated product use categories describe the consumer products a chemical may be in
  - Useful in determining chemical composition, exposure frequency, and route of exposure
- Users can visualize and extract these categories from the ICE Chemical Characterization tool
  - Over 300 categories for nearly 5000 chemicals



Bubble plots show hierarchical structure and relative abundance of categories



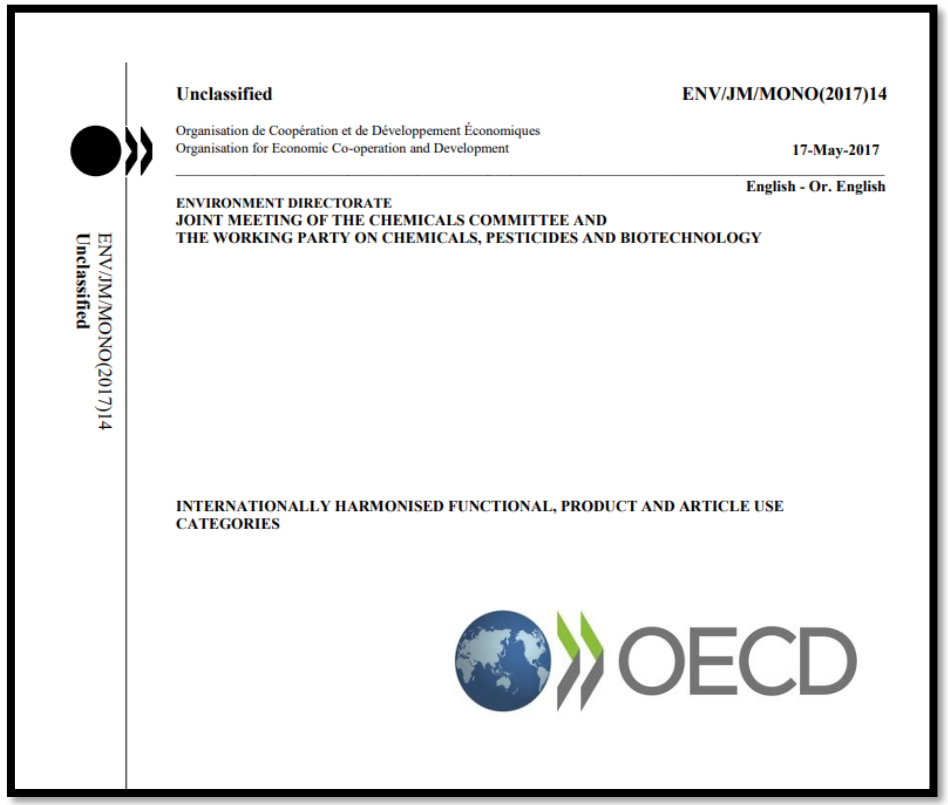
Summary tables provide breakdown by category and chemical

Chemical Characterization

Functional Use Categories

- Functional use is the role a chemical plays within a product. We pull two types of functional use from CPDat:

