



www.epa.gov

Exploring exposure pathways with Chemical/Product Categorical (CPCat) Data

KL Dionisio¹, A Frame¹, M-R Goldsmith¹, JF Wambaugh¹, A Liddell¹, T Cathey², D Smith¹, J Vail¹, RS Judson¹

¹U.S. EPA, Office of Research and Development, RTP, NC, USA

²Lockheed Martin, Durham, NC, USA



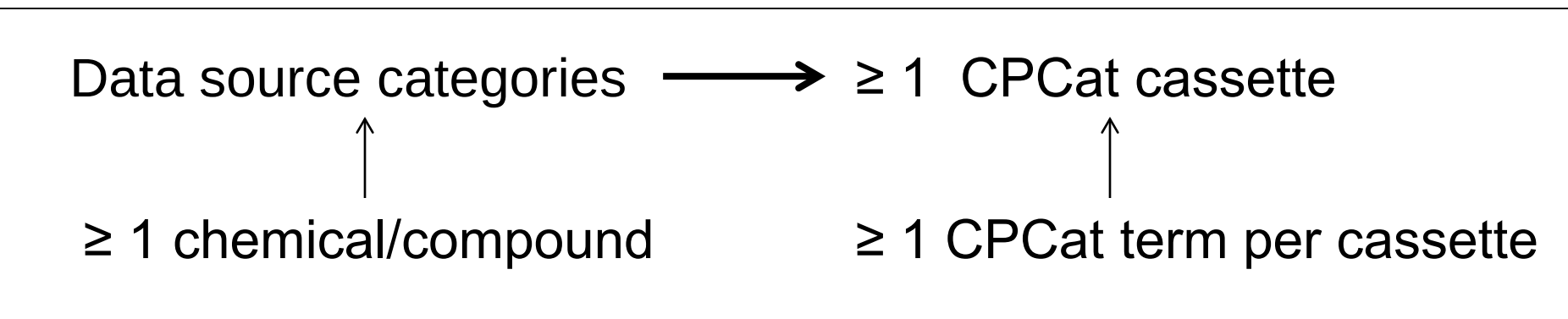
Kathie L. Dionisio | dionisio.kathie@epa.gov | 919-541-1321

Abstract

Humans are exposed to thousands of chemicals over our lifetimes. A major challenge to risk assessors is to understand how and when chemical exposures occur, and which “exposure pathways” contribute the most. An informatics-driven approach to assigning “product-use” categories to product “ingredients” will help prioritize which chemicals will be given more scrutiny for a target population (life-stage, gender, ethnicity), identifying (a) human activities that result in increased chemical exposure while (b) reducing the dimensionality of hazard assessment to a tangible subset for risk characterization. How chemicals are used is directly related to their potential exposure routes, and in the sense of near-field source apportionment, to the identification of chemical-specific use-to-receptor exposure pathways. We have developed the Chemical/Products Categories Database (CPCat) to explore these pathways. CPCat is both a front-end interface and back-end relational database comprised of chemical “use category” information, including both consumer- and industrial-process based uses for chemicals, using mined and curated data from 17 major sources from multiple countries, including data from regulatory bodies, manufacturers and retailers. CPCat contains information on >43,000 unique chemicals mapped to a set of >850 use categories. Examples are provided of how CPCat can be used to identify potential exposure routes, or prioritize chemicals for high-throughput screening. The full CPCat database is available for public download. *The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. EPA.*

CPCat methods

- Each data source has its own category hierarchy
- Category hierarchies are harmonized in CPCat into a unique set of CPCat “terms” (keywords/categories)
- Categories from data sources were manually mapped to one or more CPCat “cassettes”
- Each CPCat cassette is made up of one or more CPCat terms describing the product
- CPCat assigns 378 unique terms, excluding drug related terms (~850 terms total); and 1,319 unique cassettes
- Full database available for download at actor.epa.gov/cpcat



CPCat data sources

CPCat includes >43,000 unique chemicals

Source	Type	Source Categories	CPCat Cassettes	Chemicals
CDR 2012 Industrial Function	Functional Categories	35	25	5,023
Dow	Functional Categories	20	18	104
Keml	Functional Categories	62	44	877
Toxome	Functional Categories	17	15	443
CDR 2012 Industrial Sector	Industrial Sector Categories	48	32	5,226
SPIN Industrial Sector	Industrial Sector Categories	1,934	213	4,604
SPIN NACE	Industrial Sector Categories	138	42	7,745
Retail Product Categories	Product Categories	450	186	2,778
DrugBank	Therapeutic Categories	586	458	1,755
ACToR Assays and Lists	Use Categories	137	176	35,732
ACToR UseDB	Use Categories	15	15	31,622
CDR 2012 Consumer	Use Categories	35	33	3,321
DfE	Use Categories	12	9	444
IUR 2006	Use Categories	21	19	1,152
NICNAS	Use Categories	19	16	177
SPIN detpcat	Use Categories	1,210	276	6,491
SPIN UC62	Use Categories	63	51	9,059

