UL CHEMINFORMATICS SUITE

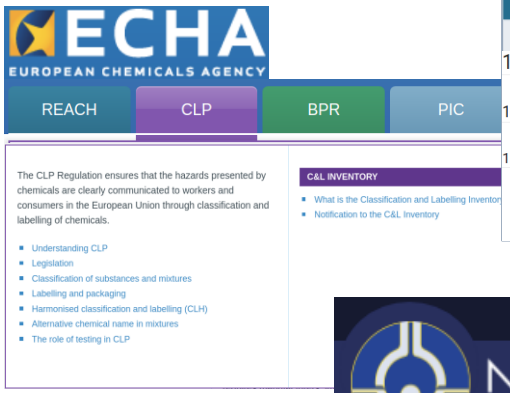
Craig Rowlands, PhD, DABT
Innovation Leader & Senior Toxicologist
UL Product Supply Chain Intelligence

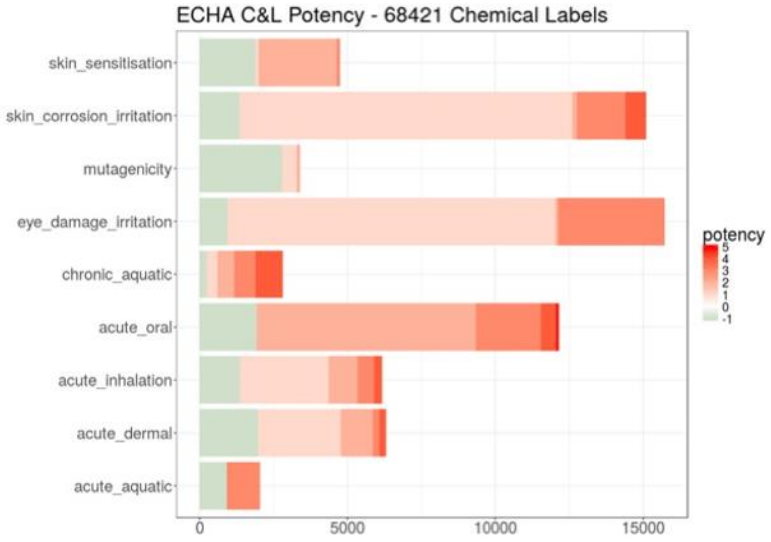
UL | Empowering Trust in a Complex World

UL CHEMINFORMATICS SUITE DESIGN

- Uses machine learning to take advantage of extensive increase in publicly available toxicology, biological, phys-chem and structural data
  - Models are trained on GHS classification hazards (H codes)
  - Predicts GHS classifications for 8 endpoints: acute oral & dermal toxicity, skin & eye irritation, skin sensitization, mutagenicity, and acute & chronic aquatic toxicity – **Acute Inhalation Toxicity endpoint coming in Fall 2018**
  - Models use Read-Across Structure Activity Relationship (RASAR) Dat-Fusion that
    - identifies the most similar chemical to the target chemical based on - structures, hazards and phys-chem properties
    - does not limit similarity searches to “toxicophore” - sub-structures with known toxic mechanisms (i.e., structural alerts)

UL CHEMINFORMATICS SUITE DATA SOURCES





- 250K chemicals with 300K endpoint labels
- 70M Chemical Structures
- GHS classification labels by weight of evidence
- discordance in a chemical label is handled by selecting the most hazardous value (aka the precautionary principle).

CHEMINFORMATICS SUITE

Hazard

1st Module

Registrations

REACH ACROSS
A DIGITAL ALTERNATIVE TO ANIMAL TESTING

ECHA
EUROPEAN CHEMICALS AGENCY

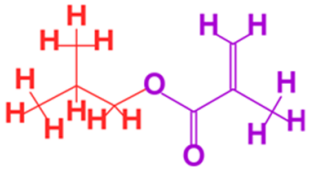
efsa

FDA

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

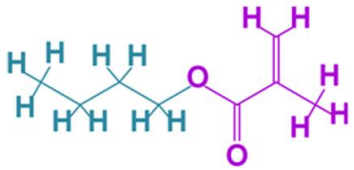
UL CHEMINFORMATICS SUITE – READ ACROSS STRUCTURE ACTIVITY RELATIONSHIP (RASAR)

