Workshop Report
"Data management and data generation for lifecycle analysis"

PROSUITE WP7 Software Development & WP8 Dissemination

Report of a
Workshop for Large Industry Stakeholders
Held in Berlin, 23 May 2013

GreenDelta (PROSUITE project); Fraunhofer IZM Institute (LCA To Go project)

PROSUITE Deliverable 8.3a Dissemination Workshop

Contents

Introduction

On the 23rd of May 2013, a workshop was organized in Berlin on the topic of data management and collection for life cycle analysis. The conveners were GreenDelta, from the PROSUITE Project www.prosuite.org together with the Fraunhofer Institute from the LCA To Go Project www.lca2go.eu.

This workshop was the follow-up to a workshop organized on 25th of January 2013 entitled "*Sustainability assessment in large enterprises – Quo vadis?*" (report of this workshop can be found here).

The January workshop focused on the challenge of conducting sustainability assessment and addressed the demand of feasible assessment methods, related requirements (including software) and the need to communicate assessment results appropriately. The follow-up workshop in May workshop focused on the issue of data generation that had been raised in January, aiming to explore and reply to ideas and statements from participants, e.g: "*Data are often only available for an operating unit; Which data can be expected from suppliers?* ".

Workshop contact: GreenDelta GmbH - Müllerstr. 135 - 13349 Berlin - Tel: 030/48496030
Andreas Ciroth, ciroth@greendelta.com

Workshop participants

Institution	Name
Henkel AG	Otte
Phoenix Contact	Husemann
P+Z	Wessels
BASF SE	Füssel
Stadler AG	Brandt
Fraunhofer IZM (LCA to Go)	Schischke
GreenDelta (PROSUITE)	Ciroth

Workshop structure and activities

- Theme A: Data for sustainability
- Theme B: Data collection and generation
- Theme C: Data management
- Summarization, conclusions, leave-taking

Minutes of the Workshop Discussions

The notes of the discussions were provided in German by GreenDelta, and translated by Institut Symlog (dissemination partner in PROSUITE) with modifications by GreenDelta.

Generic data

In the supply chain, the origin of materials and primary products used is unclear. The origin of materials and primary products is not documented, or not completely documented.

Therefore it is appropriate to use generic data.

A mapping problem is observed: how can generic data be correctly correlated with own data?

Decisions can be made on the basis of generic data—for example it is possible to identify and exclude critical supply chains (e.g. copper from Congo).

Disadvantage of generic data: in general companies do not have influence on/ access to the providers of generic data – therefore, gaps in the generic data are difficult to close.

Data collection, closing of data gaps

A compulsory data collection sheet would be very helpful. Workshop attendees agreed that this does not yet exist.

This absence is also related to the acceptance and utilization of evaluation and impact assessment methods (carbon footprint/ Tox models / social effects): the collected data must match with the methods used, or the other way round: the generated/collected data determines which methods can be applied in a meaningful way.

Overall, this shows a trade-off between standardization / harmonization of to be collected data on the one side and selecting a good assessment method on the other side. More harmonized data make data collection, communication and reuse of once collected data much. However, fewer assessment and analysis methods can then be applied.

Idea: Treat the data differently depending on characteristics, similarly to the Reach Environmental Release Categories (ERC) illustrated in the table copied below.

ERC	Lifecycle Stage	Level of containment	Intended technical fate of substance	Dispersion of emission sources	Indoor/outdoor	Release promotion during service life
1	Manufacture	Open-closed		Industrial	Indoor	n.a
2	Formulation	Open-closed	Not included into matrix	Industrial	Indoor	n.a.
3	Formulation	Open-closed	Inclusion into/onto matrix	Industrial	Indoor	n.a.
4	End use	Open-closed	Processing aid	Industrial	Indoor	n.a.
5	End use	Open-closed	Inclusion into/onto matrix	Industrial	Indoor	n.a.
6a	End use	Open-closed	Intermediate	Industrial	Indoor	n.a.
6b	End use	Open-closed	Reactive processing aid	Industrial	Indoor	n.a.
6c	End use	Open-closed	Monomers for polymers	Industrial	Indoor	n.a.
6d	End use	Open-closed	Monomers for rubbers or thermosets	Industrial	Indoor	n.a.
7	End use	Closed system	Processing aid	Industrial	Indoor	n.a.
8a	End use	Open-closed	Processing aid	Wide dispersive	Indoor	n.a.
8b	End use	Open-closed	Reaction on use	Wide dispersive	Indoor	n.a.
8c	End use	Open-closed	Inclusion into/onto matrix	Wide dispersive	Indoor	n.a.
8d	End use	Open-closed	Processing aid	Wide dispersive	Outdoor	n.a.
8e	End use	Open-closed	Reaction on use	Wide dispersive	Outdoor	n.a.
8f	End use	Open-closed	Inclusion into/onto matrix	Wide dispersive	Outdoor	n.a.
9a	End use	Closed systems	Processing aid	Wide dispersive	Indoor	n.a.
9b	End use	Closed systems	Processing aid	Wide dispersive	Outdoor	n.a.
10a	Service life	Open	Inclusion into/onto matrix	Wide dispersive	Outdoor	Low
10b	Service life	Open	Inclusion into/onto matrix Removing from matrix	Wide dispersive	Outdoor	High
11a	Service life	Open	Inclusion into/onto matrix	Wide dispersive	Indoor	Low
11b	Service life	Open	Inclusion into/onto matrix Removing from matrix	Wide dispersive	Indoor	High
12a	Service life	Open-closed	Losses from matrix during article processing	Industrial	Indoor	Low
12b	Service life	Open-closed	Losses with matrix during article processing	Industrial	Indoor	High

