United States Environmental Protection Agency

GAP-FILLING METHODS FOR EXPOSURE MODELING

Katherine Phillips

Office of Research and Development
Center for Computational Toxicology and Exposure

DISCLAIMER: The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the United States Environmental Protection Agency.

EXPOSURE PATHWAYS

- Exposure pathways provide structured ways to track exposure from a source to a receptor

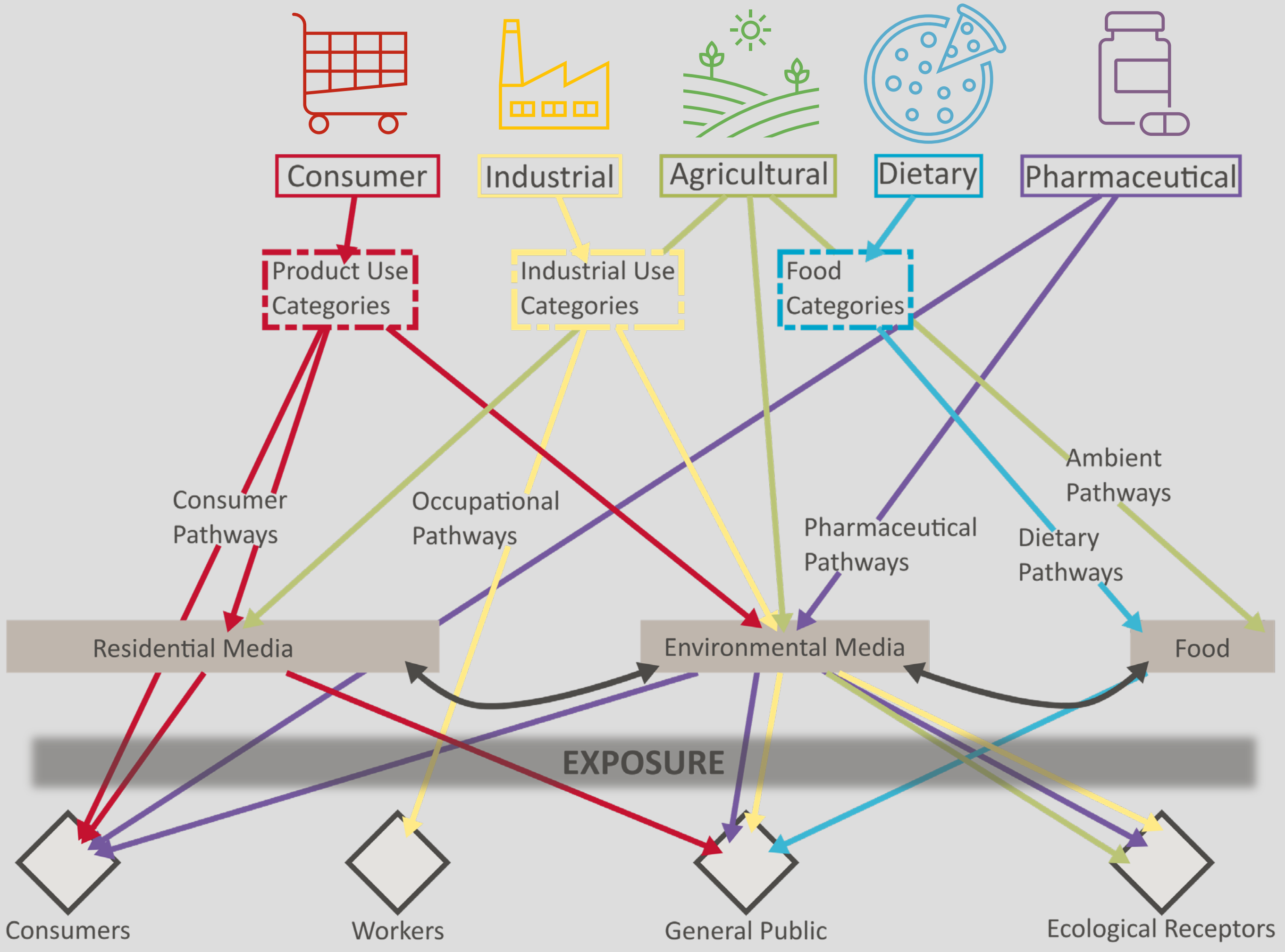


Figure adapted from Kristin Isaacs

EXPOSURE PATHWAYS

- Exposure pathways provide structured ways to track exposure from a source to a receptor
- Along the pathway there are key pieces of information that help provide an estimate of exposure:
  - What product is being used? What chemicals does it contain (at what concentrations)? How much gets used at a time? How frequently is the product used? And on, and on…
  - What media does a chemical reside in based on its source? What is state of matter (liquid, gas, solid) does it have? Do chemicals break down there and become other chemicals? At what rate? In what proportions? How long does a person contact it? How do they interact with the chemical?

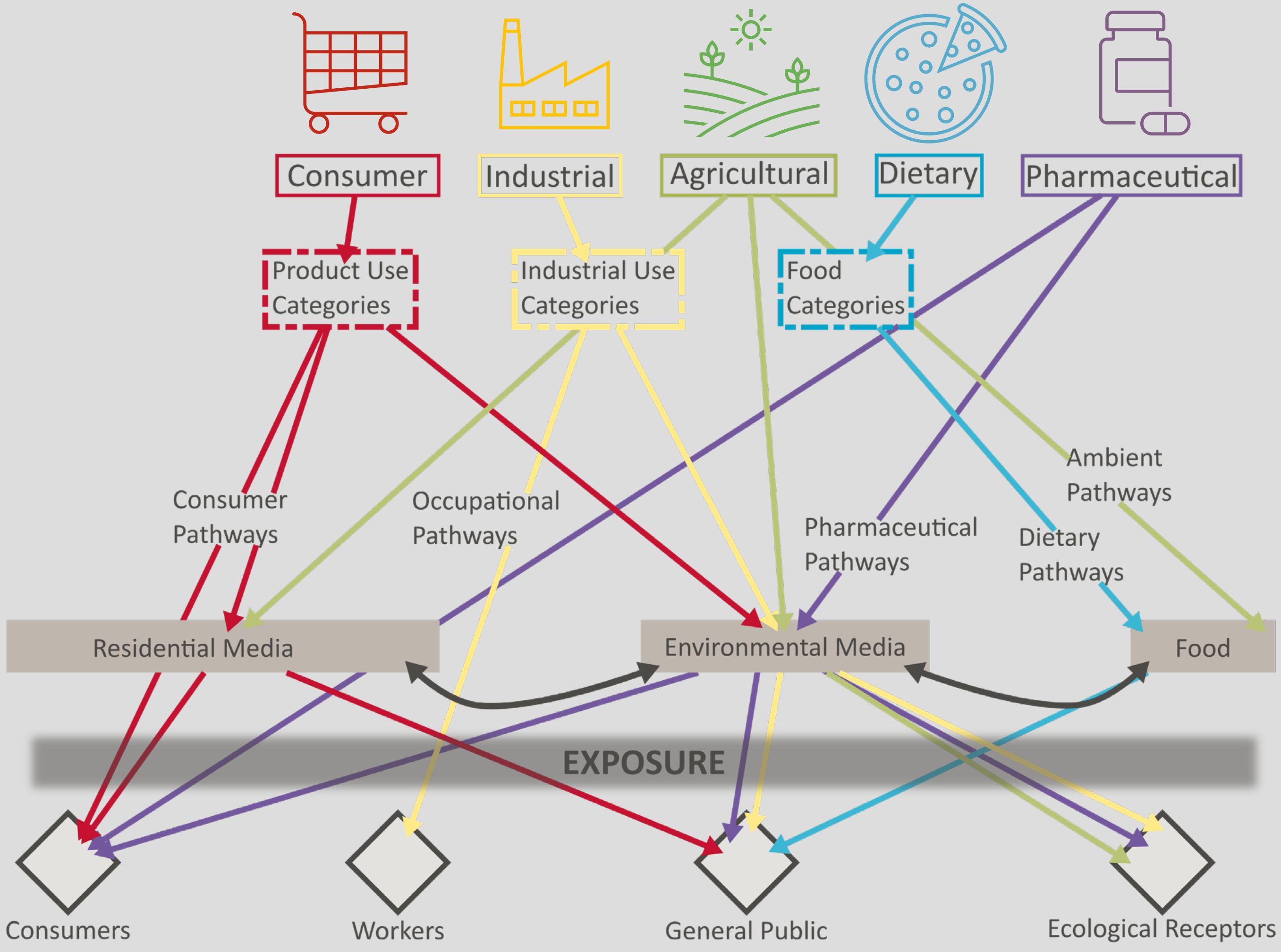


Figure adapted from Kristin Isaacs

United States Environmental Protection Agency

EXPOSURE PATHWAYS

- Exposure pathways provide structured ways to track exposure from a source to a receptor
- Along the pathway there are key pieces of information that help provide an estimate of exposure:
  - What product is being used? What chemicals does it contain (at what concentrations)? How much gets used at a time? How frequently is the product used? And on, and on…
  - What media does a chemical reside in based on its source? What is state of matter (liquid, gas, solid) does it have? Do chemicals break down there and become other chemicals? At what rate? In what proportions? How long does a person contact it? How do they interact with the chemical?
- Maybe you could get all your questions answered for one chemical if you had infinite time and resources, but what if you don't have that?...and you have thousands of chemicals?