Isobutyl Methacrylate



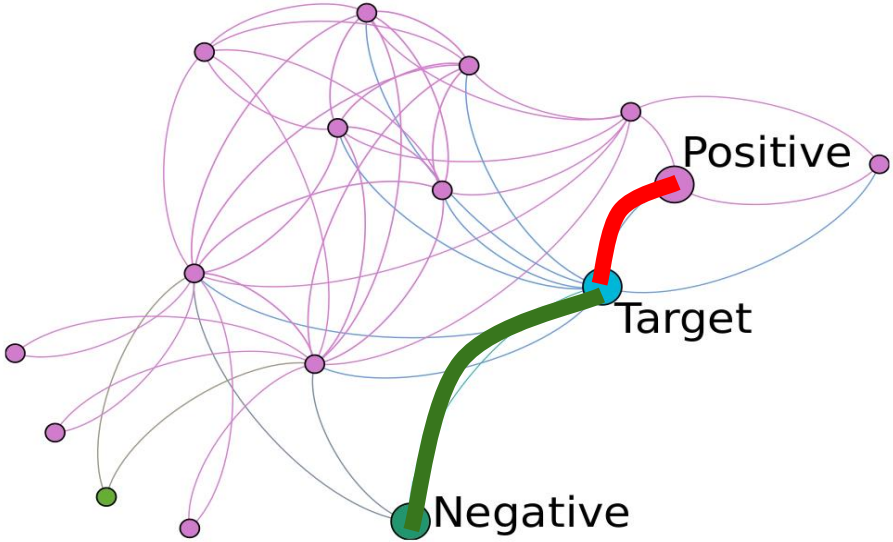
← 90% →

Butyl Methacrylate



$$\frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

Tanimoto - Heuristic

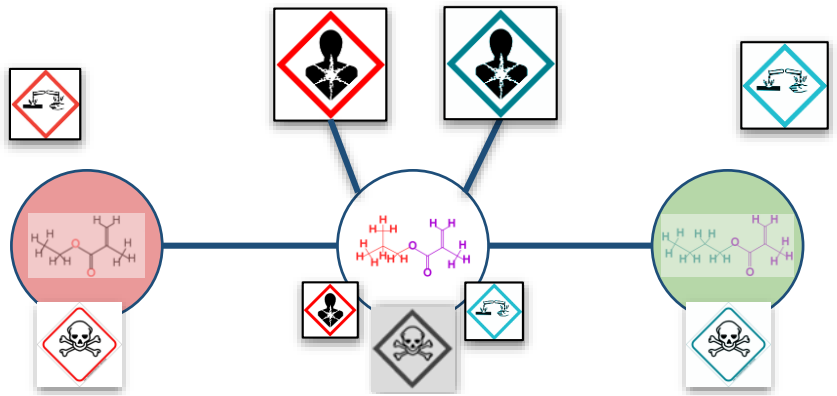


