

The International Reference Life Cycle Data System (ILCD) Format – Basic Concepts and Implementation of Life Cycle Impact Assessment (LCIA) Method Data Sets

Marc-Andree Wolf¹, Clemens Döpmeier², Oliver Kusche²

Abstract

In the context of its efforts to facilitate environmental sustainability, the European Commission is promoting and supporting the use of life cycle data and tools through its European Platform on LCA. Objectives of this project are to develop the International Reference Life Cycle Data System (ILCD) Handbook as authoritative guidance on LCA, to contribute key European scope quality data sets via the European Reference Life Cycle Database (ELCD) as well as to implement the ILCD Data Network as infrastructure for LCA data, open to all data developers. Both the ELCD database and the ILCD Data Network rely on the ILCD data format as reference format and for data exchange.

In the initial release of the ILCD data format, only a draft specification for LCIA method data sets had been included, which has been enhanced and finalized in the meantime, addressing feedback from an earlier public stakeholder consultation process and reflecting insights when documenting the ILCD-recommended LCIA methods. The finalized ILCD method dataset specification is now implemented in software and will enable tools to easily import and apply new LCIA methods documented in the ILCD data format.

In this paper, the new LCIA method dataset specification and its corresponding software implementation are presented. The history and idea behind the data format is briefly addressed as well as how the adoption among tools and databases is progressing. The basic structure of the ILCD data format is presented, with the different data set types and their relationships to each other. Then, the structure of the LCIA method data set and its modeling capabilities are explained. Furthermore, technical considerations for tool integration are discussed. Briefly, an exemplary LCIA method instance data set of a draft recommended LCIA method for Europe, foreseen for release by the European Commission's DG JRC, is presented to illustrate the use of the finalized dataset type implementation. Finally, an outlook on future developments is given.

1. Introduction

As part of its efforts to facilitate environmental sustainability, the European Commission is promoting the use of Life Cycle Assessment (LCA) methods in many decision contexts within industry and public policies. This was initiated in the Communication on Integrated Product Policy (IPP), strengthened in the Sustainable Consumption and Production and Sustainable Industrial Policy Action Plan of 2008 (COM(2008) 397). This plan further strengthens the role of life cycle thinking and assessment, reiterating the need for consistent and reliable methods and data. Finally, yet importantly, the Communication on "A resource-efficient Europe – Flagship initiative under the Europe 2020 Strategy" (COM(2011) 21) takes

¹ European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), Via E. Fermi 2479, Ispra (VA), 21027, Italy. Marc-Andree.Wolf@jrc.ec.europa.eu

² Karlsruhe Institute of Technology (KIT), Institute for Applied Computer Science (IAI), Herrmann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

these developments to the next stage. This Communication works with a life cycle approach to reduce the environmental impacts caused by resource use in the entire European Union.

To efficiently support LCA in Europe, an enhanced methodological framework based on LCA ISO standards, the ILCD Handbook, along with appropriate basic life cycle data and tools are provided through its European Platform on LCA (European Communities, 2011). One objective of the platform project is to enhance the availability of consistent and quality assured Life Cycle Inventory (LCI) data sets. This is facilitated by the International Reference Life Cycle Data System (ILCD) data format (JRC/KIT, 2011), which has been designed so serve as reference format and for data exchange, being a further development of the former ELCD format.

The ILCD data format has been released in 2009 and has already seen some adoption among tools and databases in the meantime. So far the implementation of the Life Cycle Impact Assessment (LCIA) method datasets in the ILCD format was based only on a draft specification, which has been enhanced and finalized by March 2011. This has addressed the feedback from an earlier stakeholder consultation process and considers the insight gained when documenting a wide range of LCIA methods as basis for the upcoming ILCD-recommended LCIA methods for Europe. The finalized ILCD method data set specification is now implemented in software and will enable software tools to easily import and apply new LCIA methods documented in the ILCD data format.