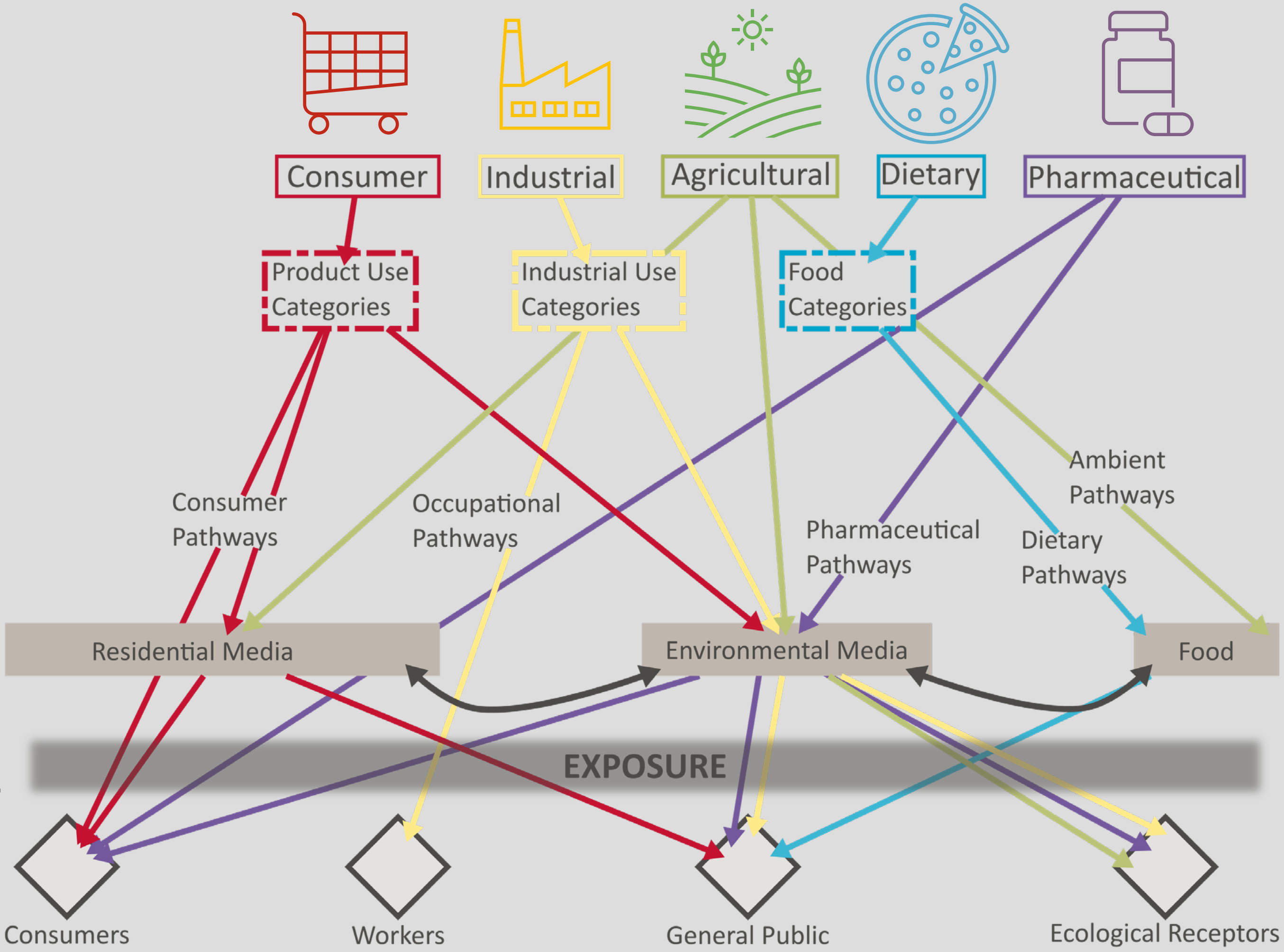


Figure adapted from Kristin Isaacs

United States Environmental Protection Agency

EXPOSURE PATHWAYS

- Exposure pathways provide structured ways to track exposure from a source to a receptor
- Along the … that help …
  - What p… contain… a time?… on…
  - What m… What is… chemica… At what… person… chemica…
- Maybe you could get all your questions answered for one chemical if you had infinite time and resources, but what if you don't have that?...and you have thousands of chemicals?

New approach methods (NAMs), similar to those in Wambaugh et al., 2019 can help address some of these gaps in exposure data


ELSEVIER
Current Opinion in Toxicology
Volume 15, June 2019, Pages 76-92

New approach methodologies for exposure science

John F. Wambaugh[1], Jane C. Bare[2], Courtney C. Carignan[3], Kathie L. Dionisio[4], Robin E. Dodson[5], Olivier Jolliet[6], Xiaoyu Liu[7], David E. Meyer[2], Seth R. Newton[4], Katherine A. Phillips[4], Paul S. Price[4], Caroline L. Ring[8], Hyeong-Moo Shin[9], Jon R. Sobus[4], Tamara Tal[10], Elin M. Ulrich[4], Daniel A. Vallero[4], Barbara A. Wetmore[4], Kristin K. Isaacs[4]

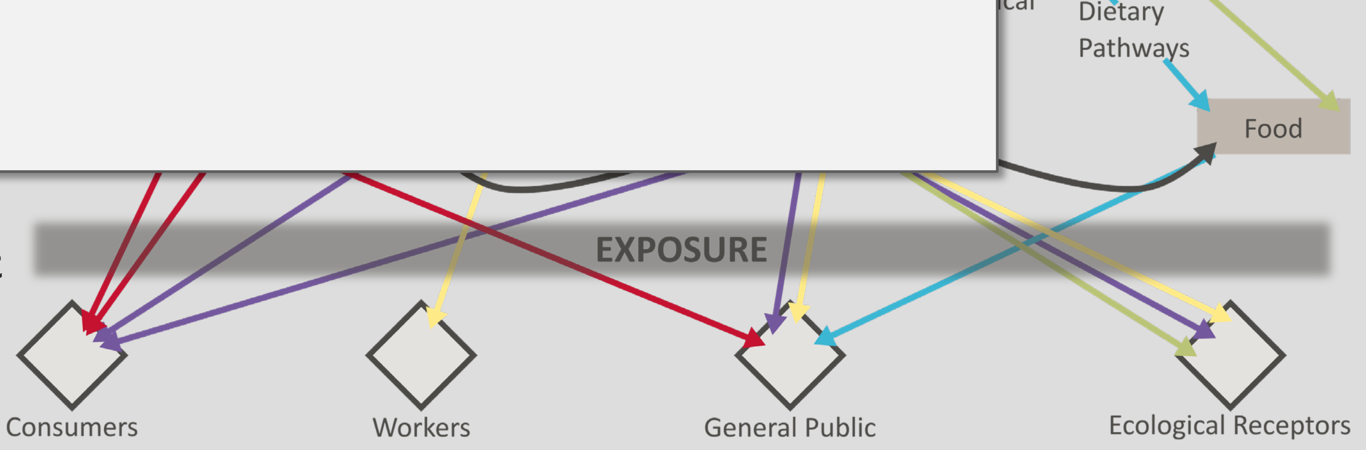
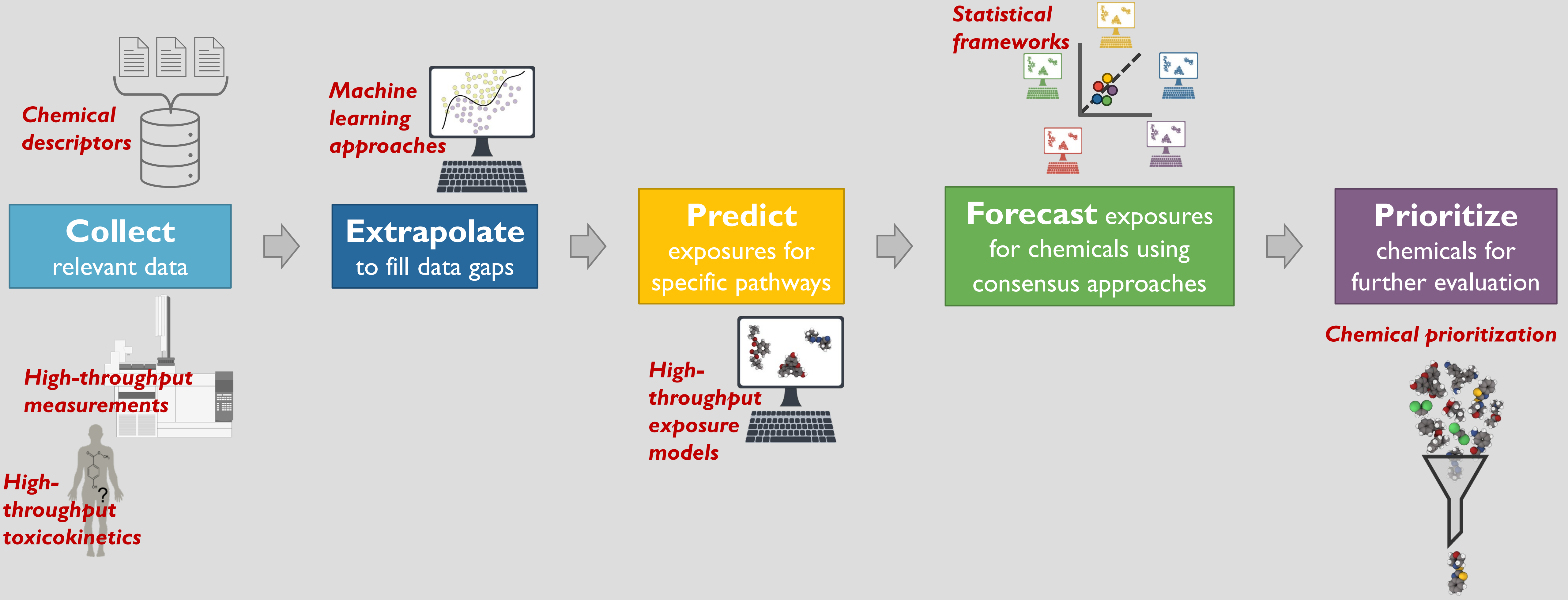


Figure adapted from Kristin Isaacs

EXPOSURE-NAMS



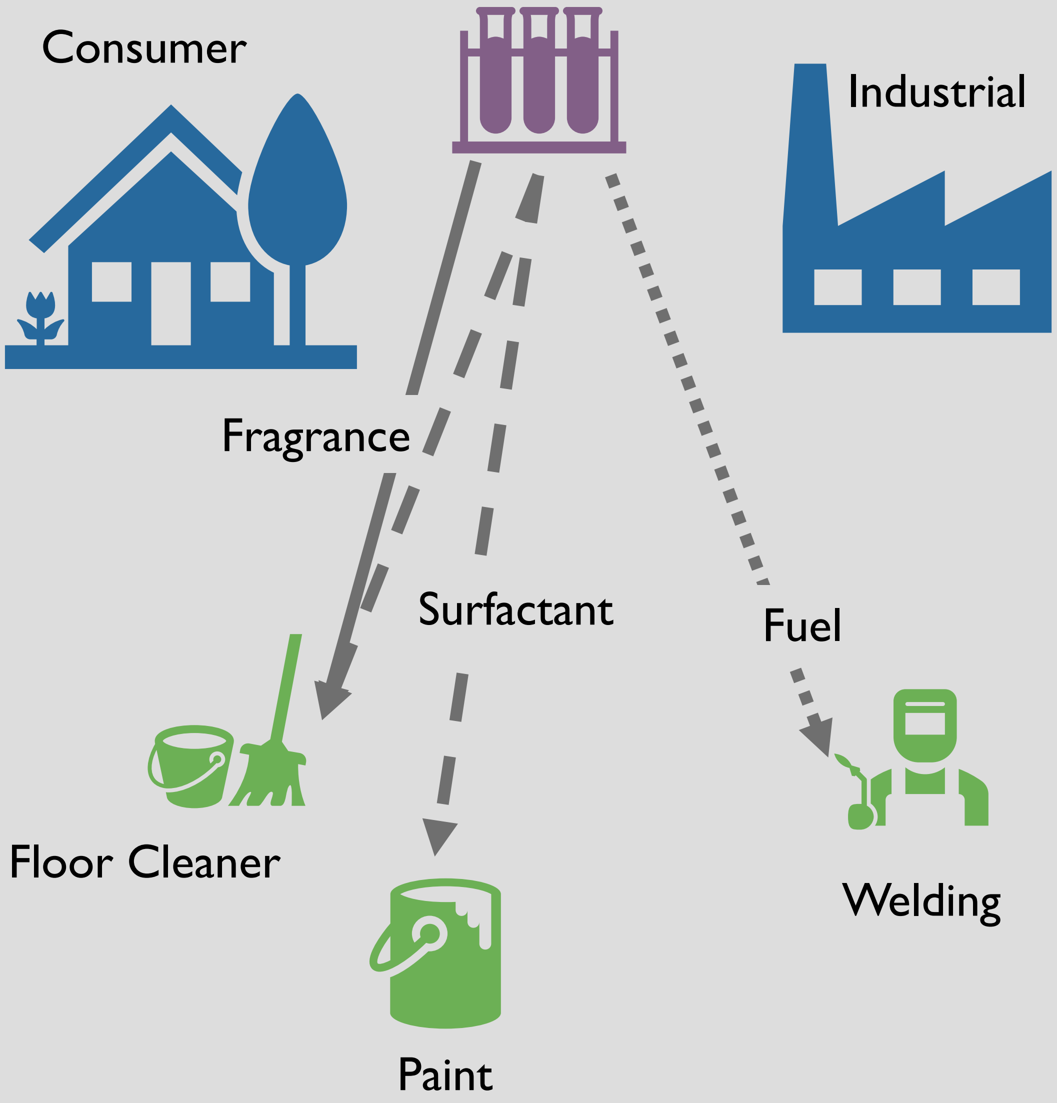
Slide Adapted from Kristin Isaacs
Based on Wambaugh et al. 2019