1. Reported functional use harmonized to Organisation of Economic Co-operation and Development (OECD) categories



Unclassified ENV/JM/MONO(2017)14

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development

17-May-2017

English - Or. English

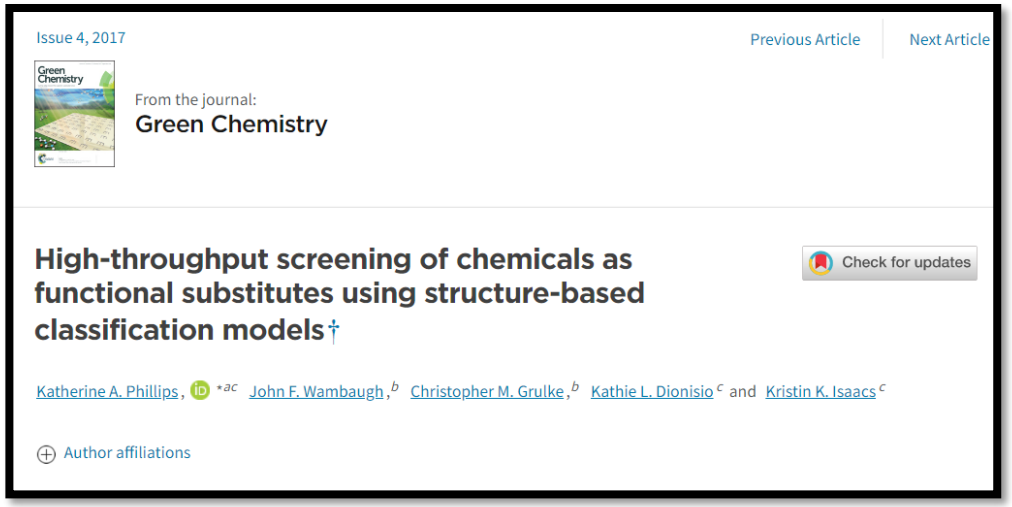
ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND
THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

ENV/JM/MONO(2017)14
Unclassified

INTERNATIONALLY HARMONISED FUNCTIONAL, PRODUCT AND ARTICLE USE CATEGORIES

OECD

2. Predicted functional use is predicted from Quantitative Structure Use Relationship (QSUR) models (Phillips et al. 2018)



Issue 4, 2017 Previous Article Next Article

From the journal:
Green Chemistry

High-throughput screening of chemicals as functional substitutes using structure-based classification models †