Table: Appendix R.12-4.2: Use characteristics by the Environmental Release Categories
ECHA Guidance on information requirements and chemical safety assessment Chapter R.12: Use descriptor system, Version 2, March 2010
http://echa.europa.eu/documents/10162/13632/information_requirements_r12_en.pdf

Another aspect: Temporal change of the data, value at time t0, improved value at time t1.

To be effective the rules of data collection must be binding. Currently this is not the case.

[Insertion: The EU Environmental Footprint Approach attempts this; however on the basis of a method that can be called comprehensive, without, in the present version, applying a smart selection or filtering (as used in Reach, see above); similar, but simplified, the World Business Council for Sustainable Development, WBCSD, the "Getting Numbers Right" (GNR) Project, used for instance in the cement industry, www.wbcsdcement.org/pdf/info%20sheet%20on%20GNR.pdf, where also rules for data collection are specified].

Some simple examples for meaningful rules:

- When generic data are used and no explicitly measured ones, then impact values should be higher by default, in order to be on the safe side, just as in a worst case assessment. Also, this rule would motivate for specific data collection and use of better quality data which is currently lacking.
- Assessment of an unknown substance (product/emission): Use a default substance/material and then change as soon as more is known.
- Pay attention to wolf in sheep's clothing molecules/substances: by themselves harmless but critical/dubious in the supply chain or in posterior products/metabolism (eg. Polyethlyenglykol, PEG).
- „Generic modelling" with defaults – without information a very general flow and process is assumed; the "data quality model" shows a relative weak quality. As a minimum goal, the modeling should be more precise, and specific where relevant.

Side theme: Process data can be more easily evaluated if other similar processes are available. To this end categorization of processes is needed: data on energy, combustion processes, commodities, etc. See also the ecoinvent pedigree process and flow types (only as an example, not necessary as a starting point). For these processes it is then possible to estimate all flows, eventually from the basis of a few "lead substances".

Data quality, data quality by data collection.

If information is provided together with a (solid, specified) uncertainty information then this information becomes more representative because it stands not for a single situation but for a wider range of values, e.g., for life cycle information, for different suppliers or for the average emissions of an entire year.

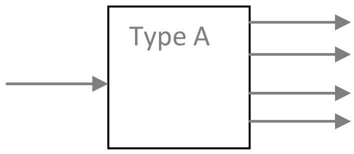
Idea: a data quality evaluation system that is used in parallel for the constitution of a model within a study and that indicates the status of data quality of the study/model. Its application would lead to a systemic quality improvement of the results, just like ecobalance hot spots which help to reveal improvement opportunities (here only related to environmental consequences).

Starting point: „overall database" that includes comprehensive generic information which is then refined, up to regionalized data on a fine scale for instance. For each new evaluation the overall database can be accessed and it can then be refined according to requirements and an evolving information basis. Each new analysis is then built upon the existing information basis, contributing to and refining and extending it further.

Which strategies of „refinement“and data quality improvements are appropriate also depends on the company type, and the production processes used in the company.

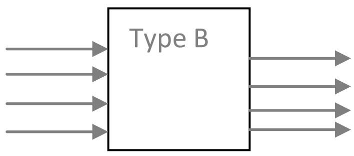
One example:

Company type A (e.g. BASF, simplified): Many different products are produced from few primary products (eg. crude oil)



Complex processes

Company type B (e.g. Henkel, simplified): many primary products and many different products.



Complex supply chain.

→ A categorization of processes can be helpful to focus in the data collection and in the data quality improvement.

Other thoughts

Individual/single studies do not scale. Goal must be to develop a solution that enables scaling. This can hardly be done by one single company or institution.

In order to follow the supply chains, it is useful to consider what companies usually track for products and primary products – in order to build on that basis.

And in the perspective of a unified approach it is also necessary to have a unified and consistent definition of the term “sustainability”.

Conclusion and outlook

Data management and collection for life cycle analyses must consider the quality of the collected data fully in parallel to the process of data collection and creation.. Only then data collection can be efficient and be focusing on relevant areas.

Workshop contact: GreenDelta GmbH - Müllerstr. 135 - 13349 Berlin - Tel: 030/48496030
Andreas Ciroth, ciroth@greendelta.com

Clearly structured, generally accepted, extensive and comprehensive instructions on how to proceed as to efficient data collection of sustainability information within the entire lifecycle of products do not exist at this time.

This workshop gathered some elements that could become components of such an approach. The participants agreed that the ideas generated should be further followed up in order to develop such a shared approach for data collection.