CHEMICAL USE INFORMS EXPOSURE PATHWAYS

- Chemical use is an excellent example of how indirect information can be applied to fill gaps in current exposure data

CHEMICAL USE INFORMS EXPOSURE PATHWAYS

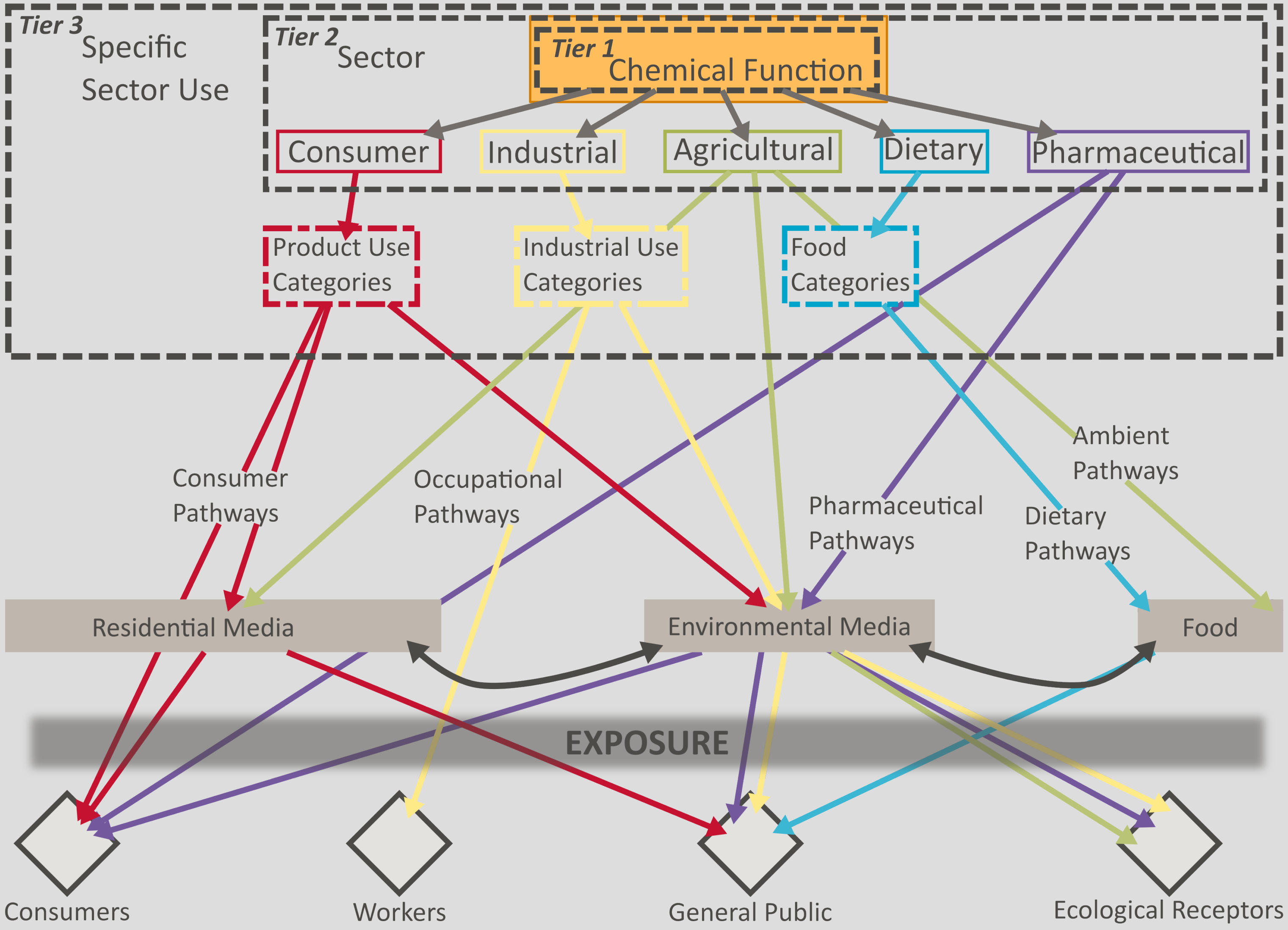
- Chemical use is an excellent example of how indirect information can be applied to fill gaps in current exposure data
  - **Functional use**: role a chemical serves in product or process
  - **Commercial sector use**: which economic sectors use a chemical (industrial vs consumer)
  - **Product use**: the specific product or process a chemical is in and how a population interacts with it



EPA United States Environmental Protection Agency

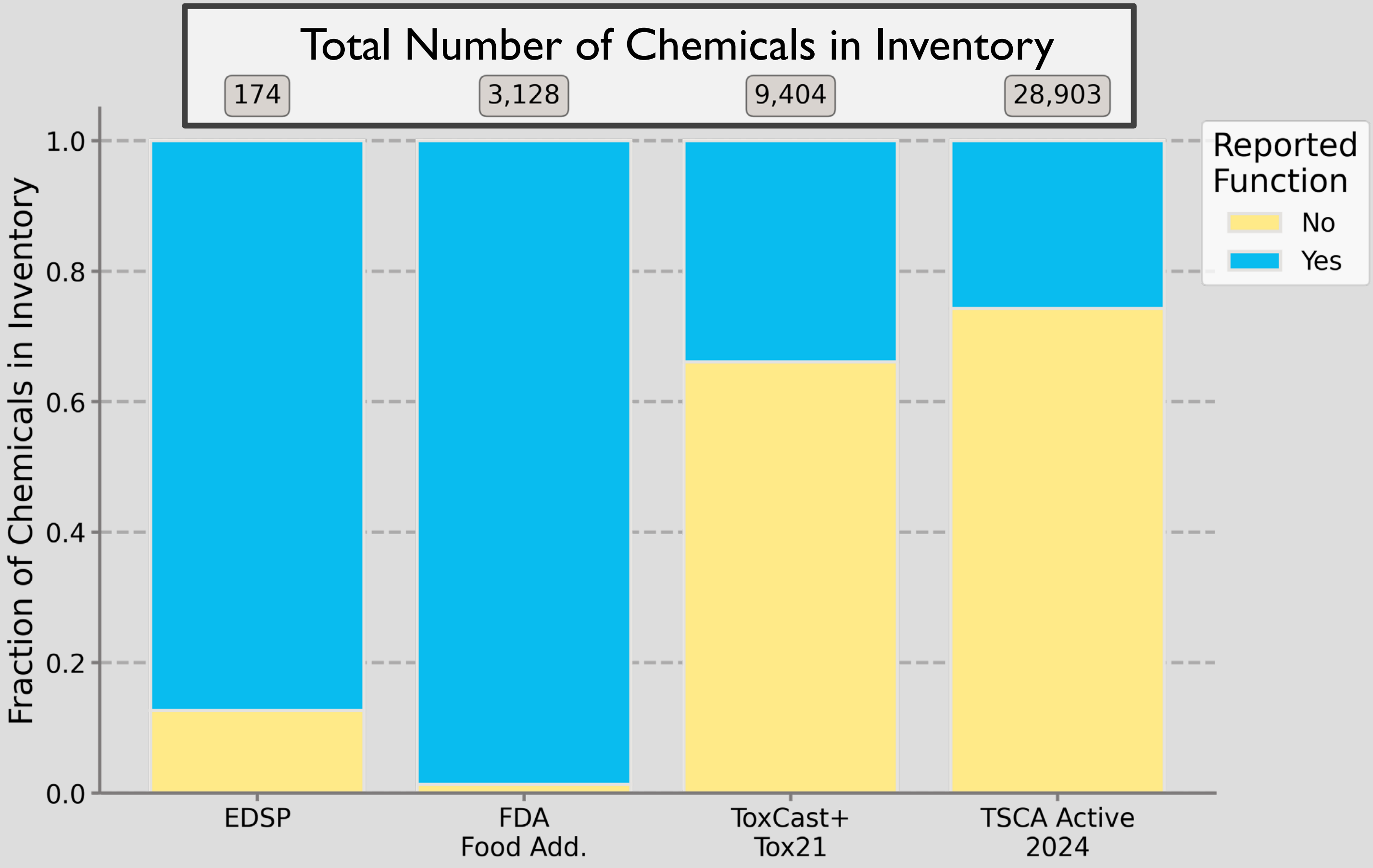
CHEMICAL USE INFORMS EXPOSURE PATHWAYS

- Chemical use is an excellent example of how indirect information can be applied to fill gaps in current exposure data
- A chemical's functional use informs both the commercial sectors and products in which it could be used



CHEMICAL USE INFORMS EXPOSURE PATHWAYS

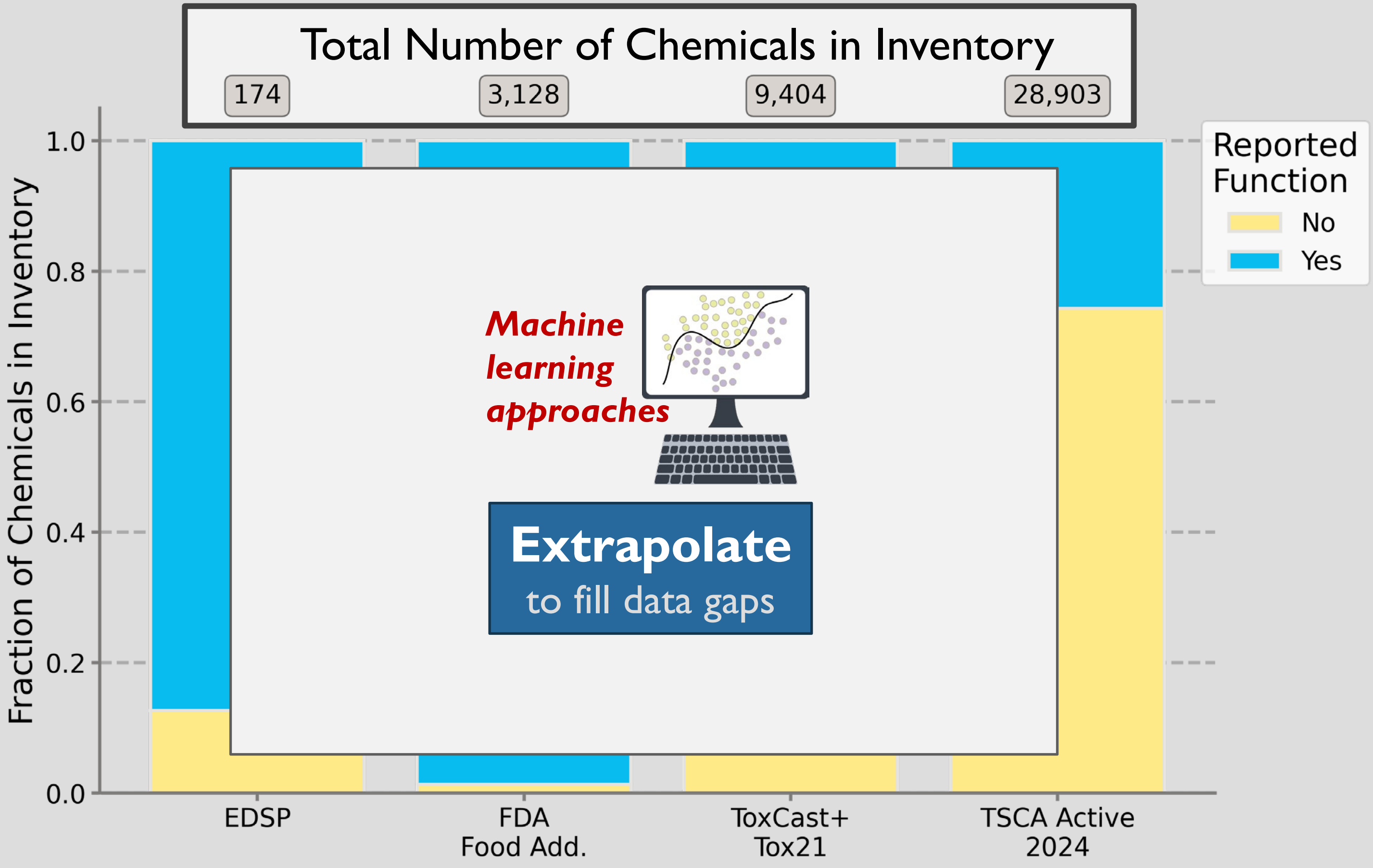
- Chemical use is an excellent example of using indirect information to fill gaps in current exposure data
- A chemical's functional use informs both the commercial sectors and products in which it could be used
- However, use data are lacking for many chemicals