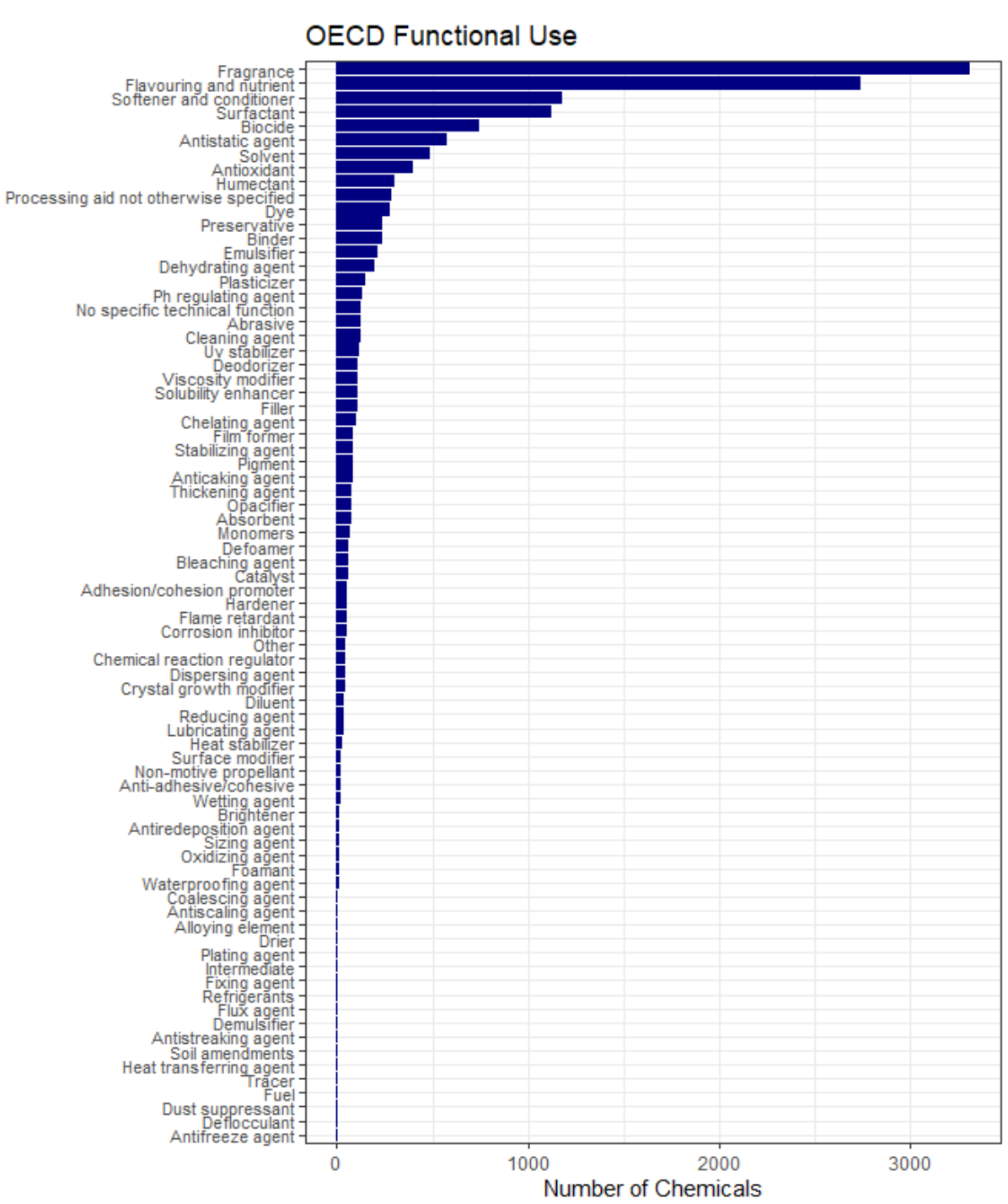
Check for updates

Katherine A. Phillips, *ac John F. Wambaugh,b Christopher M. Grulke,b Kathie L. Dionisio c and Kristin K. Isaacs c