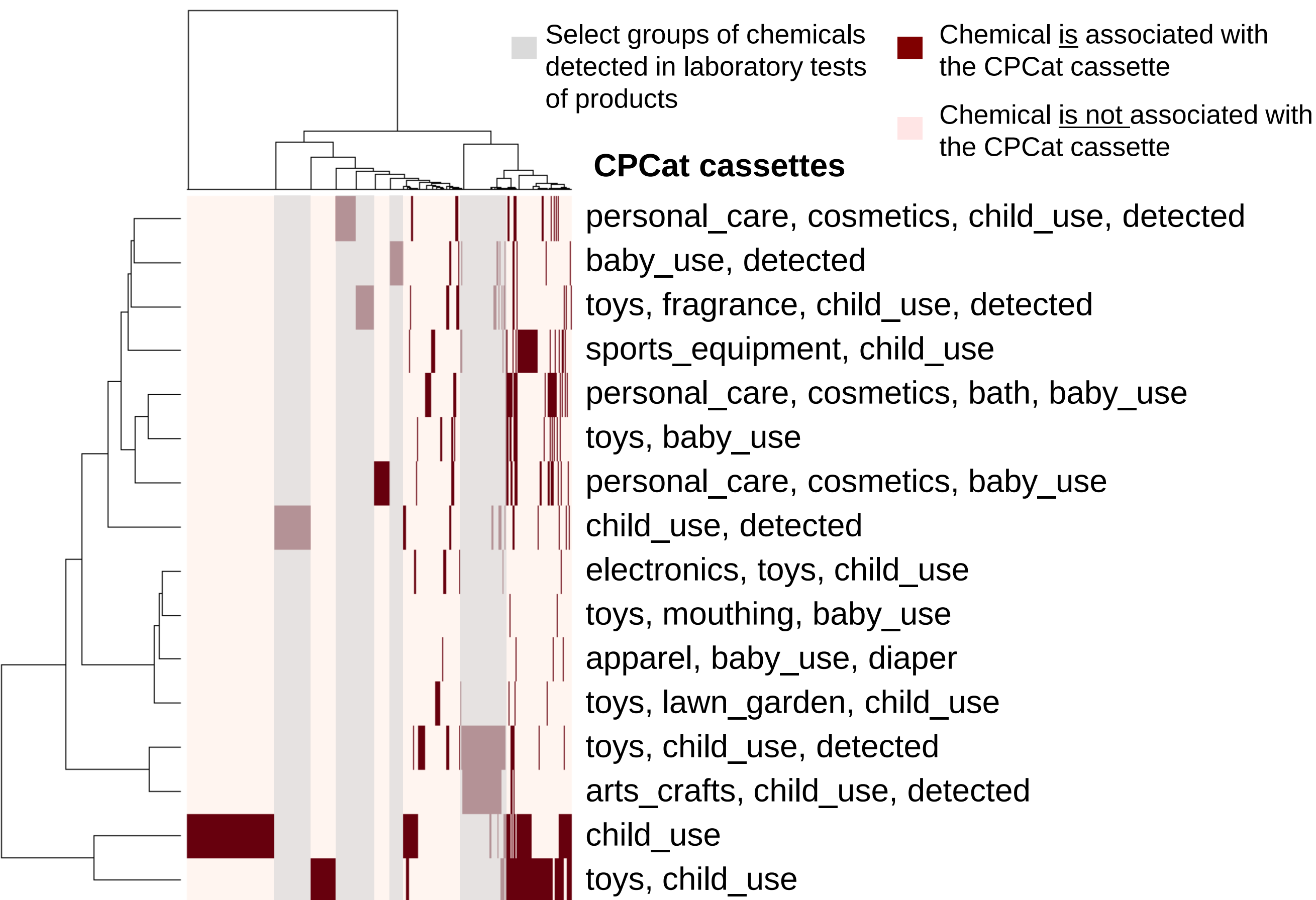
Exposure routes for SDWA chemicals

- 5,251 Safe Drinking Water Act (SDWA) compounds are candidates for exposure and hazard determination in the Endocrine Disruptor Screening Program (EDSP)
- Use CPCat for prioritization of chemicals for screening based on potential exposure routes (more “hits” → more potential exposure routes)

5,251 SDWA chemicals in EDSP ↓ ~4,000 in CPCat ↓ 3,519 associated with at least one consumer-use related CPCat cassette	CAS	Name	CPCat consumer cassette hits
	57-55-6	1,2-Propanediol	122
	64-17-5	ethanol	116
	56-81-5	glycerol	111
	67-63-0	Isopropyl alcohol	92
	99-76-3	methyl 4-hydroxybenzoate	86
	77-92-9	citric acid	85
	1310-73-2	sodium hydroxide	85
	13463-67-7	titanium dioxide	83
	7647-14-5	sodium chloride	81
	102-71-6	2,2,2-nitrilotriethanol	78
	106-97-8	butane	75
	128-37-0	2,6-di-tert-butyl-p-cresol	74
	75-28-5	isobutane	73
	94-13-3	propyl 4-hydroxybenzoate	72

Potential to interact with
estrogen receptor

Child exposure scenario



1,069 chemicals

- 1,069 chemicals associated with 16 baby/child related CPCat cassettes
- Beyond a general child use association, the largest number of chemicals are associated with children’s toys
- Gray bars highlight groups of chemicals detected in laboratory tests of children’s products; the majority of these chemicals are not identified by any other source in CPCat as being associated with potential child exposure

Conclusions

- CPCat is the first publically available, large scale database to harmonize disparate sources of chemical use categorization
- CPCat database may be useful for priority setting tasks, such as in high throughput chemical exposure prioritization
- CPCat can be used to group chemicals by potential types of exposure sources, or by diversity of potential sources