Analysis adapted from Isaacs et al., *JESEE* 2022

CHEMICAL USE INFORMS EXPOSURE PATHWAYS

- Chemical use is an excellent example of using indirect information to fill gaps in current exposure data
- A chemical's functional use informs both the commercial sectors and products in which it could be used
- However, use data are lacking for many chemicals
- Using machine learning, we can extrapolate a chemical's functional use from reported use data

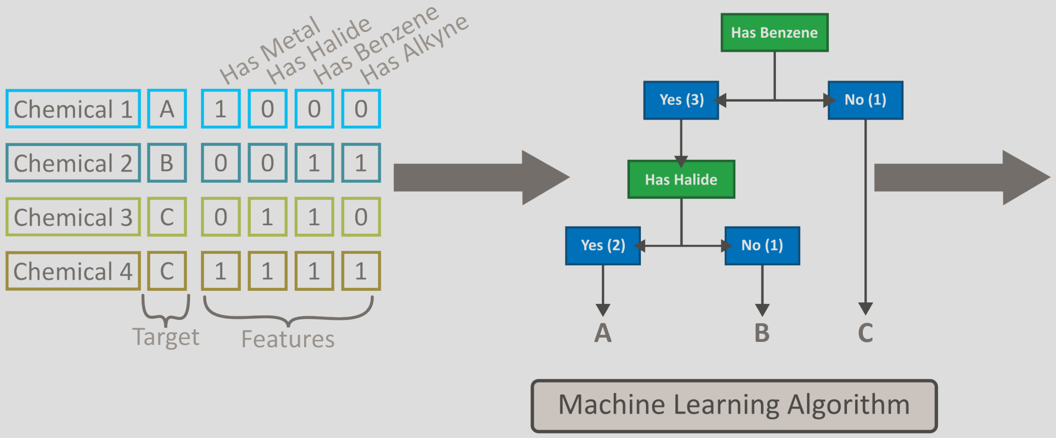


Analysis adapted from Isaacs et al., *JESEE* 2022

United States
Environmental Protection
Agency

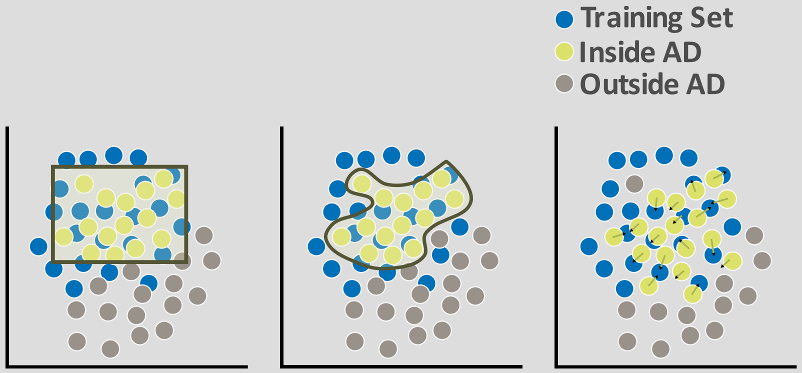
MACHINE LEARNING WORKFLOW

Train the Model

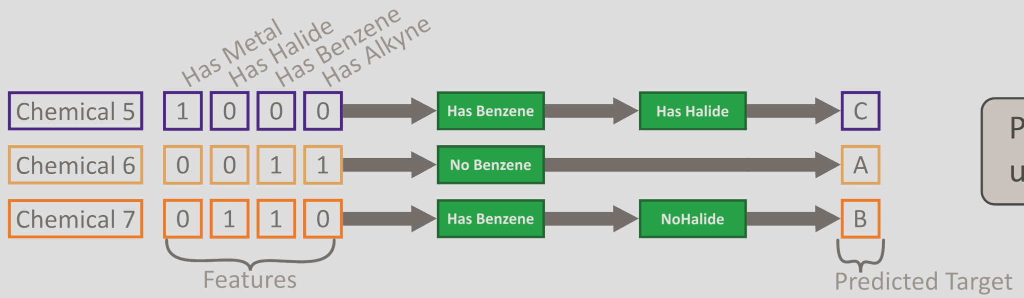


Valid models ***must***:
- accurately predict the training set
- predict beyond the training set
- predict better than an random training set

Ensure predictions made with valid models are in Applicability Domain (AD)



Predict with the Model

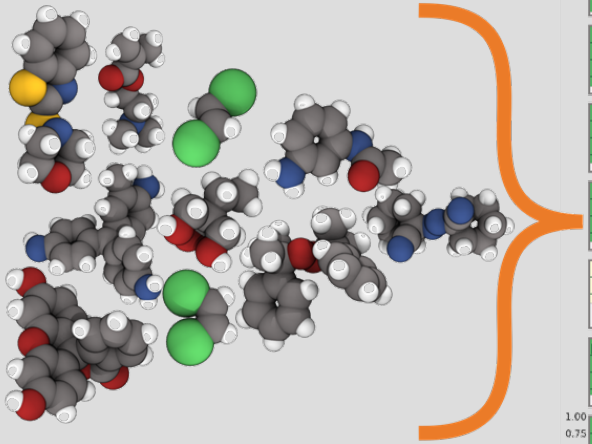


Predict targets with ***valid*** models using features

United States Environmental Protection Agency

QUANTITATIVE STRUCTURE USE RELATIONSHIPS

Chemicals with no reported use

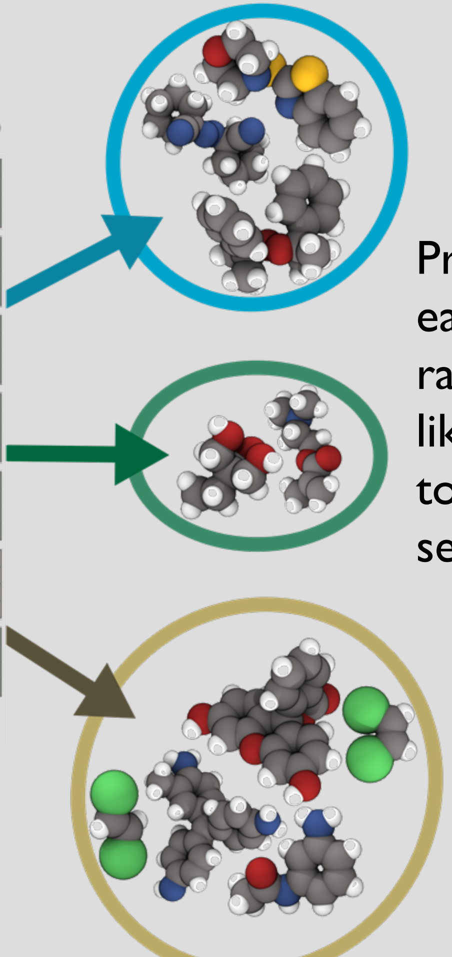


Any chemical with a QSAR-amenable structure

Suite of QSUR models

39 out of 49 QSUR models were valid using structure only as ML features

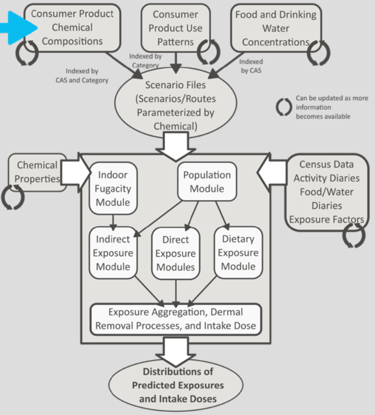
Predicted uses for chemicals



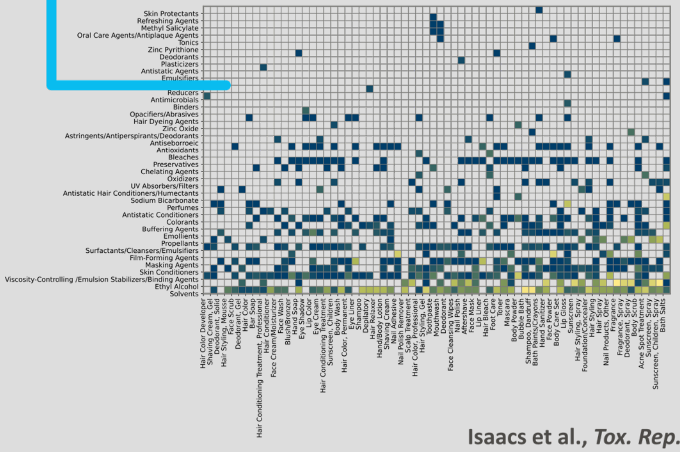
Probabilities are output for each of the 39 models and range from 0 (chemical not likely to serve the function) to 1(chemical likely to serve the function)