Author affiliations

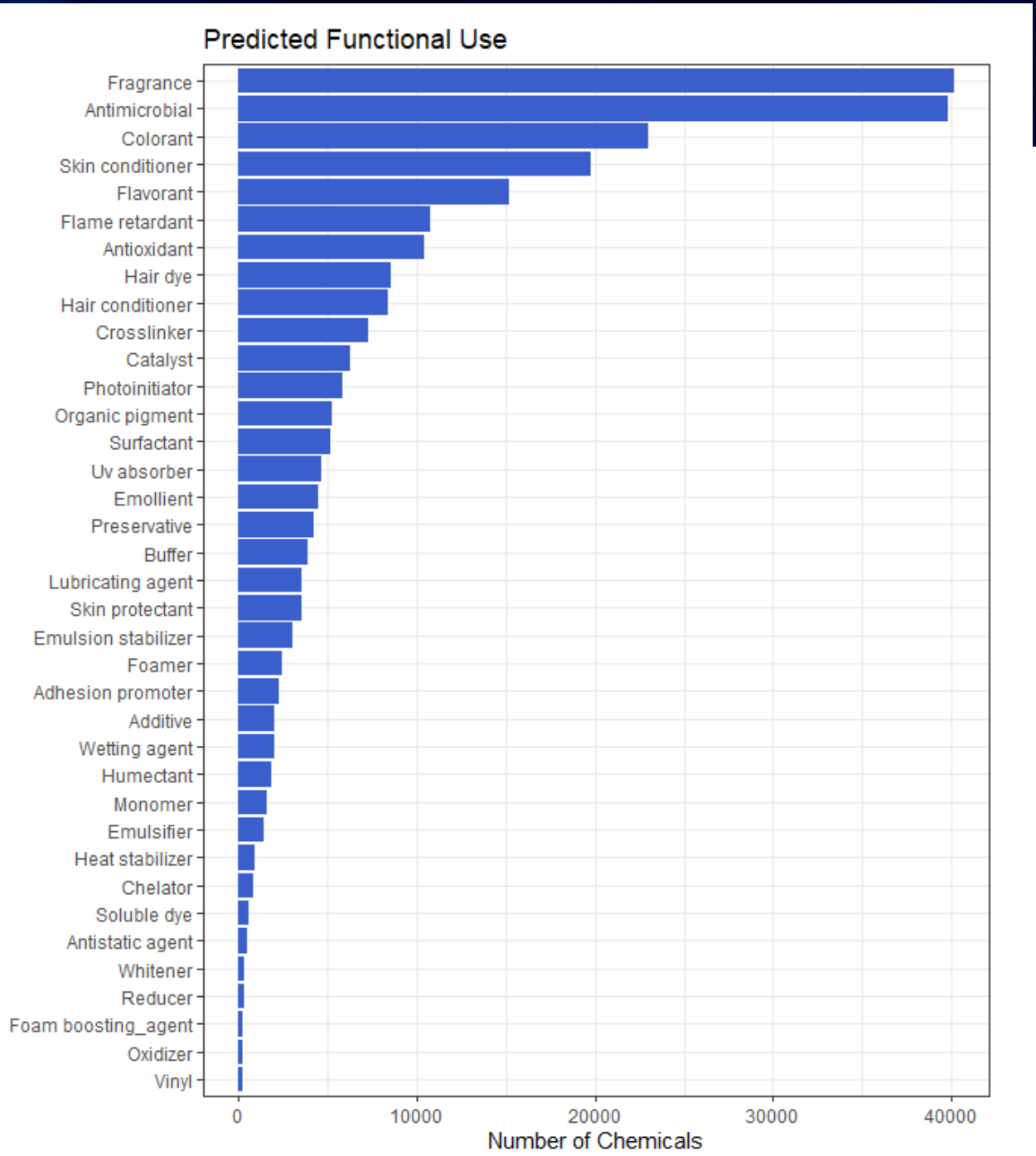
OECD Functional Use

- Within CPDat, reported functional use for ~7,500 chemicals are harmonized to uses established by the OECD
- We harmonized reported use to OECD use for ~2,000 additional chemicals
- Functional use dataset in ICE has 77 OECD uses for 9,395 chemicals