2. ILCD Data Format

2.1 Objectives and approach

While a data set format is only a technical means of storing and structuring data set information within a data stream or file, the development of the ILCD format is considered a key step towards better availability and quality as well as lower costs for LCA data sets. The previous situation of LCA data can be characterised as fragmented, with several formats being in use by the different data developers that prevented interchangeability among tools and databases. The preceding works towards a common exchange format, namely SPINE (Carlson/Löfgren/Stehen, 1995), ISO/TS 14048 (ISO, 2002) and EcoSpold (Hedemann/Meinshausen/Frischknecht, 2003) did only partly solve these interoperability problems. They did not result in a wider uptake or, in the case of EcoSpold, were not initially conceived as real tool and database system independent overall exchange format, hence not well meeting the needs of other developers, and also were found limited regarding some important documentation aspects. Some examples of these shortcomings are: lack of unique identifiers, no full multi-language capability, lack of possibility to electronically link reports, flow charts and other pictures, and limitation to document some widely used LCI methods used.

The development of the ILCD data format (that time still under the name ELCD data format) started in 2005, driven by a triple set of needs:

- Need for a data format for the European Reference Life Cycle Database (ELCD) including its use as working format,
- need for a common format to support the exchange (import and export) of the ELCD reference data sets with third party LCA databases and software tools, and
- last but not least the need for a common overall LCA exchange format to be used to exchange LCA data sets among all relevant LCA tools and databases (e.g. for LCA information transfer along supply-chain) and for Data Networks.

Accordingly, the format has been developed starting from ISO/TS 14048 and explicitly analysing a wide range of other formats in use, including the Boustead model³ (United Kingdom), CMLCA tool and databases⁴ (The Netherlands), EcoSpold (Switzerland), DEAM database and TEAM software⁵ (France), GaBi software and databases⁶ (Germany), and KCL EcoData⁷ (Finland). Last not least, the available documentation of the UNEP/SETAC Life Cycle Initiative (UNEP/SETAC, 2002), LCI programme, Taskforce 2 that had worked towards a common data exchange format but stopped without result in early 2005, was a valuable input.

The development has been accompanied by stakeholder and expert meetings and discussions in context of presentations at conferences. Based on the experience made with using the ELCD format for the ELCD database, and feedback of consultations, the format has been further developed towards the ILCD format that was published in 2009.

2.2 Basic Structure

The ILCD format is based on an Internet-aware, linked data approach. Instead of one monolithic data set format for process data sets, the ILCD format provides currently seven data set types which identify different semantic concepts in LCA modelling that are linked together via typed links called global references. These types of data set (concepts) are:

- *Process* for modelling both unit and aggregated processes and result sets. Input and output flows are modelled by global references to other data sets of type *Flow*. Process data sets may optionally contain results of an impact assessment; in this case data sets of type *LCIA Method* will be referenced in a result list.
- *Flow* describes an elementary, product or waste flow. It references one or more *Flow Property* data sets.
- *Flow Property* (quantity) describes physical or other properties of a flow that can be used to quantify it, for example mass or gross calorific value. Each instance references one Unit Group data set.
- *Unit Group* (dimension) describes a group of convertible units and the conversion factors to its reference unit.
- *LCIA Method* describes an LCIA method and its characterisation factors e.g. an impact category like global warming potential or ecotoxicity. The data set can also document an entire LCIA methodology. The data set references one Flow Property data set that identifies the quantity of the characterisation factors and - via the further reference to the Unit group - their dimension.
- *Source* represents an external source of information, such as literature or a database or data format. It can contain a reference to an external file or resource as well. It can reference a contact it is related to.
- *Contact* describes a person or organisation. It can itself again reference another contact, allowing to document hierarchical relationships (e.g. person - working group - organisation).

Each unique data set instance carries a Universally Unique Identifier (UUID) (IETF, 2005) and a version number that is incremented upon changes to the data set. Among each other, data sets are linked by global

³ <http://www.boustead-consulting.co.uk/products.htm>

⁴ <http://www.cmlca.eu/>

⁵ https://www.ecobilan.com/uk_team.php

⁶ <http://www.gabi-software.com/>

⁷ <http://www.kcl.fi/>

references. A global reference in the ILCD format is a structured and typed hyperlink which not uniquely identifies the type of semantic relationship between the linked data sets but also the type (i.e. concept) of the target data set. The referenced data set is further identified by its UUID (and optionally version number) and a Uniform Resource Identifier (URI), that serves as a semantic web compatible Internet wide identifier of the data set.