Phillips et al., *Green Chem.*, 2017

APPLICATIONS OF QSURS

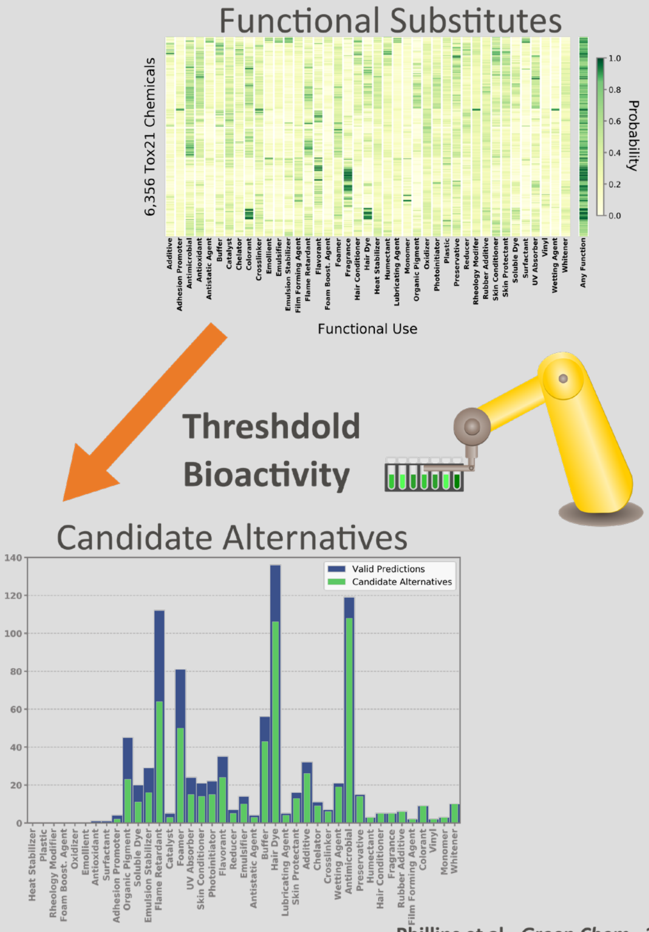


Isaacs et al., *Env. Sci. & Tech.*, 2014



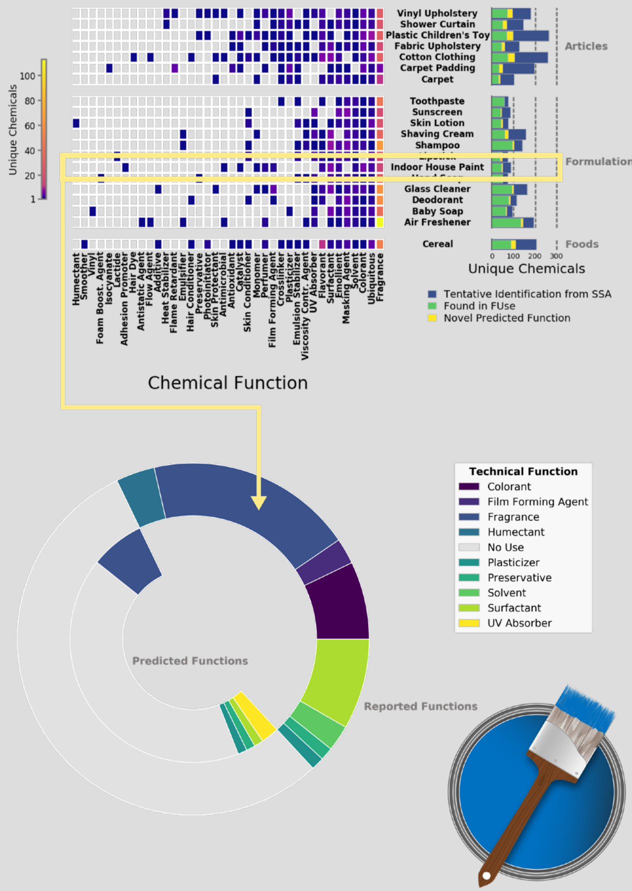
Isaacs et al., *Tox. Rep.*, 2016

Consumer Exposure Models



Phillips et al., *Green Chem.*, 2016

Alternatives Screening



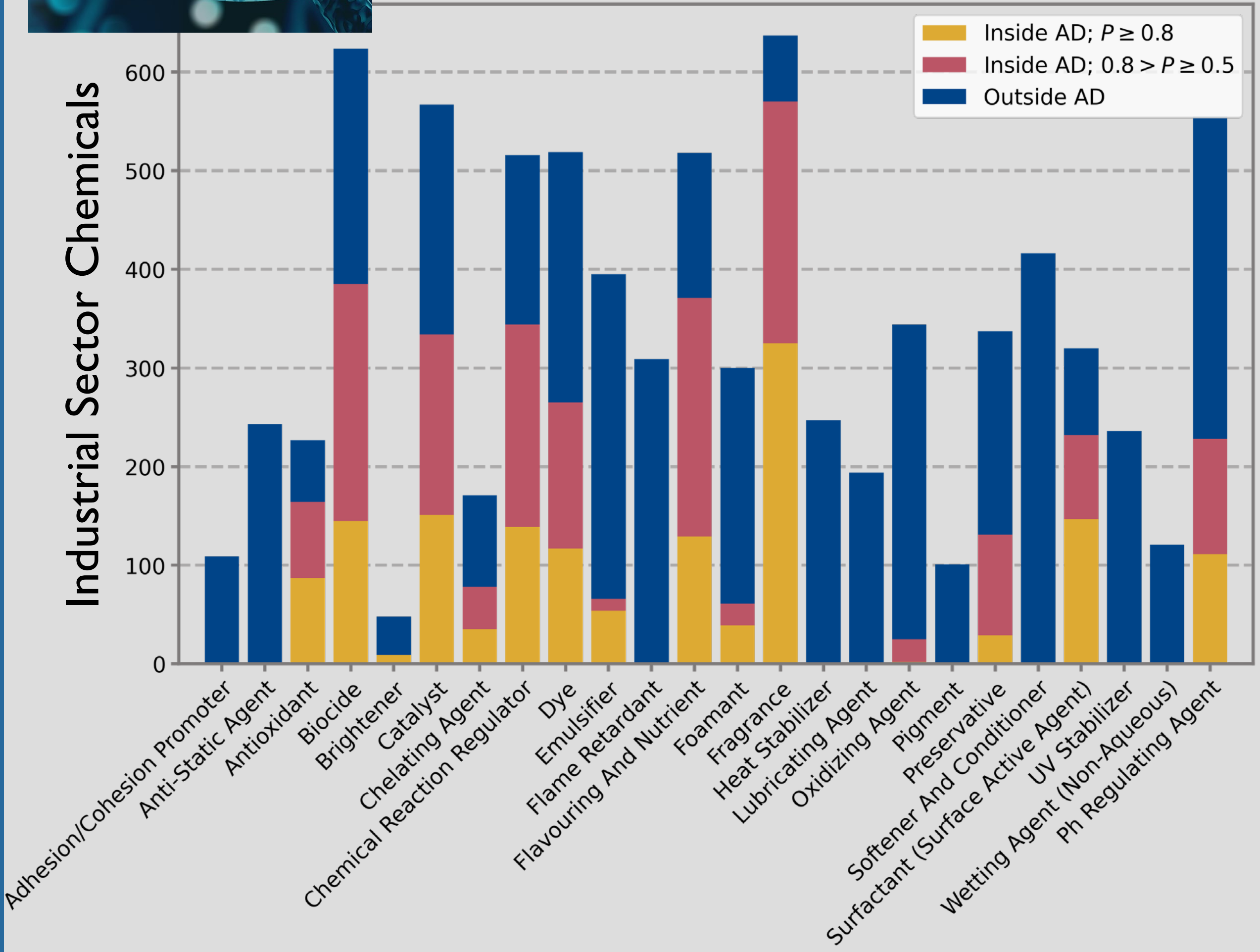
Phillips et al., *Env. Sci. & Tech.*, 2017

Suspect Screening Identification

United States Environmental Protection Agency

QSUR IMPROVEMENT

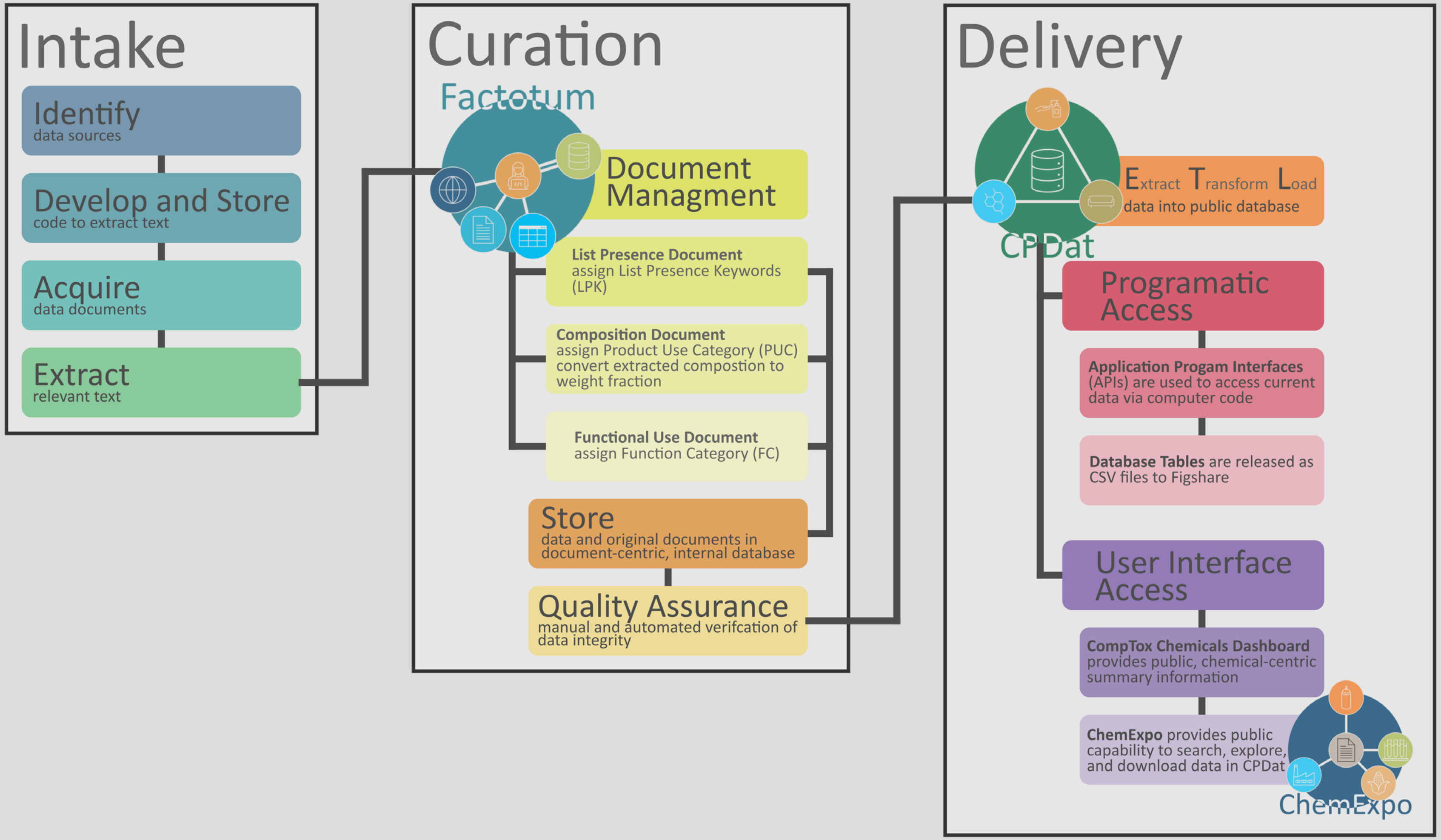
- QSUR models were trained primarily on consumer data
- QSUR predictions were evaluated against industrial chemical use data from U.S. EPA, Health Canada, and ECHA
- As part of an ongoing APCRA cases study, models are being rebuilt to account for this new space of chemicals