Similarity to closest **Positive**

Similarity to closest **Negative**

UL CHEMINFORMATICS SUITE – REACHACROSS MODULE

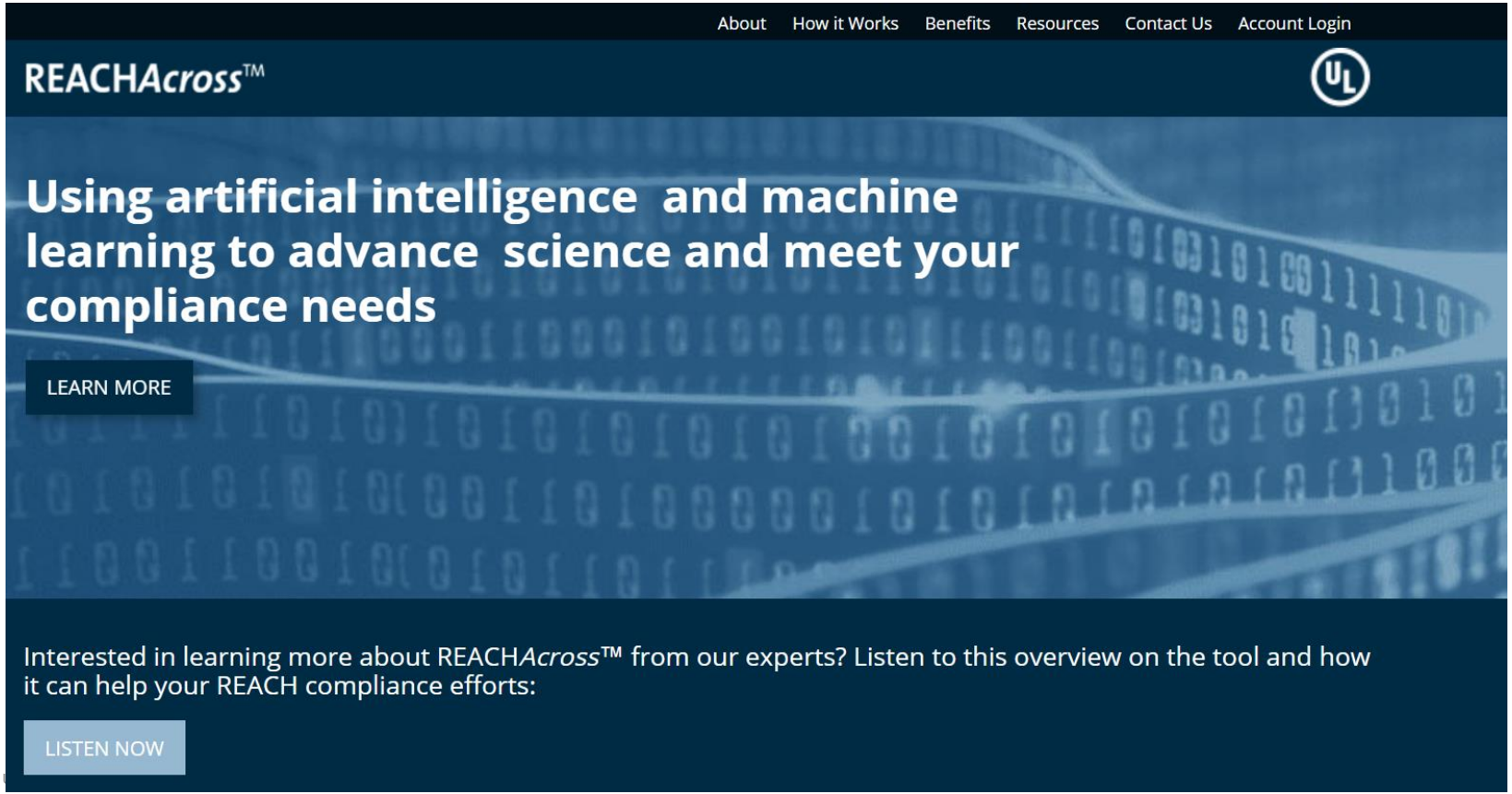
Data-Fusion Read-Across Structure Activity Relationships (RASAR)