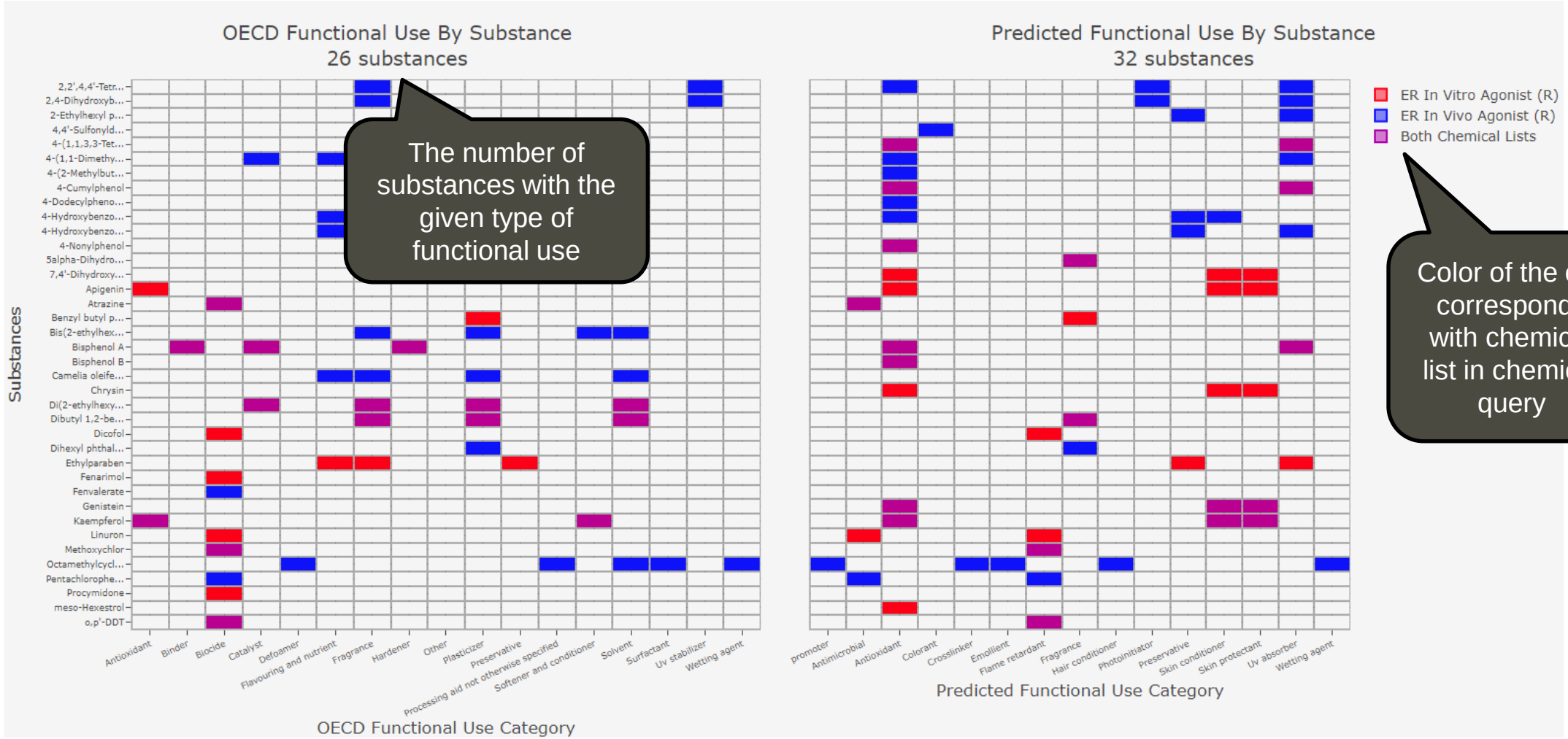
Predicted Functional Use

- Within CPDat, functional use predictions are already limited to chemicals within the QSUR model's applicability domain
  - We curate the data further by limiting the probability of prediction to ≥ 80% to ensure high confidence in results
- Functional use dataset in ICE has 37 predicted functional uses for 192,438 chemicals



Visualizations of Functional Use Data

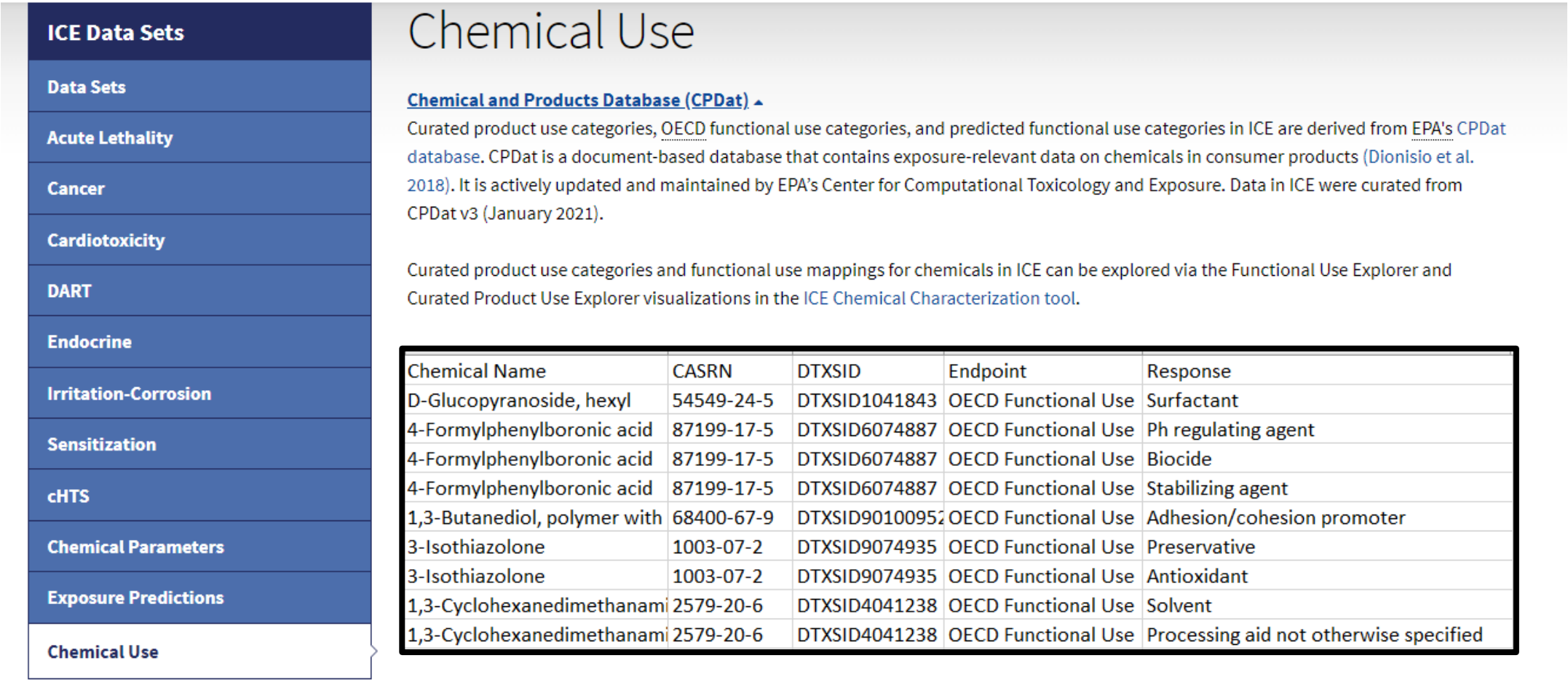
- Accessible through the ICE Chemical Characterization tool