United States Environmental Protection Agency

NEW DATA FOR VERSION 2

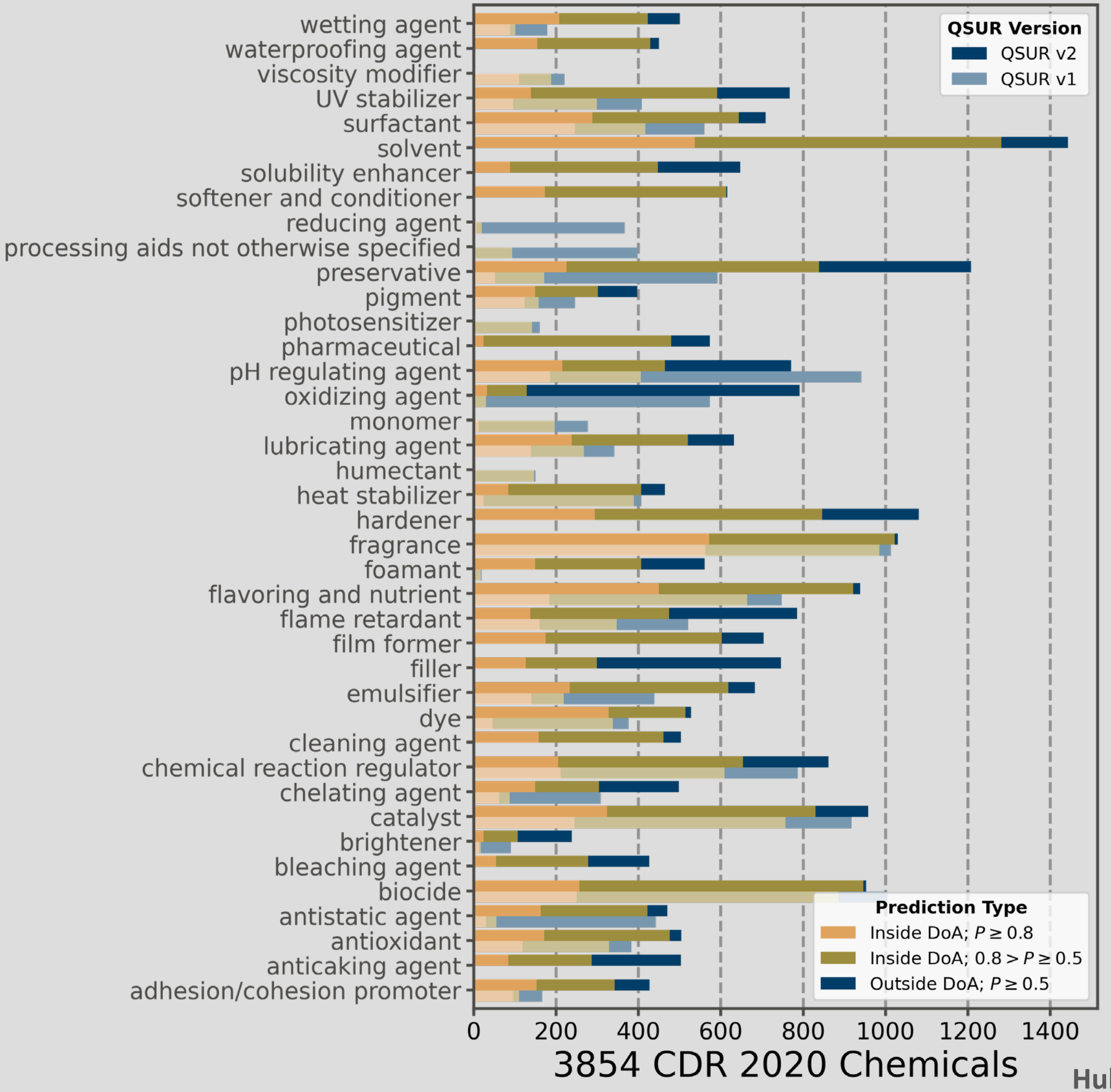
- The Chemical and Products Database (CPDat) is where we store our chemical use data. It contains:
  - Product composition and use
  - Functional use
  - Broad use across industrial sectors
- Using the updated functional use available in CPDat expands the training set for QSUR models.
- Data are available through the ChemExpo online tool: https://comptox.epa.gov/chemexpo



Handa et al., *Sci. Data.*, submitted

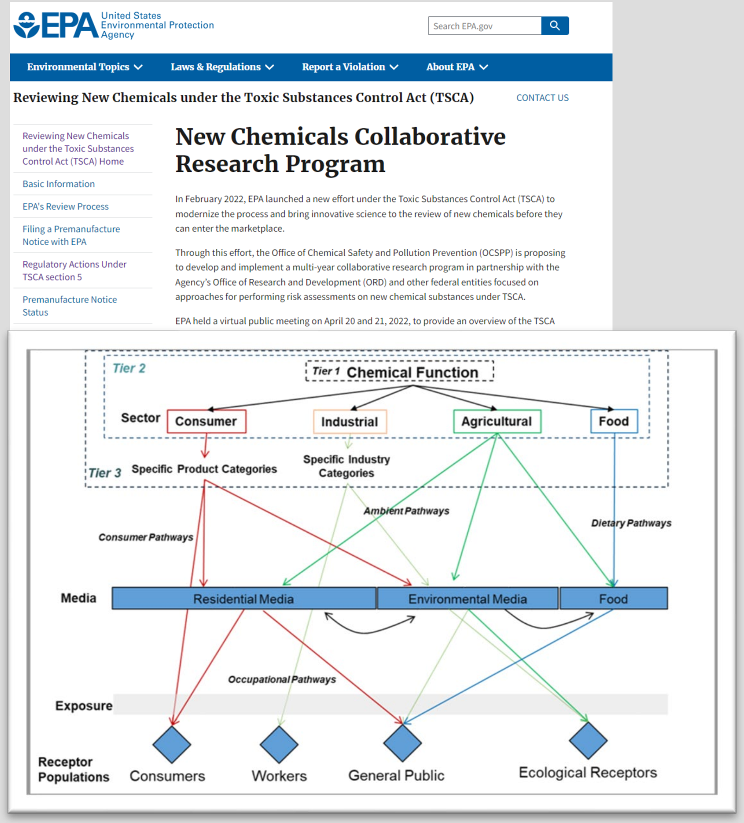
QSUR PREDICTIONS FOR 2020 CDR

- Took ~3,000 chemicals from 2020 CDR and predicted with version 1 and 2 models
- Any predictions below *p*=0.5 are not shown
- 24 overlapping models were valid
- Generally, there are more "in domain" predictions for both consumer and industrial chemicals with 2022 models



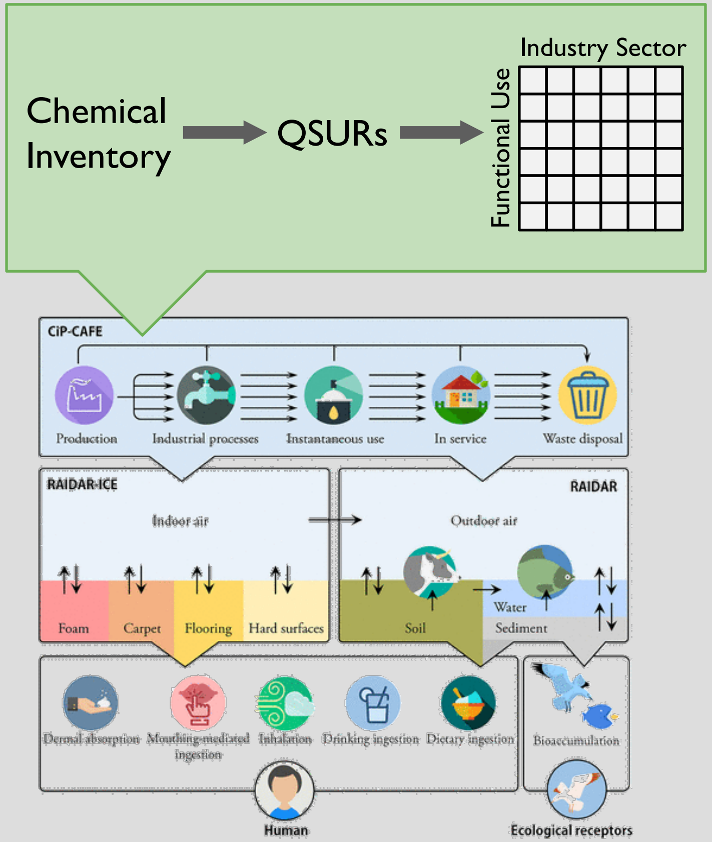
Hull et al., in preparation

CURRENT APPLICATIONS OF QSURS



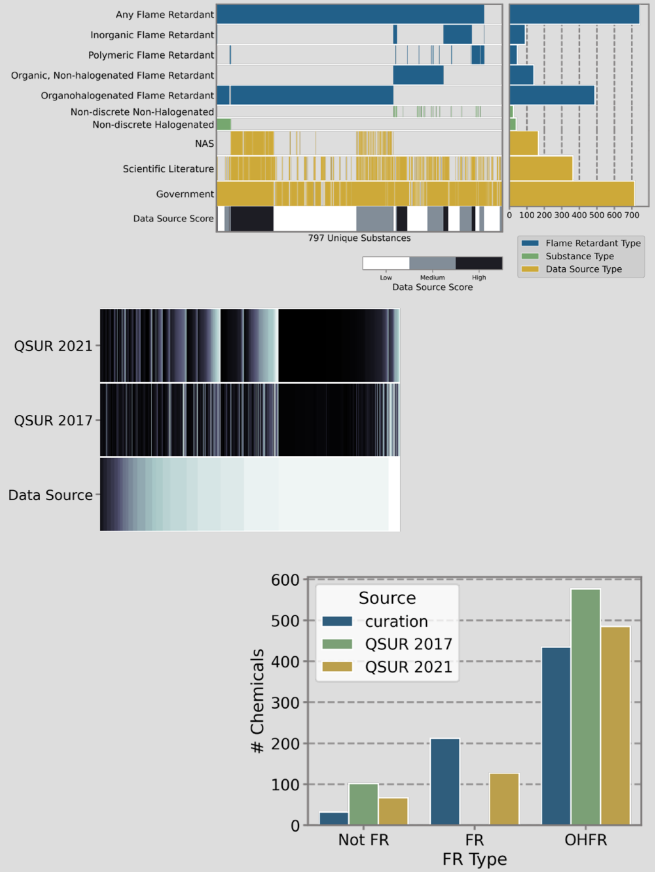
Proposed for identification of "conditions of use" and "use analogues"

