



## Method for curating chemical data submitted under REACH

**Aurisano, Nicolo; Fantke, Peter**

*Publication date:*  
2019

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Aurisano, N., & Fantke, P. (2019). *Method for curating chemical data submitted under REACH*. Abstract from 2nd European Exposure Science Strategy Workshop, Bilthoven, Netherlands.

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



European Exposure Science Strategy workshop  
**ISES Europe 2019**

**Topic of the Platform or Poster**

- |   |   |
|---|---|
| <input type="checkbox"/> Data Repositories & Analytics          | <input checked="" type="checkbox"/> Exposure Data Production & Monitoring |
| <input type="checkbox"/> Regulatory Exposure Assessment Science | <input type="checkbox"/> Exposure Assessment Methods & Tools              |
| <input type="checkbox"/> Building Partnerships & Collaboration  | <input type="checkbox"/> Exposure Education & Communication               |

## Method for curating chemical data submitted under REACH

### Data curation, Data quality, Data uncertainty, Physicochemical properties

**Nicolò Aurisano, Peter Fantke**

Quantitative Sustainability Assessment, DTU Management, Technical University of Denmark, Produktionstorvet 424, 2800 Kgs. Lyngby, Denmark

Since the European Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) Regulation entered into force, data for thousands of substances have been submitted for regulatory safety assessment to the publicly available database hosted by the European Chemicals Agency (ECHA). However, to enable a wider dissemination of these data for use in high-throughput risk screening, chemical substitution, and life cycle impact assessment, systematic methods are required for data extraction, harmonization and curation. To address this need, we propose a semi-automated and transparent data curation method based on systematically analyzing data submitted under REACH. Starting from all substances dossiers, we applied a set of aligned data selection and harmonization criteria, structured in a curation tree, to obtain a representative and high-quality mean value per substance for a given physicochemical property. We tested our method on the octanol-water partition coefficient ( $K_{ow}$ ) as widely used property for organic substances, considering both data quality and variability. This allows us to assign quantitative confidence intervals around each mean value per substance. As a result, we provide a database of harmonized and curated mean values per substance-property combination, reported along with their specific confidence intervals and the number and quality of underlying reported data points. Our data curation method can be applied to REACH data for any given set of substances, considering its specific distribution in data quality and variability, and number of data points per data quality class. For an example set of 20 REACH substances, we illustrate how to derive meaningful estimates for physicochemical properties using our semi-automated curation method. With that, our proposed method constitutes a valuable starting point; still, further research is required to extend the method to consider mining information also from text fields, and to adapt it to curate data available in other chemical data sources.