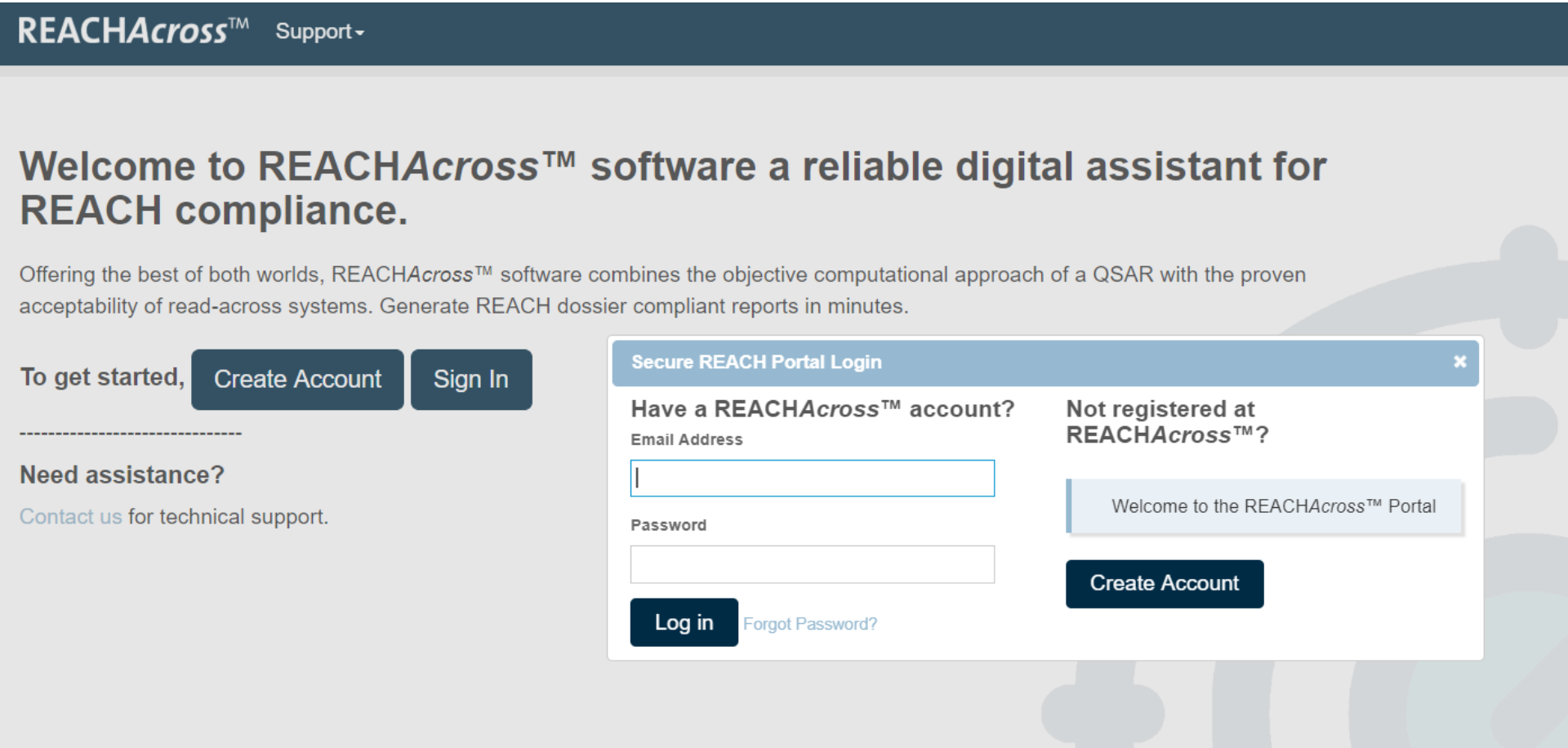
Target Features: 19 categories:

1. Acute Toxicity - Dermal / Inhalation / Oral
2. Hazardous to the aquatic environment - acute / chronic
3. Skin or Respiratory Sensitization / Corrosion / Irritation
4. Serious Eye Damage or Irritation
5. Water contact flammable
6. Substances and Mixtures corrosive to Metals
7. Self-heating Substances and Mixtures
8. Reproductive Toxicity
9. Pyrophoric Solids / Liquids
10. Oxidizing Solids / Gases
11. Organic Peroxides
12. Hazardous to the ozone layer
13. Germ Cell Mutagenicity
14. Gases Under Pressure
15. Flammable Solids / Liquids / Gases
16. Explosives
17. Effects on or via Lactation
18. Carcinogenicity
19. Aspiration Hazard