Chemical Characterization

Direct Access to Functional Use Data Set

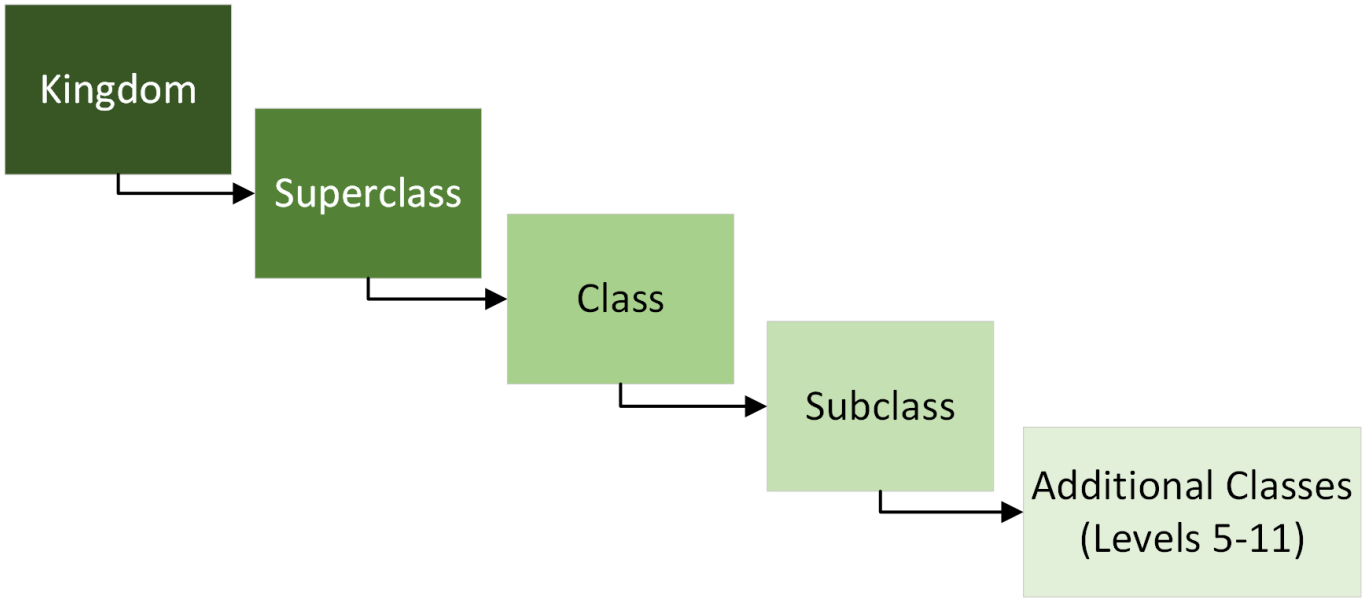
- Accessible through the ICE Datasets Page