The linked data and object-oriented approach helps to reduce redundancy and eases maintenance and updating of the content of databases. Furthermore, background data sets those are not subject to frequent change but are commonly used by all data developers (such as flow, flow property and unit group data sets) can be stored and provided at a central location and shared from there, with an. In Figure 1, the relationships among the different data set types are shown.

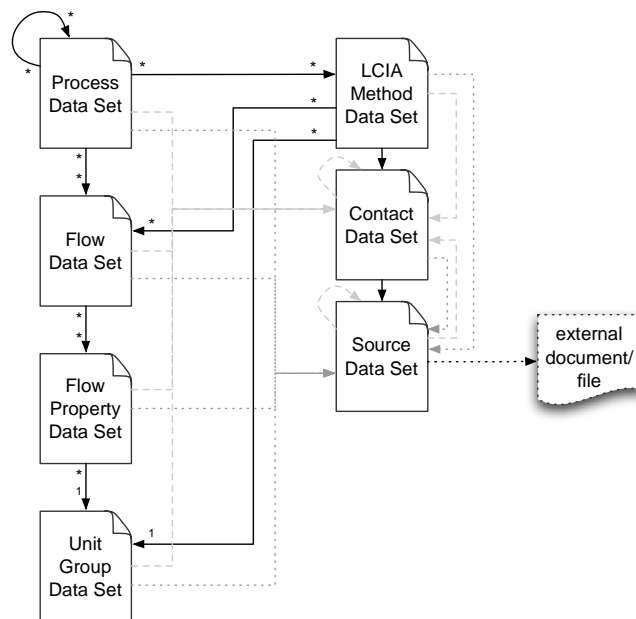


Figure 1
ILCD data set types and relationships

2.3 Adoption

The first application to support the data format was the ILCD data set editor (JRC/KIT, 2011), which is available from the JRC and has been implemented at KIT to provide a basic tool to work with the new format. Next to documenting all data set types, it can perform a supportive check for ILCD documentation compliance of a data set, i.e. whether all required fields are in use and contain valid data. Furthermore, a converter allows converting the data sets to MS Excel spreadsheet documents.

Currently the ILCD format is in use by the following databases (note that this information is non-binding and is based on written statements made by the respective tool and database developers, who remain solely responsible for their developments):

- the ELCD database at the JRC (EU)⁸
- the Malaysian National database at SIRIM⁹

⁸ <http://ict.jrc.ec.europa.eu/assessment/data>

- the Thai National database at MTEC (format extended based on ILCD) 10

The ILCD format is also the technical foundation of the upcoming ILCD Data Network which will use a service API to connect different LCA databases together to a harmonized network of LCA databases.

The French National LCA database that is currently being developed by the French environmental agency ADEME in support of implementing the national labelling of mass market product programme under Grenelle 2 law also uses the ILCD format and ILCD reference elementary flows. Further, industrial associations like PlasticsEurope works and worldsteel will provide additional LCI data sets in ILCD format.

ILCD formatted data sets can be imported and/or exported with the following LCA software tools:

- eBalance software, IKTE (China)¹¹
- EIME software, CODDE/BureauVeritas (France)¹²
- GaBi software and databases (PE International, Germany)
- openLCA software (GreenDelta TC, Germany)¹³
- TEAM software and DEAM database (under implementation), Ecobilan at PriceWaterhouseCoopers (France)

Some of these tools allow to document data sets inside the software, i.e. support working with the format either as basis of the tool or regarding the documentative meta data.

A data format converter (openLCA converter¹⁴) has been developed by GreenDelta TC in support by the UNEP/SETAC Life Cycle Initiative and other sponsors, that supports the conversion of data in the EcoSpold version 1 and 2 format to the ILCD format and vice versa.

With the recent developments of the LCIA method format completion, the availability of a complete set of ILCD-recommended LCIA methods for the whole range of impact categories and the availability of an updated and extended set of ILCD reference elementary flows, it is expected that the attractiveness to work with a common format and increased availability of compatible data sets.