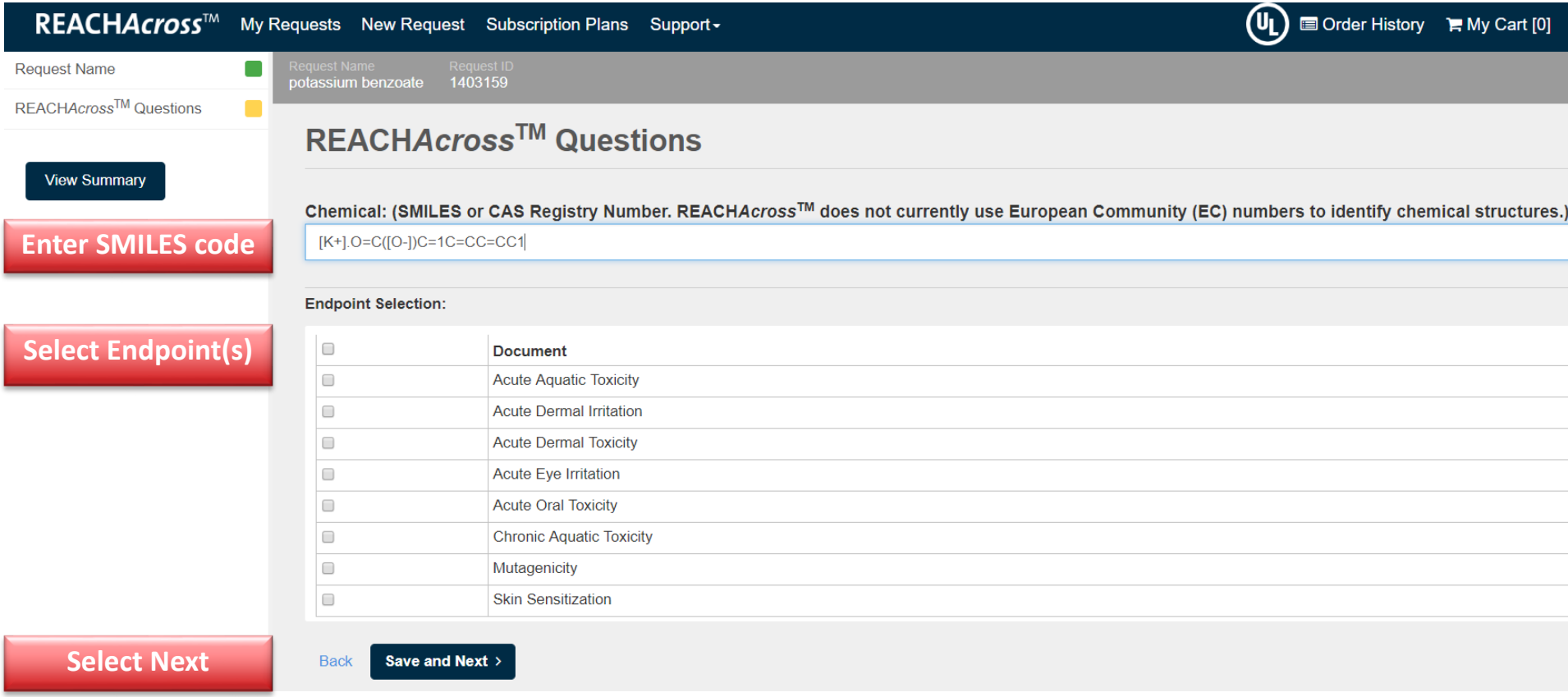
HTTPS://WWW.ULREACHACROSS.COM/



HTTPS://WWW.ULREACHACROSS.COM/



HTTPS://WWW.ULREACHACROSS.COM/



Enter SMILES code

Select Endpoint(s)

Select Next

UL CHEMINFORMATIC SUITE: DATA-FUSION RASAR

Target: potassium;benzoate

[K+].O=C([O-])C=1C=CC=CC1

Hazard Analogues		Non-Hazard Analogues	
96% Similar	azanium;benzoate	93% Similar	2-methylbutan-2-yl benzenecarboperoxoate
94% Similar	sodium;2-fluorobenzoate	88% Similar	benzyl benzoate
94% Similar	tributylstannyl benzoate	88% Similar	(4-methylbenzoyl) 4-methylbenzenecarboperoxoate
93% Similar	benzoyl benzenecarboperoxoate	83% Similar	4-(4-ethylcyclohexyl)benzoic acid
93% Similar	magnesium;dibenzoate	83% Similar	zinc;4-tert-butylbenzoate

Conclusions:

REACH*Across*™ v3.1.4 estimates a 61% probability that [K+].O=C([O-])C=1C=CC=CC1 causes acute eye irritation

Hazard	Probability
Eye Irritation Hazard	61%
Not Eye Irritation Hazard	39%

Potency:

Potency Category	Probability
H319: Eye Irrit. 2	41%
H319: Eye Irrit. 2a	0%
H318: Eye Dam. 1	20%

UL CHEMINFORMATICS SUITE – FALL 2018 U

Accuracy similar to animal test reproducibility

Nine endpoints: adding Acute Inhalation Toxicity

Broad Domain of Applicability

Correctly Identified — **True Positives** / **True Negatives**

UL Chem Suite

Animal Studies

Endpoint	Total	Coverage %	Se%	Sp%	AC%	AC%
Acute Toxicity-Oral	32411	100	94	86	93	94
Acute Toxicity-Dermal	11252	100	89	94	90	88
Acute Toxicity-Inhalation	11369	100	90	91	90	
Skin Sensitization	7670	100	80	96	84	91
Skin Irritation	46331	100	98	75	97	78
Eye Irritation	15760	100	99	70	98	88
Mutagenicity	3703	100	76	92	88	96
Acute Aquatic Toxicity	10541	100	95	94	95	
Chronic Aquatic Toxicity	17295	100	98	66	98	