Chemical Name	CASRN	DTXSID	Endpoint	Response
D-Glucopyranoside, hexyl	54549-24-5	DTXSID1041843	OECD Functional Use	Surfactant
4-Formylphenylboronic acid	87199-17-5	DTXSID6074887	OECD Functional Use	Ph regulating agent
4-Formylphenylboronic acid	87199-17-5	DTXSID6074887	OECD Functional Use	Biocide
4-Formylphenylboronic acid	87199-17-5	DTXSID6074887	OECD Functional Use	Stabilizing agent
1,3-Butanediol, polymer with	68400-67-9	DTXSID90100952	OECD Functional Use	Adhesion/cohesion promoter
3-Isothiazolone	1003-07-2	DTXSID9074935	OECD Functional Use	Preservative
3-Isothiazolone	1003-07-2	DTXSID9074935	OECD Functional Use	Antioxidant
1,3-Cyclohexanedimethanami	2579-20-6	DTXSID4041238	OECD Functional Use	Solvent
1,3-Cyclohexanedimethanami	2579-20-6	DTXSID4041238	OECD Functional Use	Processing aid not otherwise specified

Future Additions: ClassyFire Chemical Taxonomies

- ClassyFire (Djoumbou et al. 2016, Wishart Research Group) is an automated, structure-based, hierarchical chemical taxonomy with up to 11 levels of classification and 4,285 classifications across all levels of the hierarchy



Djoumbou Feunang *et al. J Cheminform* (2016) 8:61
DOI 10.1186/s13321-016-0174-y

Journal of Cheminformatics

SOFTWARE Open Access

ClassyFire: automated chemical classification with a comprehensive, computable taxonomy

Yannick Djoumbou Feunang[1], Roman Eisner[2], Craig Knox[3], Leonid Chepelev[5], Janna Hastings[6], Gareth Owen[6], Eoin Fahy[7], Christoph Steinbeck[6], Shankar Subramanian[7], Evan Bolton[8], Russell Greiner[3,9] and David S. Wishart[1,3,4,10*]

Future Additions: ClassyFire Chemical Taxonomies

- ClassyFire (Djoumbou et al. 2016, Wishart Research Group) is an automated, structure-based, hierarchical chemical taxonomy with up to 11 levels of classification and 4,285 classifications across all levels of the hierarchy
- One potential use case for these classifications is to link certain chemical use categories within ICE to ClassyFire taxonomies
  - Can help identify chemical classes and structures that are most abundant in their chemical query
- Identifying chemical groups of interest can help focus follow-up investigations or aid in the selection of alternative chemicals