3. LCIA Method Data Set

3.1 Introduction

Past discussions about data formats in LCA have evolved exclusively around process (LCI) data sets. With the implementation of the LCIA method data set, a first comprehensive and dedicated format specification for LCIA methods and the underlying data and models used has been provided and is now being put in use, with the upcoming recommended ILCD2011 LCIA methodology for Europe and its individual LCIA methods.

⁹ <http://lcamalaysia.sirim.my/introduction.asp>

¹⁰ <http://www.thailcidatabase.net/>

¹¹ <http://www.itke.com.cn/cn/product/?CategoryID=1>

¹² <http://www.codde.fr/page.php?rubrique=20>

¹³ <http://www.openlca.org/>

¹⁴ <http://www.openlca.org/Converter.8.0.html>

3.2 Structure

Similarly to a process data set, a LCIA method datasets consists of four major sections:

- *LCIA method information* with metadata describing the LCIA method given its name, classification, quantitative reference, time and geographical and general representativeness;
- *Modelling and validation* with detailed information on the fate model, underlying data, geography and time related aspects of the model etc., together with information about normalisation and weighting, data sources, completeness, validation and compliance declarations;
- *Administrative information* with information required for data set management and administration,
- *Characterisation factors* that lists all flows with their corresponding impact factors according to the respective LCIA method.

A specificity is how geographically differentiated LCIA methods are implemented: the "Location" of e.g. where an emission takes place is an additional attribute that differentiates the very same elementary flow within the LCIA method data set (e.g. "FR" for France). The location of an intervention can then be derived from the "Location" of the unit process data set that has the emission, i.e. also here is no necessity to have country-wise differentiated elementary flows. This allows for geographical differentiation without multiplying the number of elementary flow data sets. Still, it is possible, if wished so, to store geographically differentiated input and output lists in aggregated LCI result process data sets, as the input and output list allows to add an "Location" information (own field) to each input and output. Alternatively, the inputs and outputs can be aggregated without geographical differentiation (as currently wide practice) and (what is new) the LCIA results that were calculated with geographical differentiation can be stored in the aggregated process data set in the respective "LCIA results" section. This allows keeping also the aggregated inventory manageable / lean while still allowing working with geographically differentiated inventories and impact assessment.

For modelling normalisation and weighting, two additional data set types will be used, as normalisation and weighting data can be reused and therefore be stored in a separate data set. These additional data set types have not yet been fully defined and implemented in software.

3.3 Modelling LCIA Methodologies

The XML Schema for the LCIA method data set has been designed in such a way so that it can be used to either model a single LCIA method (e.g. "Photochemical ozone formation; midpoint - human health; van Zelm et al. (2008)") or alternatively describe a group of LCIA methods belonging to the same LCIA methodology (for example, "ILCD2011"). In the latter case, the name field carries exclusively the name of the methodology and exclusively the descriptive data set fields (meta data) are used, while all LCIA method data sets belonging to this methodology are referenced using the *referenceToIncludedMethods* field.

3.4 Example

Figure 2 shows an extract of the draft implementation of a LCIA method for Photochemical ozone formation, midpoint - human health effects as it is rendered in a web browser when using the format's XSLT transformation to HTML, figure 3 the characterisation factors section of the same dataset.

LCIA Method data set: ILCD2011; Photochemical ozone formation; midpoint - human health; POCP; Van Zelm et al (2008) (en)

Table of Contents: [LCIA method information](#) - [Modelling and validation](#) - [Administrative information](#) - [Characterisation factors](#)

LCIA method information

Data set information

Name of LCIA method or methodology	ILCD2011; Photochemical ozone formation; midpoint - human health; POCP; Van Zelm et al (2008)
Belongs to: Name of LCIA methodology/ies	ILCD2011
	ReCiPe2008
Impact category/ies	<i>Photochemical ozone creation</i>
Impact indicator	Expression of the potential contribution to photochemical ozone formation.
General comment	Only for Europe. Includes spatial differentiation.
External documentation / files source (source data set)	Van Zelm et al. (2008)

Quantitative reference

Reference quantity (flow property data set)	C2H4-equivalents
---	----------------------------------

Time representativeness

Reference year	2009
Duration of modelled impact	indefinite

Geographical representativeness

Intervention location	RER
Impact location	RER
Geographical representativeness description	Valid for Europe (spatial differentiation included).