UL CHEMINFORMATICS SITE IS PEER REVIEWED

Models published in peer reviewed journals

Making Big Sense from Big Data in Toxicology by Read-Across

Thomas Hartung

Johns Hopkins Bloomberg S
University of Konstanz, CA

Global Analysis of Publicly Available Safety Data for 9,801 Substances Registered under REACH from 2008-2014

Thomas Luechtefeld[1], Alexandra Maertens[1], Daniel P. Russo[2], Costanza Rovida[4], Hao Zhu[2,3] and Thomas Hartung[1,4]

[1]Center for Alternatives to Animal Testing (CAAT), Johns Hopkins Bloomberg School of Public Health, Environmental Health Sciences, Baltimore,
Camden, NJ, USA; [3]Departm
Konstanz, Konstanz, Germa

Analysis of Publically Available Skin Sensitization Data from REACH Registrations 2008-2014

Thomas Luechtefeld[1], Alexandra Maertens[1], Daniel P. Russo[2], Costanza Rovida[4], Hao Zhu[2,3] and Thomas Hartung[1,4]

[1]Center for Alternatives to Animal
Health Sciences, Baltimore, MD,
Camden, NJ, USA; [3]Department o
Konstanz, Konstanz, Germany

Analysis of Draize Eye Irritation Testing and its Prediction by Mining Publicly Available 2008-2014 REACH Data

Thomas Luechtefeld[1], Alexandra Maertens[1], Daniel P. Russo[2], Costanza Rovida[4], Hao Zhu[2,3] and Thomas Hartung[1,4]

[1]Center for Alternatives to Animal
Health Sciences, Baltimore, MD,
Camden, NJ, USA; [3]Department of
Konstanz, Konstanz, Germany

Analysis of Public Oral Toxicity Data from REACH Registrations 2008-2014

Thomas Luechtefeld[1], Alexandra Maertens[1], Daniel P. Russo[2], Costanza Rovida[4], Hao Zhu[2,3] and Thomas Hartung[1,4]

[1]Center for Alternatives to Animal Testing (CAAT), Johns Hopkins Bloomberg School of Public Health, Environmental Health Sciences, Baltimore, MD, USA; [2]The Rutgers Center for Computational & Integrative Biology, Rutgers University at Camden, NJ, USA; [3]Department of Chemistry, Rutgers University at Camden, NJ, USA; [4]CAAT-Europe, University of Konstanz, Konstanz, Germany

Toxicology Research

REVIEW

View Article Online
View Journal

Check for updates

Cite this: DOI: 10.1039/c8tx00051d

Big-data and machine learning to revamp computational toxicology and its use in risk assessment

Thomas Luechtefeld,[a] Craig Rowlands[b] and Thomas Hartung

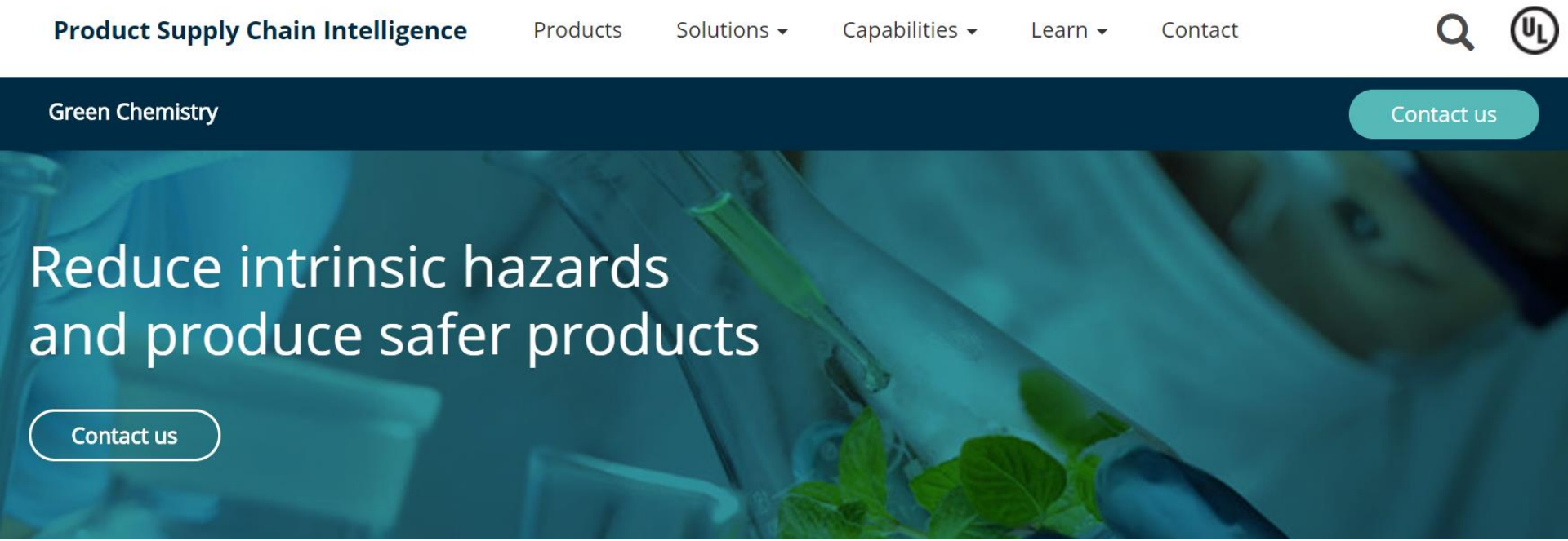
Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility

Tom Luechtefeld[*†], Dan Marsh[†], Craig Rowlands[‡] and Thomas Hartung[*§]

[*]Johns Hopkins University Bloomberg School of Public Health, Center for Alternatives to Animal Testing (CAAT), Baltimore, MD, 21205, TLuechte1@jhu.edu, THartun1@jhu.edu
[†]ToxTrack, Baltimore, MD, 21209, tom@toxtrack.com, danmarsh@toxtrack.com
[‡]Underwriters Laboratories (UL), UL Product Supply Chain Intelligence, Northbrook, IL, 60062, Craig.Rowlands@ul.com
[§]University of Konstanz, CAAT-Europe, Konstanz, Germany, Thomas.Hartung@uni-konstanz.de

CURRENT: HTTPS://WWW.ULREACHACROSS.COM/

SEPTEMBER 2018: HTTPS://PSI.UL.COM/EN//



REFERENCES

- Luechtefeld T, Marsh D, Rowlands C and Hartung T. Machine learning of toxicological big data enables read-across structure activity relationships (RASAR) outperforming animal test reproducibility. Toxicological Sciences, 2018, in press.
- Luechtefeld T, Rowlands C and Hartung T. Big-data and machine learning to revamp computational toxicology and its use in risk assessment. Toxicological Research 2018, in press, doi:10.1039/C8TX00051D.
- Luechtefeld T and Hartung T. Computational Approaches to Chemical Hazard Assessment. ALTEX 2017, 34:459-478.
- Hartung T. Making big sense from big data in toxicology by read-across. ALTEX, 2016, 33:83-93.
- Luechtefeld T, Maertens A, Russo DP, Rovida C, Zhu H and Hartung T. Global analysis of publicly available safety data for 9,801 substances registered under REACH from 2008-2014. ALTEX 2016, 33, 95-109. http://doi.org/10.14573/altex.1510052.
- Luechtefeld T, Maertens A, Russo DP, Rovida C, Zhu H and Hartung T. Analysis of public oral toxicity data from REACH registrations 2008-2014. ALTEX 2016, 33, 111-122. http://dx.doi.org/10.14573/altex.1510054.
- Luechtefeld T, Maertens A, Russo DP, Rovida C, Zhu H and Hartung T. Analysis of Draize eye irritation testing and its prediction by mining publicly available 2008-2014 REACH data. ALTEX 2016, 33, 123-134. http://dx.doi.org/10.14573/altex.1510053.
- Luechtefeld T, Maertens A, Russo DP, Rovida C, Zhu H and Hartung T. Analysis of publically available skin sensitization data from REACH registrations 2008-2014. ALTEX 2016, 33, 135-148. http://dx.doi.org/10.14573/altex.1510055.


Thank you

CHEMINFORMATICS FOR HAZARD DATA NEEDS

- No test material available (e.g., R&D, purchased chemical)
- Animal testing ban (i.e., cosmetics ingredients)
- Reduce animal testing (e.g., sustainability goals)
- Intelligent testing plan design → increase success
- Support risk assessment → weight-of-evidence