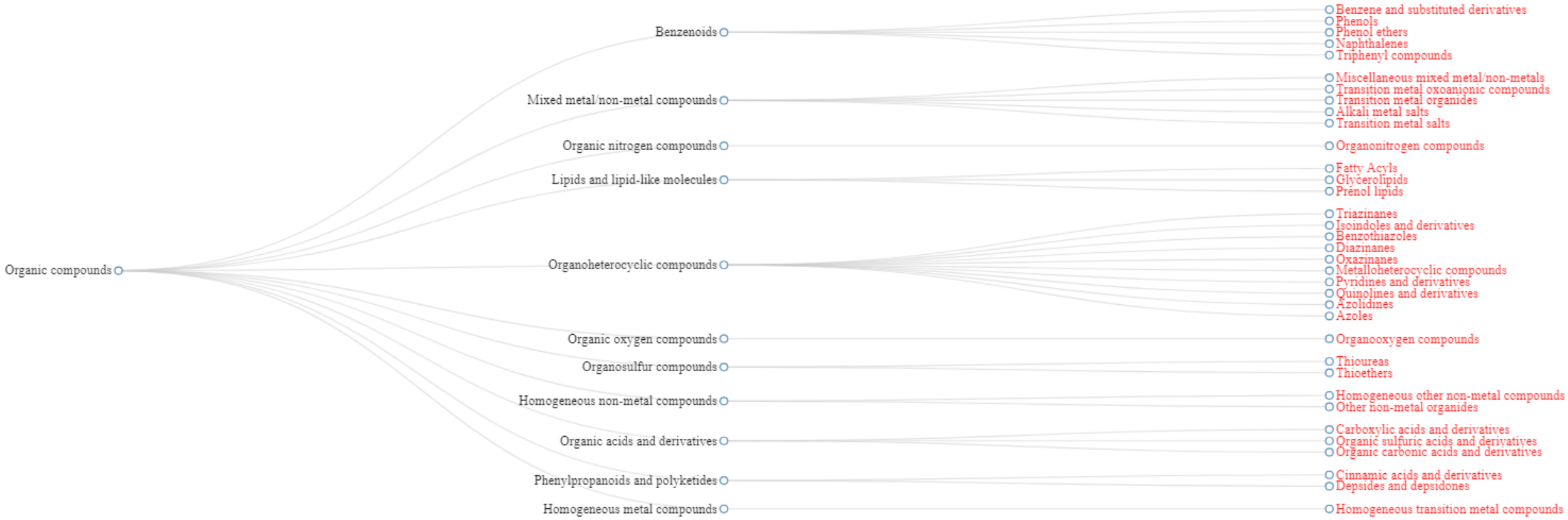
ClassyFire Case Study

- We conducted a case study of 100 chemicals with the OECD functional use of biocides to demonstrate a potential ClassyFire output, which mapped to 88 chemicals within the ClassyFire Kingdom "Organic Compounds"

Superclass	Class	# Chemicals
Benzenoids	Benzene and substituted derivatives	27
	Naphthalenes	1
	Phenol ethers	4
	Phenols	4
	Triphenyl compounds	1
Lipids and lipid-like molecules	Fatty Acyls	2
	Glycerolipids	1
	Prenol lipids	3
Organic acids and derivatives	Carboxylic acids and derivatives	2
	Organic carbonic acids and derivatives	2
	Organic sulfuric acids and derivatives	1
Organic nitrogen compounds	Organonitrogen compounds	10
Organic oxygen compounds	Organooxygen compounds	9
Organoheterocyclic compounds	Azoles	1
	Azolidines	3
	Benzothiazoles	1
	Diazinanes	1
	Isoindoles and derivatives	1
	Metalloheterocyclic compounds	1
	Oxazinanes	1
	Pyridines and derivatives	3
	Quinolines and derivatives	3
	Triazinanes	1
Organosulfur compounds	Thioethers	1
	Thioureas	1
Phenylpropanoids and polyketides	Cinnamic acids and derivatives	1
	Depsides and depsidones	1

ClassyFire Case Study

- We conducted a case study of 100 chemicals with the OECD functional use of biocides to demonstrate a potential ClassyFire output, which mapped to 88 chemicals within the ClassyFire Kingdom "Organic Compounds”



Summary and Future Directions

- In our continuing efforts to provide high-confidence, high-quality toxicologically relevant data, we curated exposure predictions from EPA's SEEM3 model and functional use data from EPA's CPDat
  - This data was made publicly available with the release of ICE v4.0 (March 2023) and ICE v4.0.1 (August 2023)
- With the inclusion of these data, users will be able to better explore how human populations may interact with chemicals and their potential levels of exposure
- The addition of these new data types into ICE facilitates the potential addition of new data sources, exposure models, and types of use
  - Occupational exposure models, models with demographic information, other types of use like sector of use, etc.

National Institute of Environmental Health Sciences
Division of Translational Toxicology

Acknowledgments

The NICEATM Group

NIEHS/DTT Contributors

Integrated Chemical Environment

https://ice.ntp.niehs.nih.gov/

Subscribe to NICEATM News
https://ntp.niehs.nih.gov/go/niceatm

My email: Victoria.Hull@inotivco.com
ICE contact email: ICE-support@niehs.nih.gov