Impact model

LCIA characterisation model(s) name(s)	LOTOS-EUROS
LCIA characterisation model description and included sub-models	Considering a marginal increase in ozone formation, the LOTOS-EUROS spatially differentiated model averages over 14000 grid cells to define European factors.
LCIA characterisation model source (source data set)	Van Zelm et al. (2008)
Considered environmental or other mechanisms along the impact chain	Consideration of VOC and NOx emissions.
LCIA method(ology) flowchart (source data set)	Photochemical ozone figure

Modelling and validation

Use advice for data set	Representative for Europe
-------------------------	---------------------------

LCIA method, normalisation, weighting

Type of data set	<i>Mid-point indicator</i>
LCIA method principle(s)	<i>other</i>
Deviation from LCIA method principle(s)	None
Normalisation included?	false
Weighting included?	false

Data sources

Data sources (source data set)	Van Zelm et al. (2008)
--------------------------------	--

Completeness

Number of basic inventory items covered	132
---	-----

Validation

Type of review	<i>Independent review panel</i>
Scope of review	
Method(s) of review	

Figure 2
Extract of the draft implementation of a LCIA method for
“Photochemical ozone formation, midpoint - human health effects”.

Reference to flow data set	Exchange direction	Mean value (of LCIA characterisation factor)	Data derivation type / status
1,2,3-trimethylbenzene (Mass, kg, Emissions to non-urban air or from high stacks)	Output	2.14	Calculated
1,2,3-trimethylbenzene (Mass, kg, Emissions to urban air close to ground)	Output	2.14	Calculated
1,2,3-trimethylbenzene (Mass, kg, Emissions to air, unspecified (long-term))	Output	2.14	Calculated
1,2,3-trimethylbenzene (Mass, kg, Emissions to air, unspecified)	Output	2.14	Calculated
1,2,4-trimethylbenzene (Mass, kg, Emissions to non-urban air or from high stacks)	Output	2.16	Calculated
1,2,4-trimethylbenzene (Mass, kg, Emissions to urban air close to ground)	Output	2.16	Calculated
1,2,4-trimethylbenzene (Mass, kg, Emissions to air, unspecified (long-term))	Output	2.16	Calculated
1,2,4-trimethylbenzene (Mass, kg, Emissions to air, unspecified)	Output	2.16	Calculated
1,3,5-trimethylbenzene (Mass, kg, Emissions to non-urban air or from high stacks)	Output	2.33	Calculated
1,3,5-trimethylbenzene (Mass, kg, Emissions to urban air close to ground)	Output	2.33	Calculated
1,3,5-trimethylbenzene (Mass, kg, Emissions to air, unspecified (long-term))	Output	2.33	Calculated
1,3,5-trimethylbenzene (Mass, kg, Emissions to air, unspecified)	Output	2.33	Calculated
1-butene (Mass, kg, Emissions to non-urban air or from high stacks)	Output	1.82	Calculated
1-butene (Mass, kg, Emissions to urban air close to ground)	Output	1.82	Calculated

Figure 3
 Characterisation factors section in the draft implementation of a LCIA method for “Photochemical ozone formation, midpoint - human health effects”.

3.5 Technical Considerations

Because the data model of LCIA method data sets can contain large lists of flows and characterization factors, their implementation in XML provide some challenges for tools, which are described in this chapter.

As in the ILCD format the environmental compartment of elementary flows is not modelled in an orthogonal fashion to the flow, but there is rather one elementary flow for each substance-compartment combination, the total number of elementary flows in the ILCD data system nowadays slightly exceeds 40,000 (considering the extension now being published; formerly it was about 19,000). This implementation reflects the explicit feedback of LCA software and database developers, since this is the way elementary flows are implemented in most tools.