New Chemicals Collaborative



Li et al., *Environ. Health Perspect.*, 2021

Environmental Exposure Models

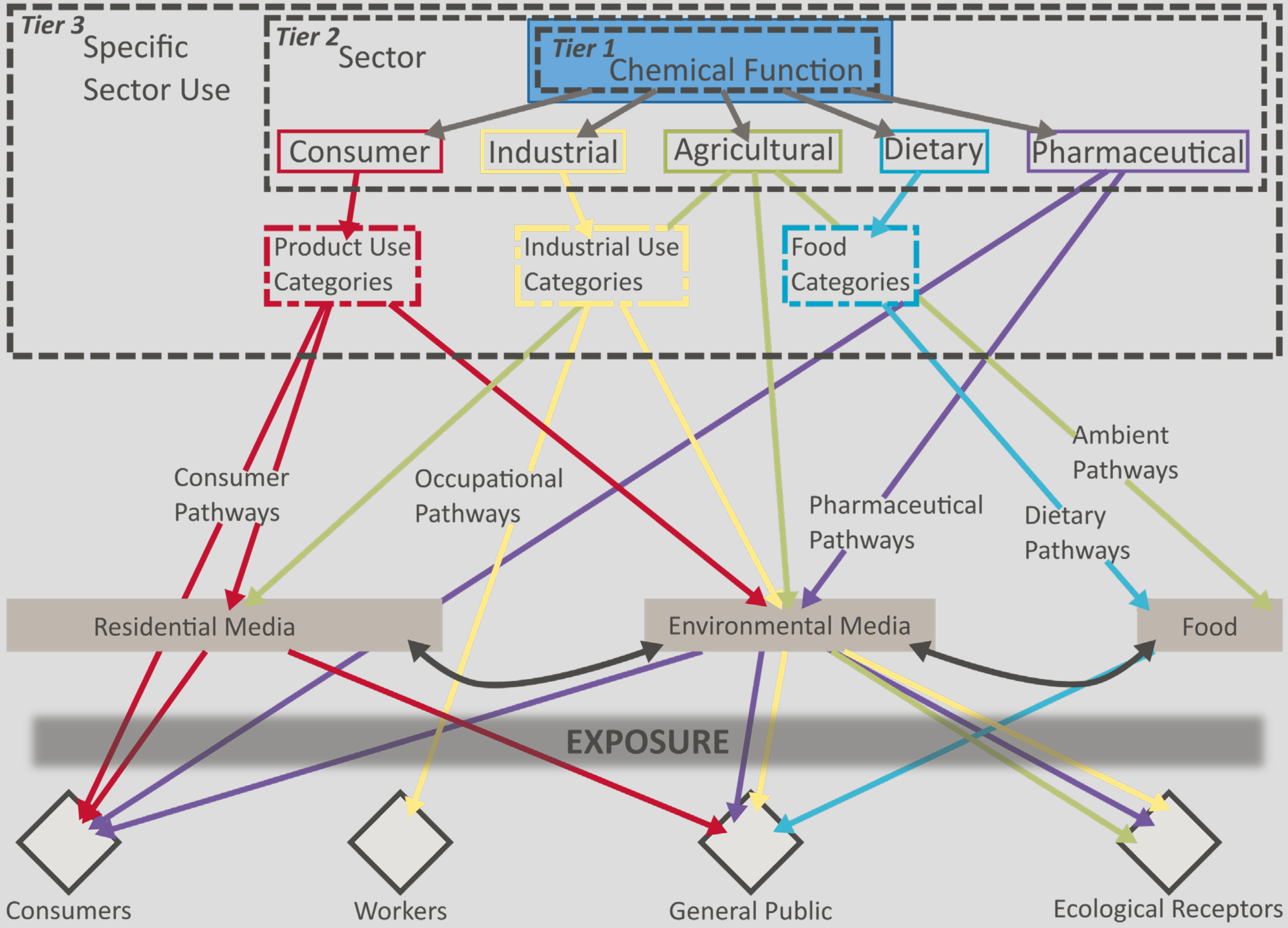


Bevington et al., *Sci. Data*, 2022

Refine for Specific Needs

GAPS IN EXPOSURE PATHWAYS

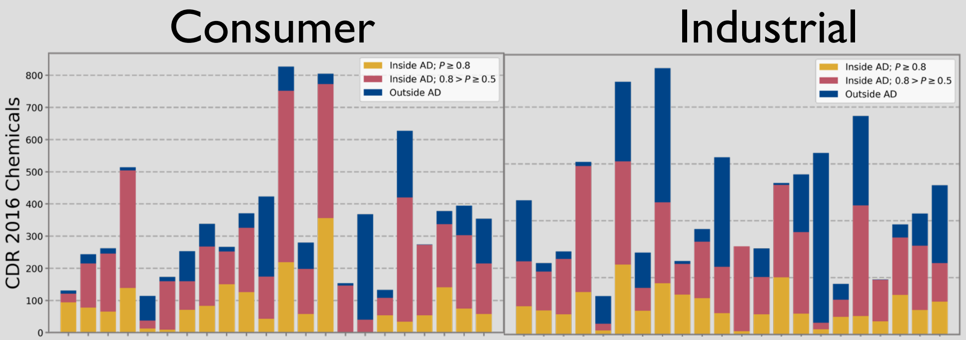
- While QSURs fill consequential gaps, there still areas in to fill along exposure pathways

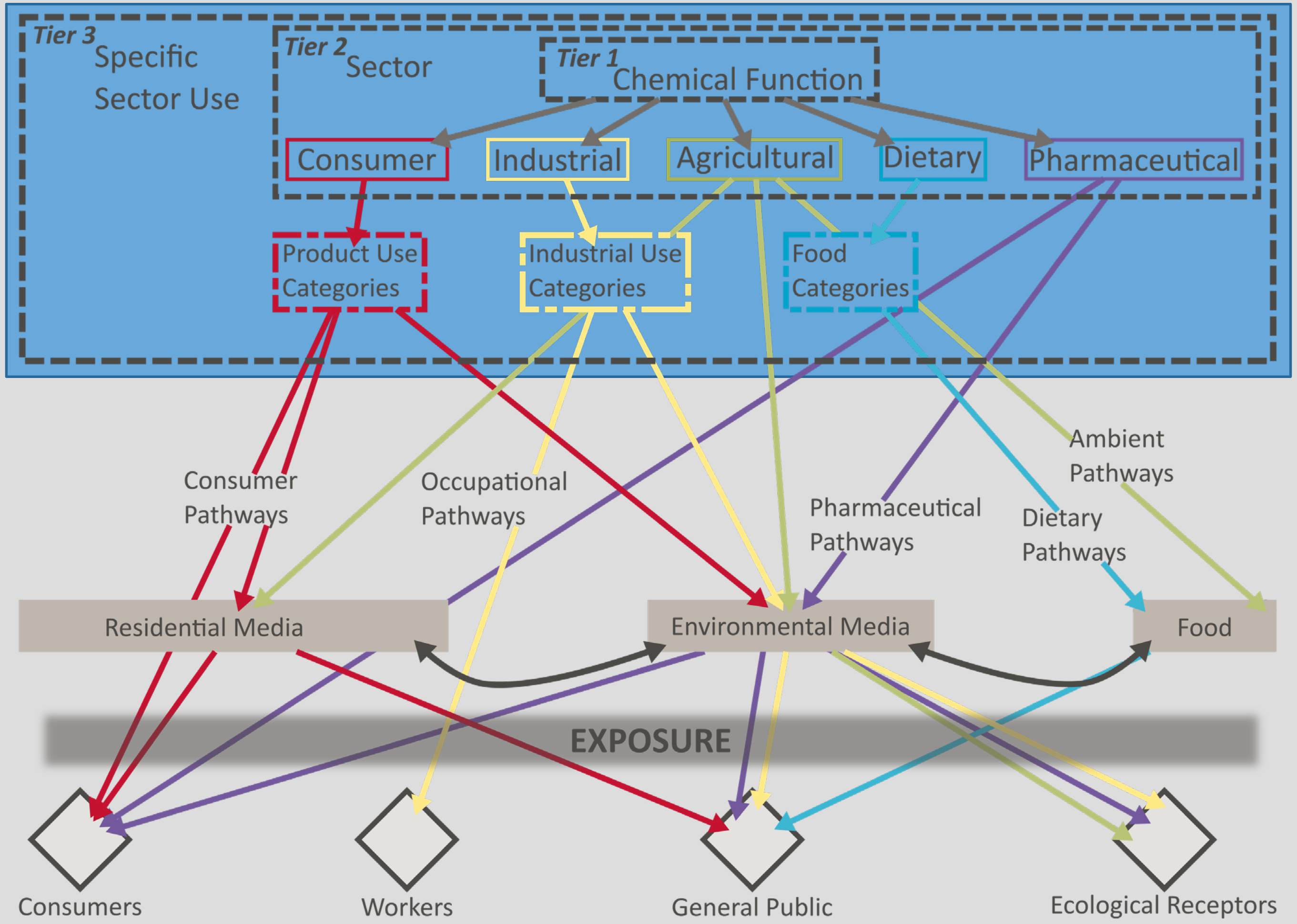


United States Environmental Protection Agency

GAPS IN EXPOSURE PATHWAYS

- While QSURs fill consequential gaps, there still areas in to fill along exposure pathways
- Models are currently being developed (using the same QSUR training set) to predict industrial sectors and products

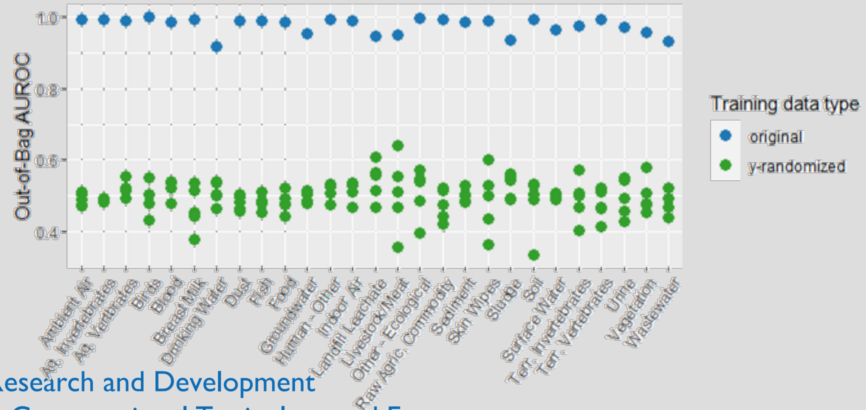


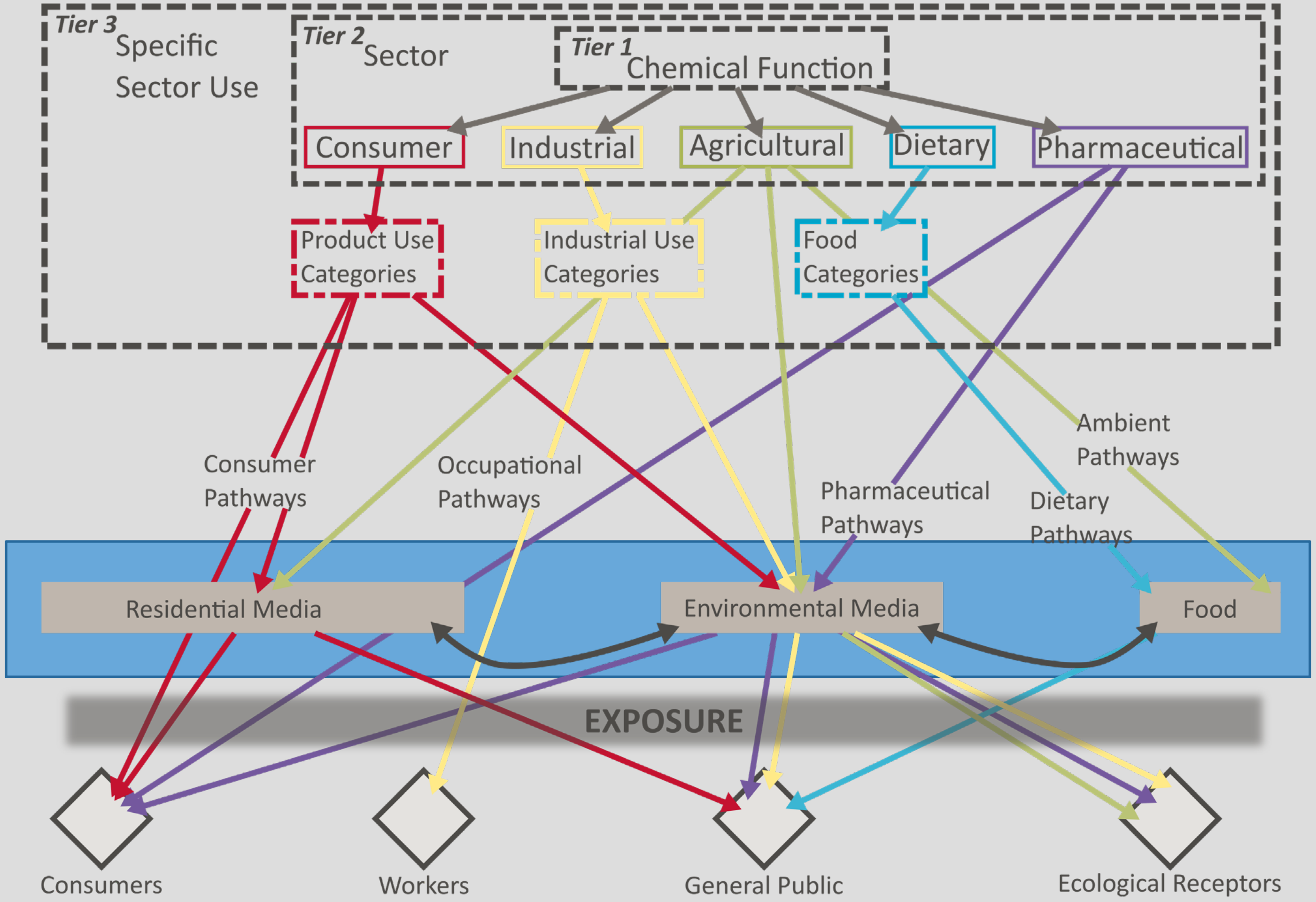


Hull et al., in preparation

GAPS IN EXPOSURE PATHWAYS

- While QSURs fill consequential gaps, there still areas in to fill along exposure pathways
- Models are currently being developed (using the same QSUR training set) to predict commercial sectors and products
- Similar structure-based models can also predict
  - chemical occurrence in media.