As there are some LCIA methods that include thousands of substances in several compartments each, such as toxicity methods, the XML markup for those datasets will consist of very many elements and use a considerable amount of disk space. As most extreme current example, the “ecotoxicity freshwater endpoint” method of the draft ILCD2011 contains about 200,000 XML elements in the characterisationFactors section, with a total data set size on disk of 18.2MB. Processing such large XML documents with conventional methods like DOM (Document Object Model) and SAX (Simple API for XML) requires excessive processing time and impractically large amounts of memory, even if only a few pieces of information from the metadata section are needed. A pull parser implementing the StAX API (Streaming API for XML) (BEA Systems 2003) is a far better parsing tool for such large data sets to achieve significant performance improvements over DOM and SAX parsers, as they provide a cursor based approach where only the relevant elements are read and the processing of the document can be easily terminated once all required information has been extracted, skipping possibly large sections of the document that contain irrelevant data. Furthermore, it reduces the amount of main memory needed for the processing of the data set.

With data sets which are several megabytes in size, it is advantageous and may even be necessary to compress the XML data prior to storage in a database or transmission over a network.

4. Summary and Outlook

The ILCD data format has been developed for electronic documentation and compatible exchange of LCI data and, with the extended specification, now also LCIA data sets. It is implemented in XML as an linked data format and supported by an original ILCD data set editor as well as XSLT stylesheets for conversion to HTML and MS Excel spreadsheets. A set of about 40,000 elementary flows and the now upcoming ILCD2011 complete LCIA methodology with the ILCD-recommended LCIA methods for Europe are now being provided, completing the system.

Several national LCA databases are documented in the ILCD format and the upcoming ILCD Data Network is relying on it as well. The format is implemented as import/export interface and also as internal working/documentation format in a number of key LCA software tools.

The ILCD format can be expected to help overcoming the situation of several, incompatible LCI databases by acting as exchange format and moreover improve documentation practice in LCA, easing access to well documented data for LCA practitioners.

Future development is expected to address the implementation of the normalisation and weighting data sets types and take into account practical experience from working with the format from LCA experts, database developers and software vendors.

Bibliography

BEA Systems, Inc. (2003): Streaming API for XML (JSR-173) for Java Specification.

<http://jcp.org/en/jsr/summary?id=173>

Carlson R., Löfgren G., Steen B. (1995): SPINE, A Relation Database Structure for Life Cycle Assessment. Swedish Environmental Research Institute, IVL-Report B 1227, Göteborg, Sweden

European Commission - Joint Research Centre (JRC) - Institute for Environment and Sustainability, Karlsruhe Institute of Technology (KIT) - Institute for Applied Computer Science (2011): International Reference Life Cycle Data System (ILCD) data format and editor. (Accessible via <http://lct.jrc.ec.europa.eu>)

European Commission - Joint Research Centre (JRC) - Institute for Environment and Sustainability: (2011): International Reference Life Cycle Data System (ILCD) Handbook. Series of guidance documents for good practice in Life Cycle Assessment. First edition 2010 - 2011. Luxembourg. Publications Office of the European Union; 2010 and 2011. (Accessible via <http://lct.jrc.ec.europa.eu>)<http://lct.jrc.ec.europa.eu>

European Commission - Joint Research Centre (JRC) - Institute for Environment and Sustainability (2009): ELCD core database version II.

European Communities (2011): European Platform on Life Cycle Assessment. <http://lct.jrc.ec.europa.eu/eplca/>

Hedemann K., Meinshausen I., Frischknecht R. (2003): ecoinvent 2000 - Documentation EcoSpold; Version 1.2; Swiss Centre for Life Cycle Inventories, Dübendorf / ifu Hamburg GmbH, Hamburg

International Organisation for Standardisation (ISO) (2002): ISO/TS 14048: Environmental management - Life cycle assessment - Data documentation format. International Organisation for Standardisation (ISO), Geneva

Internet Engineering Task Force (IETF) (2005): A Universally Unique Identifier (UUID) URN Namespace. <http://www.ietf.org/rfc/rfc4122.txt>

United Nations Environment Programme (UNEP), Society for Environmental Toxicology and Chemistry (SETAC) (2002): The Life Cycle Initiative. <http://lcinitiative.unep.fr/>