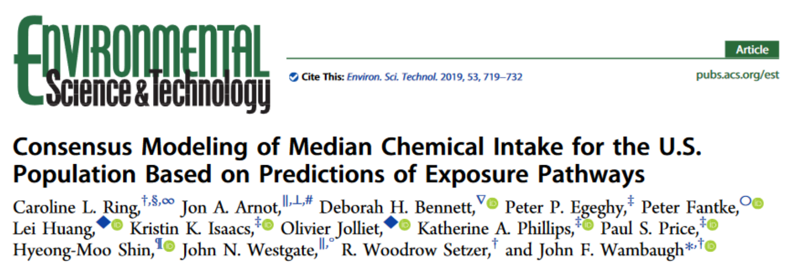




Eddy et al., in preparation

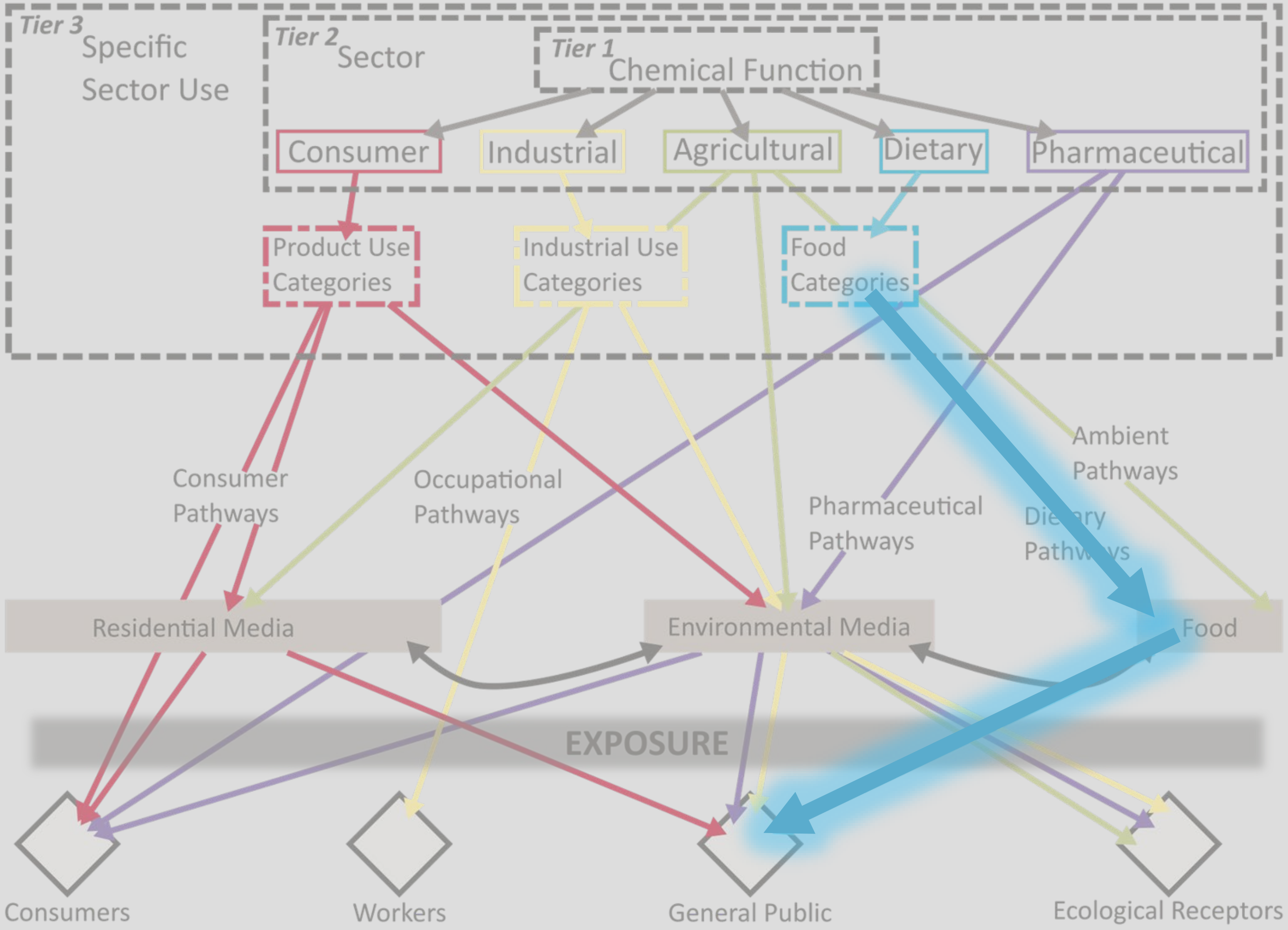
GAPS IN EXPOSURE PATHWAYS

- While QSURs fill consequential gaps, there still areas in to fill along exposure pathways
- Models are currently being developed (using the same QSUR training set) to predict commercial sectors and products
- Similar structure-based models can also predict
  - chemical occurrence in media
  - exposure pathways


Environmental Science & Technology
Article
Cite This: Environ. Sci. Technol. 2019, 53, 719–732
pubs.acs.org/est

Consensus Modeling of Median Chemical Intake for the U.S. Population Based on Predictions of Exposure Pathways

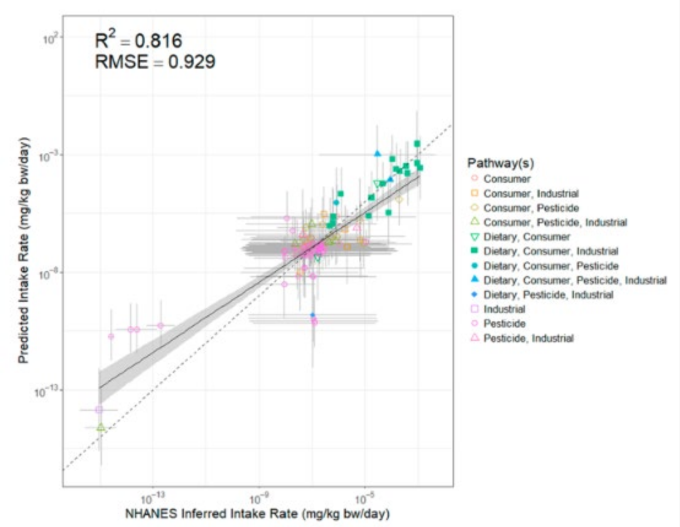
Caroline L. Ring,[†,§,∞] Jon A. Arnot,[∥,⊥,#] Deborah H. Bennett,[∇] Peter P. Egeghy,[‡] Peter Fantke,[○] Lei Huang,[◆] Kristin K. Isaacs,[‡] Olivier Jolliet,[◆] Katherine A. Phillips,[‡] Paul S. Price,[‡] Hyeong-Moo Shin,[¶] John N. Westgate,[∥,°] R. Woodrow Setzer,[†] and John F. Wambaugh[*,†]



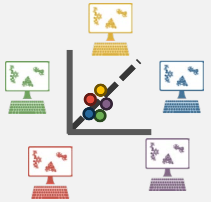
Ring et al., 2018, *ES&T*

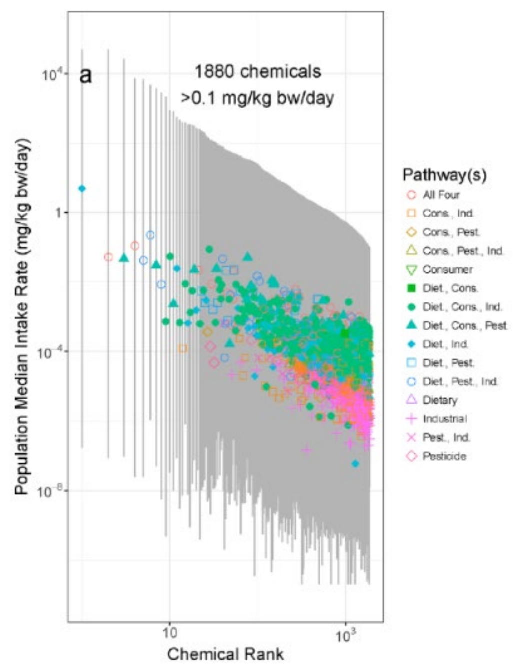
GAPS IN EXPOSURE PATHWAYS

- While Q
  there sti
  pathways
- Models a
  (using th
  predict c
- Models a
  - predict
  - exposu

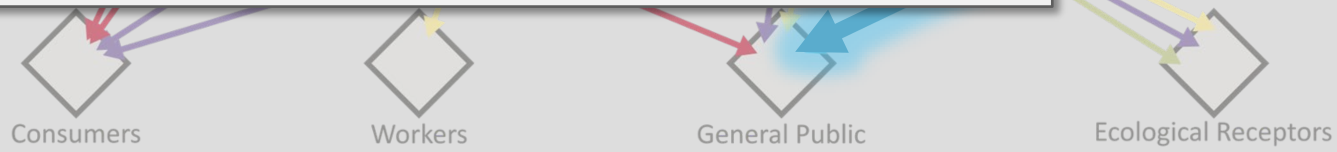


These models provide crucial input into statistical frameworks that forecast exposures along different pathways.





Ring et al., 2018, *ES&T*



Environmen
Science & Techn

Consensus Modeling of Median Chemical Intake for the U.S. Population Based on Predictions of Exposure Pathways

Caroline L. Ring, Jon A. Arnot, Deborah H. Bennett, Peter P. Egeghy, Peter Fantke, Lei Huang, Kristin K. Isaacs, Olivier Jolliet, Katherine A. Phillips, Paul S. Price, Hyeong-Moo Shin, John N. Westgate, R. Woodrow Setzer, and John F. Wambaugh

Ring et al., 2018, *ES&T*

United States Environmental Protection Agency

THANKS

New QSUR Models

- Victoria Hull (former trainee)
- Kristin Isaacs

Training Data from CPDat

- Allison Larger (GDIT)
- All former trainees from the data curation team
- Sakshi Handa
- Kristin Isaacs

Refined Flame Retardant QSURs

- Colin Guider (former trainee)
- Charles Bevington (U.S. CPSC)
- Michael Babich (U.S. CPSC)

Media Prediction Models

- Lindsey Eddy (former trainee)
- Kristin Isaacs